JOURNAL OF VIROLOGY, July 1984, p. 181–191
0022-538X/84/070181-11$02.00/0

Vol. 51, No. 1

Mapping of the Cohesive Overlap of Duck Hepatitis B Virus DNA and of the Site of Initiation of Reverse Transcription

KATHERINE L. MOLNAR-KIMBER,* JESSE W. SUMMERS, AND WILLIAM S. MASON
Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Received 27 October 1983/Accepted 24 February 1984

The hepatitis B-like viruses have a ~3.2 kilobase, partially double-stranded DNA genome that is held in a circular conformation by a cohesive overlap between the 5′ ends of the two strands. In addition, a protein is covalently bound to the 5′ end of the minus strand of virion DNA. The sequence of the cohesive overlap region and its location relative to open reading frames and to the initiation site for minus-strand DNA synthesis, which occurs by reverse transcription of viral RNA, were investigated in duck hepatitis B virus. The 5′ ends of virion DNA were mapped by restriction endonuclease analysis of labeled virion DNA, S1 nuclease digestion, and primer extension, using avian myeloblastosis virus DNA polymerase. The cohesive overlap region was shown to be 69 ± 4 base pairs in length. It contained a 10-base pair inverted repeat in approximately the middle and a 12-base pair direct repeat near each end. The apparent initiation site of reverse transcription was determined by partial sequence analysis of dideoxynucleotide-truncated minus-strand DNA intermediates and comparison of their lengths with the length of a known DNA sequence. It mapped within two to four nucleotides of the 5′ end of the minus strand of virion DNA. The results are consistent with the interpretation that the 5′ end of the minus strand of virion DNA is the origin of reverse transcription and that the protein covalently bound to virion DNA is the primer of reverse transcription.

The hepatitis B viruses, which include human hepatitis B virus (HBV), woodchuck hepatitis virus (WHV), ground squirrel hepatitis virus, and duck hepatitis B virus (DHBV), have a small, circular, partially double-stranded DNA genome (11, 15, 17, 32, 33; see Fig. 1). One strand, the minus strand, is full length (~3.2 kilobases) and has a protein covalently bound to its 5′ end (8, 10, 20). The other strand, the plus strand, varies in length, up to unit length. The virion contains a DNA polymerase that can complete the synthesis of the plus strand (11, 32). The circular conformation of the genome is maintained by a cohesive overlap between the 5′ ends of the two strands (26). Comparison of the sequences of HBV and WHV reveals (i) four open reading frames on the plus strand, which account for all but a small noncoding region (ca. 70 to 100 bases) in the vicinity of the cohesive overlap, and (ii) the presence of a 10 to 12 base pair (bp) inverted repeat located in the approximate region of the cohesive overlap (6, 7, 35).

Although the elucidation of the mechanism of replication of these viruses has been hindered by the lack of a tissue culture system, we have recently isolated virus core-like particles from livers of DHBV-infected ducks (31). These particles, or replication complexes, contain a DNA polymerase which in the presence of added deoxynucleotide triphosphates can elongate both the minus and plus strands. These studies showed that the minus strand is reverse transcribed from an RNA template and the plus strand is copied from the minus strand, similar to the mechanism of replication of retroviruses. The site of initiation of minus-strand synthesis by reverse transcription is unique in DHBV (20) and retroviruses (for a review, see reference 36). Moreover, the circular genome of hepatitis B viruses superficially resembles a type of circular DNA intermediate of retroviruses, i.e., that which contains one copy of the long terminal repeat (LTR). From this view, the cohesive overlap of hepatitis B viruses might be analogous to the LTR of the retroviruses. Initiation of minus-strand synthesis of retroviruses occurs at the right end of the LTR, and initiation of plus-strand DNA synthesis occurs at the left end. In evaluating possible similarities between the cohesive overlap of the hepatitis B viruses and the LTR of the retroviruses, questions arise as to (i) the structure and organization of the cohesive overlap region of the hepatitis viruses and (ii) the location of the site of initiation of reverse transcription in relation to the cohesive overlap region.

To investigate these questions, we mapped the 5′ ends of both strands of DHBV virion DNA and the site of initiation of minus-strand synthesis by reverse transcription of viral RNA. Mapping of the 5′ ends of virion DNA revealed that the cohesive overlap region of DHBV was 69 ± 4 bp in length, in contrast to the estimate of ca. 200 to 300 bp for HBV (3, 4, 26, 27). The cohesive overlap region contained a 10-bp inverted repeat located approximately in the middle, as well as 12-bp direct repeats near the 5′ ends. The site of initiation of minus-strand DNA synthesis mapped within two to four nucleotides of the 5′ end of the minus strand of virion DNA. Comparison of the sequence of DHBV (13) with these results showed that the first ATG start codon in the (5 + 8) open reading frame mapped within the cohesive overlap, approximately 19 bp downstream from the origin of reverse transcription. The (5 + 8) open reading frame appears to code for the major structural protein of DHBV virion and liver cores, which has a size of ca. 35,000 daltons (J. Newbold, W. Mason, and J. Summers, unpublished data) and appears to be analogous to a fusion of the putative X gene products and core gene products of HBV and WHV (13).

MATERIALS AND METHODS

Experimental animals. One-day-old ducklings were purchased from a commercial supplier or hatched from eggs congenitally infected with DHBV (generously provided by Anna O'Connell). Ducklings were bled by cardiac puncture, and sera were assayed for the presence of DHBV DNA by dot hybridization. The nitrocellulose filters were hybridized

* Corresponding author.

with [^{32}P]dTTP-labeled bacteriophage DNA, which was prepared from a recombinant clone containing the entire DHBV genome in a bacteriophage lambda Charon 27 vector (16). Ten- to 14-day-old infected ducklings were sacrificed by injection with pentobarbital sodium and exsanguination. The livers were immediately removed and processed as described below.

Virion DNA isolation. Serum samples (17 to 20 ml) were layered onto 10 to 20% (wt/vol) sucrose gradients containing 150 mM NaCl and 20 mM Tris-hydrochloride (pH 7.4). The virus was pelleted by centrifugation at 20,000 rpm for 16 h at 4°C in a Beckman SW27 rotor. Two alternative methods were then used to purify virus away from the cellular DNA present in serum, presumably due to erythrocyte lysis. In one procedure, virus pellets were suspended with gentle sonication in 150 mM NaCl–20 mM Tris-hydrochloride (pH 7.4), layered onto a 7.0-ml, 15 to 55% (wt/vol) sucrose gradient, overlaid with mineral oil, and subjected to isopycnic centrifugation at 24,000 rpm for 16 h at 4°C in a Beckman SW40 rotor (17). Alternatively, the virus pellet was suspended in 2 ml of CsCl solution (density = 1.18 g/ml) containing 10 mM Tris-hydrochloride (pH 7.4), and virions were banded to equilibrium by centrifugation at 40,000 rpm for 16 to 40 h at 20°C in an SW60 rotor (17). Approximately 20 to 30 fractions were collected, and the position of virions was assayed by the endogenous DNA polymerase reaction (17). The fractions containing virions were diluted at least threefold with 150 mM NaCl–20 mM Tris-hydrochloride (pH 7.4), and the virions were pelleted by centrifugation at 23,000 rpm for 16 h at 4°C in an SW40 rotor. In some cases, it was necessary to perform a second isopynic centrifugation on the virions to further reduce contamination of cellular RNA and DNA. Virion DNA was isolated by proteinase K-sodium dodecyl sulfate (SDS) treatment, phenol-chloroform (1:1) extraction, and ethanol precipitation (17). Contamination of the virion DNA with cellular RNA and DNA was not detected when monitored by 0.7% agarose gel electrophoresis and subsequent staining with ethidium bromide (0.5 μg/ml). Where indicated, the endogenous polymerase reaction (17), using nonradioactive nucleotides, was performed before DNA extraction.

Isolation of viral cores from infected liver. Replication complexes, virus core-like structures that synthesize both plus- and minus-strand DHBV DNA, were extracted from 50 g (wet weight) of liver as described previously (31), with the following modifications (J. Newbold and J. Summers, unpublished data). Briefly, liver tissue was homogenized at 4°C in HCB1 buffer (2 to 3 ml per g of liver) containing 50 mM NaCl, 20 mM Tris-hydrochloride (pH 7.4), 7 mM $MgSO_4$, 0.1% (vol/vol) 2-mercaptoethanol, and 8% (wt/vol) sucrose, and clarified by centrifugation at 10,000 rpm for 20 min at 4°C in a Sorvall HB4 rotor and then by centrifugation at 27,000 rpm for 90 min at 4°C in an SW27 rotor. EDTA (0.5 M) was added to the supernatant to a final concentration of 20 mM. Aliquots (17 ml) were then layered onto 20-ml, 10 to 20% (wt/vol) sucrose gradients containing HCB2 buffer (50 mM NaCl, 20 mM Tris-hydrochloride [pH 7.4], 1 mM EDTA, 0.1% [vol/vol] 2-mercaptoethanol, and 0.01% [vol/vol] Triton X-100), and the cores were pelleted by centrifugation at 20,000 rpm for 16 h at 4°C in an SW27 rotor. Cores were suspended in HCB2 (no sucrose) (1 ml per g of liver), and polyethylene glycol 6000 was added to a final concentration of 5% (wt/vol). The mixture was clarified by centrifugation at 10,000 rpm for 20 min at 4°C in an HB4 rotor. The supernatant was adjusted to 0.55 M NaCl final concentration, additional polyethylene glycol 6000 was added to yield a 15% final concentration, and the solution was incubated on ice for at least 2 h. The cores were then pelleted by centrifugation at 10,000 rpm for 20 min at 4°C in a Sorvall HB4 rotor and suspended in 1 to 2 ml of HCB2 buffer by the use of a Dounce homogenizer (B pestle), and 0.5 to 0.75 ml was layered onto an 11.4-ml, 15 to 30% (wt/vol) sucrose gradient containing HCB2 buffer. Gradients were centrifuged at 40,000 rpm for 2.8 h at 4°C in an SW40 rotor, 0.5-ml fractions were collected and assayed for cores active in the endogenous DNA polymerase reaction (31), and appropriate fractions were pooled.

Gel electrophoresis, transfer, hybridization, and autoradiography. Samples were denatured by (i) glyoxalation as described previously (1), (ii) incubation at 90°C for 30 s in urea buffer (6 M urea, 40 mM NaOH, 4 mM EDTA, 0.03% xylenecyanol, and 0.03% bromophenol blue), or (iii) incubation at 22°C in 90 mM NaOH–6 mM EDTA for 30 min, chilling to 4°C, and threefold dilution in water. DNA fragment length was determined by electrophoresis into (i) 8% polyacrylamide gels containing 25 mM sodium phosphate buffer (pH 6.0) (1), (ii) 6% polyacrylamide gels containing 8 M urea, 50 mM Tris-borate (pH 8.3), and 1 mM EDTA (19), or (iii) 1.5% agarose gels containing 30 mM NaOH and 2 mM EDTA. The gel types and percentages are indicated in the figure legends. The gels were either dried for autoradiography or blotted to a nitrocellulose filter for subsequent hybridization by the techniques of Southern (30), essentially as modified by Wahl et al. (38). Molecular weight markers were *Hae*III-digested bacteriophage ϕX174 replicative form DNA or *Hin*dIII-digested bacteriophage λ DNA. Hybridization of the nitrocellulose filters was performed with [^{32}P]dTTP-labeled probes specific to the plus or minus strand, as previously described (16). Autoradiography was at −80°C with Kodak XAR film and an intensifying screen.

Recombinant DNAs. Recombinant bacteriophage or plasmids containing DHBV DNA were used to generate quantities of DNA adequate for sequence analysis. Clones for these experiments were derived from covalently closed circular DHBV DNA isolated from infected liver rather than from virion DNA, to avoid the possibility that base deletions or alterations at the 5′ ends of virion DNA were generated during cloning. Briefly, frozen liver (0.2 g) from a DHBV-infected duck was homogenized in 10 ml of TES buffer (10 mM Tris-hydrochloride [pH 7.4], 10 mM EDTA, 0.15 M NaCl) in a Dounce homogenizer with a B pestle. Large tissue fragments were allowed to settle, and the supernatant fluid was removed. Nuclei were pelleted by centrifugation at 1,000 × *g* for 5 min, and the nuclear pellet was washed twice with 10 ml of 10 mM Tris-hydrochloride (pH 7.4)–10 mM EDTA–0.3 M NaCl–0.1% (vol/vol) Triton X-100. The washed nuclei were suspended in 5 ml of TES, digested for 30 min at 37°C with 0.5 mg of pronase per ml and 0.1% (wt/vol) SDS, and extracted once with phenol-chloroform (1:1), and the nucleic acids were precipitated with 2 volumes of absolute ethanol. The nucleic acids (100 μg) were loaded onto a preparative 1% low-melting-temperature agarose gel (Seaplaque) containing 0.5 μg of ethidium bromide per ml, 40 mM Tris-acetate (pH 7.4), 20 mM sodium acetate, and 1 mM EDTA, and fractionated by electrophoresis. The region of the gel corresponding to molecular lengths of 2.0 to 2.5 kilobase pairs was excised and melted at 68°C, and the DNA was recovered by phenol extraction and ethanol precipitation (16). DNA from the preparative gel was suspended in 1 ml of water and adjusted to a density of 1.55 g/ml with CsCl and 200 μg of ethidium bromide per ml, overlaid with silicon oil, and centrifuged at 40,000 rpm at 20°C for 48 h in a

Beckman SW56 rotor. Gradient fractions were collected, and the closed circular DNA peak was located by Southern blotting and concentrated by ethanol precipitation. The closed-circular DNA fraction was digested with *Eco*RI (New England Biolabs, Beverly, Mass.) and cloned into the bacteriophage lambda vector Charon 27 (24). DHBV DNA from one of these clones, Charon 27 RI, was subcloned into the *Eco*RI site of pBR322 (this subclone was a gift of C. Rogler). Both Charon 27 and pBR322 recombinant vectors were used in the experiments described below. Bacteriophage and plasmid DNAs were isolated as described by Maniatis et al. (14).

Preparation of ^{32}P-labeled restriction fragments from cloned DNA. Charon 27 DHBV DNA (50 μg) was digested at 37°C for 2 h with 50 U of *Ava*I (New England Biolabs). pBR322 DHBV DNA (5 μg) was digested for 2 h at 37°C with 10 U of *Acc*I (New England Biolabs). Digestion buffers were those recommended by the manufacturer. The samples were deproteinized by extraction with an equal volume of phenol–CCl_3 (1:1) and concentrated by ethanol precipitation. The DNA was dephosphorylated with bacterial alkaline phosphatase and labeled at the 5′ end with [γ-^{32}P]ATP (3,000 Ci/mmol, New England Nuclear Corp., Boston, Mass.) and polynucleotidyl kinase (New England Biolabs) (14). Alternatively, the DNA was labeled at the 3′ end with [α-^{32}P]dTTP (800 Ci/mmol, New England Nuclear) and the Klenow fragment of *Escherichia coli* DNA polymerase I (New England Biolabs) (14). To separate the two labeled ends, *Ava*I and *Acc*I fragments were digested with either *Eco*RI or *Bam*HI restriction endonuclease (New England Biolabs), respectively, using the conditions recommended by the suppliers, and the fragments were separated by electrophoresis into a 1.0%, low-melting-point agarose gel containing 40 mM Tris-acetate (pH 7.5), 20 mM sodium acetate, and 1 mM EDTA. Relevant fragments were extracted (14) and either used in S1 nuclease digestion experiments or cleaved by the Maxam and Gilbert technique (19) with slight modifications (14).

S1 nuclease digestion mapping of the termini of virion DNA. Approximately 4 ng of 5′-end-labeled fragment derived from cloned DNA and 40 ng of virion DNA were mixed, denatured by incubation at 100°C for 3 min in water, and reannealed at 68°C for 2 h in 100 μl of 0.75 M NaCl. Three of four 25-μl aliquots were then incubated at 37°C in 0.5 ml of 30 mM sodium acetate (pH 4)–300 mM NaCl–3 mM $ZnSO_4$–40 μg of single-stranded calf thymus DNA per ml with 300 U of S1 nuclease (Miles Laboratories, Inc., Elkhart, Ind.) for 3, 10, or 30 min. The fourth aliquot was incubated for 30 min as described above, but in the absence of S1 nuclease. The reaction was halted on ice by the addition of 20 μl of 0.5 M EDTA and 10 μl of 10% SDS, and the samples were concentrated by ethanol precipitation.

RESULTS

Approximate size of the cohesive overlap of virion DNA. A previous study suggested an approximate location for the cohesive overlap region of DHBV relative to several restriction endonuclease sites (17). As a first step in the present study, the location, orientation, and approximate size of the cohesive overlap of DHBV virion DNA were determined by mapping the 3′ and 5′ termini of the plus and minus strands relative to the cleavage sites of several restriction endonucleases. Virion DNA was digested with four restriction enzymes, *Eco*RI, *Bgl*II, *Kpn*I, and *Bam*HI, whose positions are shown in Fig. 1A (13). The DNA was then denatured in alkali, and each sample was divided in half and electrophoresed in duplicate into a single alkaline agarose gel. The DNA was then transferred to a nitrocellulose filter, the filter was cut in half, and each half was hybridized with a strand-specific DHBV probe. The detection in the *Kpn*I and *Bam*HI digests of a significant quantity of full-length minus strand (Fig. 2A) and only minute amounts of full-length plus strand (Fig. 2B) showed that in some virions, the synthesis of the plus strand was not completed in the region of these endonuclease sites. The plus-strand results (Fig. 2B) indicated that the endonuclease digestions were themselves completed (*Bam*HI and *Kpn*I will not cut single-stranded DNA). Synthesis of the plus strand proceeded past the *Eco*RI and *Bgl*II sites in the majority of virions, since no full-length DNA was detected with either probe. In agreement with a previous report (18), these data showed that >80% of DHBV virions contained genomes with full-length plus and minus strands.

The results shown in Fig. 2 are summarized in Table 1, which shows the measured distance (fragment size) from each endonuclease site to the 3′ and 5′ termini of both the plus and minus strands of virion DNA. Estimates of the size of the cohesive overlap from these data varied from ~30 to 110 bp. Although only approximate, these results suggested that the cohesive overlap of DHBV was considerably smaller than values of 200 to 300 bp published for HBV (3, 4, 26, 27) and served as a basis for the more precise mapping experiments described below.

Restriction endonuclease mapping and partial sequencing of the region between the *Eco*RI and *Bam*HI sites had revealed *Ava*I and the *Acc*I sites (data not shown). (These restriction endonuclease sites have since been observed in the sequence of the full genome of DHBV by Mandart et al. [13]). Fig. 1A shows the position of the *Ava*I and *Acc*I sites of the complete genome (Mandart et al. [13]). In Fig. 1B, the

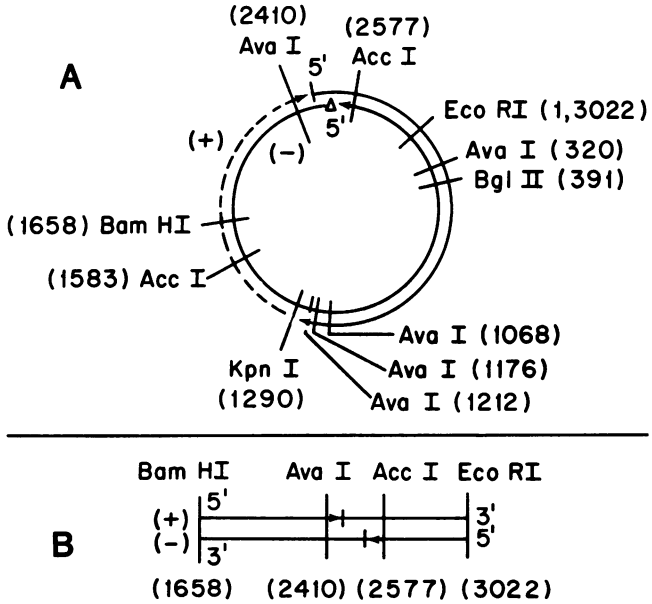

FIG. 1. Restriction enzyme maps of DHBV DNA. (A) Genomic DNA. The positions of the restriction enzyme sites as determined by the sequence of DHBV by Mandart et al. (13) are listed in parentheses. Since the *Eco*RI site was the cloning site for DHBV (13), it is listed as nucleotide 1 and 3022. (B) *Bam*HI-*Eco*RI fragment containing the cohesive overlap region. The approximate sites of the 5′ and 3′ ends of the virion DNA are marked by the short vertical slashes and arrows, respectively.

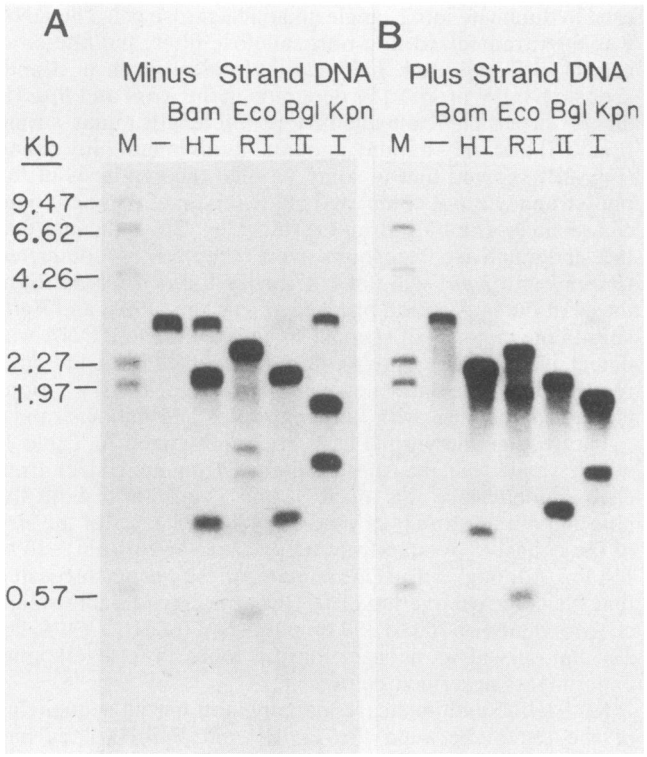

FIG. 2. Agarose gel electrophoresis of restriction enzyme-digested and denatured virion DNA. Virion DNA was digested with four different restriction enzymes (*Bam*HI, *Eco*RI, *Bgl*II, and *Kpn*I) as recommended by the supplier, deproteinized by SDS-proteinase K digestion and phenol-chloroform extraction, ethanol precipitated, and denatured by resuspension in alkali. Samples were electrophoresed in duplicate into a single alkaline 1.5% agarose gel, blotted onto a nitrocellulose filter, and hybridized with strand-specific probes (see the text). (A) Hybridization with a probe that detects minus strands. (B) Hybridization with a probe that detects plus strands. The *Eco*RI digest yielded, in addition to the two expected major bands, several minus-strand bands whose origin has not been determined.

location of the *Acc*I site (2577) and *Ava*I site (2410) relevant to the following experiments are shown in relationship to the approximate location of the 5′ ends of the plus and minus strands of virion DNA.

Mapping the 5′ end of the minus strand of virion DNA. In a first approach, S1 nuclease digestion was used to locate the 5′ end of the minus strand. The *Ava*I to *Eco*RI fragment derived from cloned DHBV DNA (Fig. 1B) was 5′ end labeled by polynucleotidyl kinase at the *Ava*I site and annealed to one of two isolates of virion DNA (see above). Samples were incubated for various times with S1 nuclease to degrade single-strand regions of the labeled fragment. The DNAs were then denatured and electrophoresed into an 8% polyacrylamide gel (Fig. 3A). Three bands were observed. (i) Band *c* (ca. 118 bases [b]) was present in all samples, indicating that it was a contaminating fragment. This was similar but not identical in size to a 108-b fragment predicted by Mandart et. al (13). Examination of the sequence (13) revealed that single base changes at nucleotides 1059 and 1071 would eliminate the 108-b fragment and replace it with a 120-b fragment. (ii) Band *a*, ca. 160 b in size, was released during a short incubation with S1 but disappeared with longer incubation periods, suggesting that although a preferred S1 cleavage site may result in this intermediate, longer S1 nuclease treatment degraded this fragment (5). (iii) Band *b*, ca. 128 b in size, increased in intensity with longer S1 nuclease digestion. The maximum intensity of band *b* in the S1 digestion containing virion DNA isolate 1 (lanes 1, 3, 5, and 7) or virion DNA isolate 2 (lanes 2, 4, 6, and 8) was at 30 and 10 min, respectively. We concluded that band *b* resulted from S1 cleavage of the 5′-labeled *Ava*I to *Eco*RI fragment. To map more precisely the length of fragment *b*, similarly treated samples were electrophoresed into a 6% sequencing gel with the same preparation of *Ava*I to *Eco*RI fragment that had been cleaved by the modified Maxam and Gilbert technique (14, 19). The sequence in the vicinity of the S1 nuclease cleavage site is shown at the right of Fig. 3B. The S1 nuclease generated fragment *b* (lanes 1 to 4) was resolved into four to five S1 bands 128 to 132 b in length, whose 3′ termini were localized to the ACCCC region of the

$$\begin{bmatrix} 5'(+)\ \text{TTAC}\overline{\text{ACCCC}}\text{TCTCC} \\ 3'(-)\ \text{AATGTGGGG} \end{bmatrix}$$

sequence, i.e., to positions 2539 to 2542 on the DHBV genome (Mandart et al. [13]; see Table 2 and Fig. 8). It appears, therefore, that this sequence is complementary to the 5′ terminus of the minus strand of virion DNA. The contaminating band (*c*) had an apparent length of 122 b (see above).

Since DHBV has a relaxed circular DNA genome, the 3′ ends of each strand can be used as primers for displacement synthesis by avian myeloblastosis virus (AMV) DNA polymerase through the cohesive overlap region extending to the 5′ ends of the complementary strand. Restriction endonuclease digestion of the radioactive products of this reaction with *Ava*I should then produce a small, labeled DNA fragment complementary to the 5′ end of the minus strand, as well as a much larger fragment complementary to the 5′ end of the plus strand (Fig. 1). With the use of dideoxynucleotides in the elongation reaction, this technique should provide an independent estimate not only of the distance from the *Ava*I site to the 5′ end of the minus strand, but also of the sequence at the 5′ end, which can then be used to establish the location of the 5′ terminus on the nucleotide map of DHBV (13). Virions were isolated free of contaminating nucleic acids, incomplete plus strands were elongated in an endogenous reaction with unlabeled deoxynucleotides, and the genomic DNA was extracted (see above). It should be noted that >80% of DHBV virions isolated from serum contain complete genomes, and the endogenous reaction on

TABLE 1. Sizing of the cohesive overlap region of DHBV

Restriction enzyme	Fragment length kilobases				Deduced overlap (bp)	
	Minus strand		Plus strand			
	5′	3′	5′	3′	Formula 1[a]	Formula 2[b]
*Bam*HI	0.83	2.10	2.3	0.80	30	110
*Eco*RI	2.5	0.48	0.56	2.6	80	40
*Bgl*II	2.15	0.88	0.94	2.15	60	70
	3.0		3.0			

[a] Calculated by the difference in length of the smallest fragments of the minus and plus strands. By assuming the orientation shown in Fig. 1A, all calculations gave a positive number for the length of the overlap.

[b] Calculated by adding the length of the 5′ fragment of the plus strand and the length of the 5′ fragment of the minus strand and subtracting the full length of the DHBV genome (ca. 3.02 kilobases; reference 13).

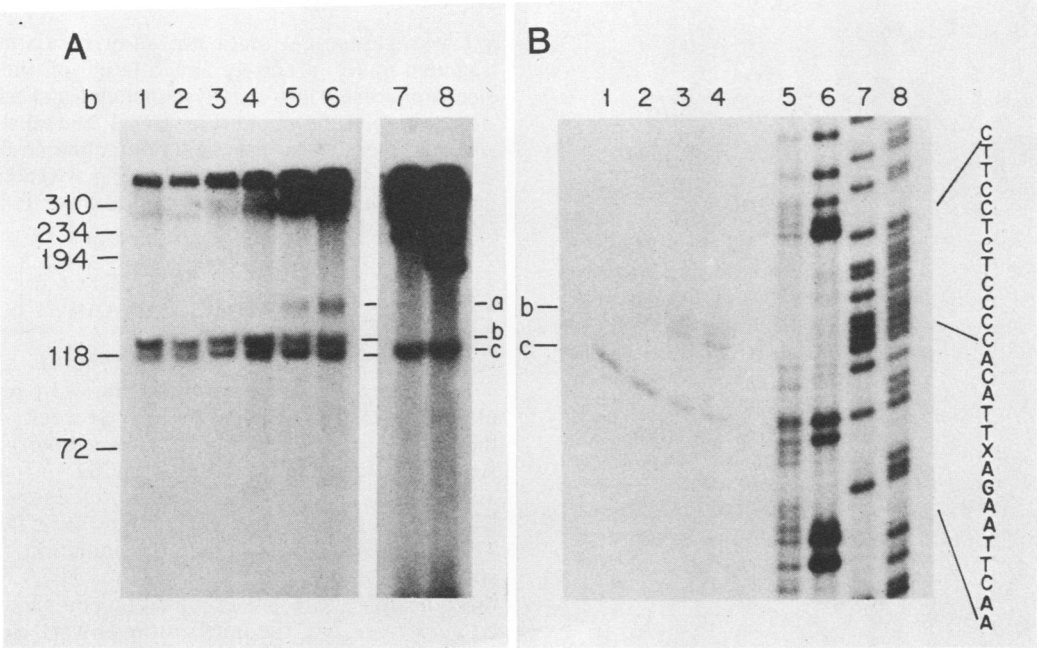

FIG. 3. S1 nuclease mapping of the 5′ end of the minus strand of virion DNA. The *Ava*I to *Eco*RI fragment was 5′ labeled at the *Ava*I site and annealed to virion DNA, and samples were incubated either in buffer or with S1 nuclease for 3, 10, 30 min (see the text). (A) Samples were glyoxalated and electrophoresed into an 8% polyacrylamide gel (1). Lanes 1, 3, 5, and 7 and lanes 2, 4, 6, and 8 contained different batches of virus. Lanes 1 and 2, 30-min digestion; lanes 3 and 4, 10-min digestion; lanes 5 and 6, 3-min digestion; lanes 7 and 8, no enzyme. (B) Samples were denatured in urea buffer and electrophoresed into a 6% polyacrylamide-urea gel at 2,000 V at 45 to 50°C. Lane 1, no enzyme; lane 2, 3-min digestion; lane 3, 10-min digestion; lane 4, 30-min digestion. Lanes 5 to 8 are the *Ava*I to *Eco*RI fragment of cloned DHBV DNA, 5′ labeled at the *Ava*I site and cleaved by the modified Maxam and Gilbert technique (14, 19). Lane 5, dATP plus dGTP; lane 6, dGTP; lane 7, dCTP, lane 8, dTTP plus dCTP.

this virion population increased this proportion of virions with complete genomes to >95%. AMV DNA polymerase was used to extend the 3′ ends of plus and minus strands in the presence of dideoxynucleotides by strand displacement. Cleavage by *Ava*I resulted in a sequencing ladder sufficiently readable to be identified as complementary to the minus strand, which stopped abruptly at the AAGAAXXXX sequence, presumably at the 5′ end (cf. Fig. 4 and 8 and Table 2; see below). Sequences complementary to the 5′ end of the plus strand would be too large (ca. 1,200 b) to be resolved by this gel. Some regions of the ladder were not readable because of termination at more than one base. The technique of primer extension mapped the 5′ end of the minus strand to be 123 b from the *Ava*I cleavage site of the minus strand, i.e., at G of the sequence:

$$\begin{bmatrix} 5'(+)\ \text{AAGAATTACACCCCTCT} \\ 3'(-)\ \text{TTC TTAATG} \end{bmatrix}$$

(see Fig. 1B and 8). In comparison, S1 mapping placed the 5′ end of the minus strand 129 to 132 b from the *Ava*I cleavage site of the plus strand, i.e., at GGGG of the

$$\begin{bmatrix} 5'(+)\ \text{AAGAATTACACCCCTCT} \\ 3'(-)\ \text{TTCTTAATGTGGGG} \end{bmatrix}$$

sequence. (Note that the *Ava*I cleavage site differed by 4 b in the plus and minus strands.) Radiolabeling of fragments smaller than ca. 57 b was rarely observed, consistent with the idea that the 3′ end of the presumably unit-length plus strands that served as the primer for AMV DNA polymerase were downstream from the *Ava*I site (cf. Fig. 4 and 8, 5′ end of plus strand).

Mapping the 5′ end of plus-strand DNA. S1 nuclease digestion was also used as an approach to mapping the 5′ end of the plus strand of virion DNA. The *Acc*I to *Bam*HI fragment (Fig. 1B) was 5′ end labeled by polynucleotidyl kinase at the *Acc*I site, annealed to virion DNA, and digested with S1 nuclease for various times (see above). The bands generated by S1 nuclease digestion for 10 and 30 min were measured to be ca. 110 and 104 b, respectively, by denaturation and electrophoresis into an 8% polyacrylamide gel (Fig. 5A). Similar samples were electrophoresed into a 6% sequencing gel adjacent to the same preparation of *Acc*I to *Bam*HI fragment that had been cleaved by the Maxam and Gilbert technique (14, 19). The bands generated by S1 digestion were 104 to 110 b in length (Fig. 5B); moreover, their 3′ termini corresponded to the GGAAAA region of the

$$\begin{bmatrix} 5'(+)\ \text{TTTTCCA} \\ 3'(-)\ \text{CCGAAAAGGT} \end{bmatrix}$$

sequence of the minus strand (Fig. 5; see Table 2 and Fig. 8).

In the second approach to mapping the 5′ end of the plus strand, virion DNA was digested with *Acc*I, and the Klenow fragment of *E. coli* DNA polymerase I was used to label the 3′ ends in a reaction containing dTTP and dCTP (14). The restriction fragments were then denatured and electrophoresed into an 8% polyacrylamide gel (Fig. 6A). The band observed at ca. 114 b was the size expected for a fragment extending from the 5′ end of the plus strand to the *Acc*I cleavage site (Fig. 1B). The 47-b fragment corresponds to the distance from this *Acc*I site to the 3′ end of the minus strand,

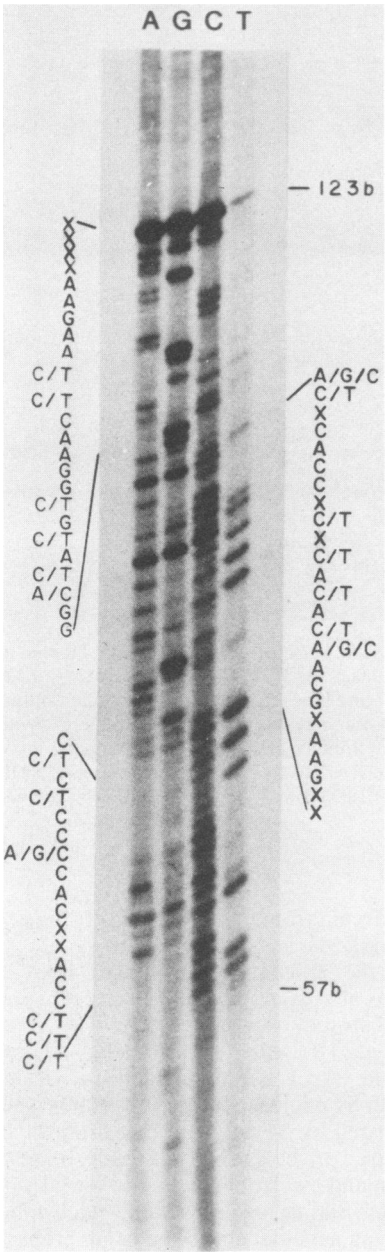

FIG. 4. Mapping the 5′ end of the minus-strand of virion DNA by primer extension. Virion DNA (25 ng per reaction) was incubated in four 5-μl reactions, each containing 0.5 μM [^{32}P]dTTP, 2 μM dCTP, 2 μM dGTP, 2 μM dATP, 50 mM Tris-hydrochloride (pH 8.1), 5 mM $MgCl_2$, 2 mM dithiothreitol, 1 U of AMV DNA polymerase (Bethesda Research Laboratories, Bethesda, Md.), and a single dideoxynucleotide at 0.2 mM for 30 min at 37°C, followed by the addition of 0.05 mM dTTP and a 15-min incubation at 37°C (28). Samples were heated at 68°C for 5 min to inactivate AMV DNA polymerase, diluted in 25 μl of *Ava*I endonuclease buffer, and digested with 1 U of *Ava*I at 37°C for 1 h. Samples were deproteinized, concentrated by ethanol precipitation, denatured in urea buffer, and electrophoresed into a 6% polyacrylamide-urea gel at 2,000 V at 45 to 50°C (see the text).

and its presence suggests that the Klenow fragment was able to cause limited synthesis at this site. Other *Acc*I fragments (ca. 2030 b, 960 b, 880 b) were too large to enter the gel, although they were readily observed in an autoradiography of a 1.5% agarose gel (data not shown). To map the 114-b fragment more precisely, an aliquot of the sample was electrophoresed into a 6% sequencing gel adjacent to the *Acc*I to *Bam*HI fragment (Fig. 1B), 3′ end labeled at the *Acc*I site, and cleaved by the modified technique of Maxam and Gilbert (14, 19). The band at 114 b (Fig. 6A) actually resolved into two plus-strand fragments with 5′ ends located in the GC region of the

$$\begin{bmatrix} 5'(+) & \overline{GC}TTTTCCA \\ 3'(-) & TGCCGAAAAGGT \end{bmatrix}$$

sequence (Fig. 6b; see Table 2 and Fig. 8).

Initiation site of minus-strand synthesis by replication complexes from DHBV-infected liver. We recently demonstrated that (i) the site of initiation of DNA synthesis of the minus strand is unique and (ii) minus-strand DNA found in virions, as well as nascent minus-strand intermediates from infected liver, has a protein covalently bound near the 5′ end (20). The question arises whether the initiation site of minus-strand synthesis is the 5′ end of virion DNA. To investigate this question, we isolated replication complexes from infected duck liver, and the minus strands were elongated in the presence of dideoxynucleotides, resulting in a unique sequencing ladder (Fig. 7, lanes 1 to 4). The lack of labeled fragments smaller than ca. 30 b suggests that initiation occurred already in vivo (20). The length of each of the dideoxynucleotide-truncated fragments was estimated by comparison with an unrelated fragment of cloned DNA that had been end labeled and cleaved by the Maxam and Gilbert technique (Fig. 7, lanes 5 to 8). The *Ava*I cleavage site of the minus strand, previously used for the S1 and primer extension experiments, was located ca. 127 b from the site of initiation of reverse transcription (see Table 2 and Fig. 8).

DISCUSSION

The cohesive overlap region of DHBV has been mapped by determining the 5′ ends of the plus and minus strands of virion DNA. The 5′ end of the plus strand was determined by S1 nuclease digestion (Fig. 5) to be located in the TTTTCC (2470 to 2476) region of the

$$\begin{bmatrix} 5'(+) & \overline{TTTTCC}ATAC \\ 3'(-) & CCGAAAAGGTATG \end{bmatrix}$$

sequence (Table 2 and Fig. 8). Some of the six S1 nuclease bands may be generated by cleavage of the first few bases of the duplex molecule. In fact, it has previously been observed that S1 nuclease digestion of single-stranded DNA annealed to mRNA may result in gradual cleavage of the first few bases at the end of the double-strand region under standard conditions, especially (A+T)-rich sequences (5). In an independent approach involving restriction analysis of virion DNA (Fig. 6), we again mapped the 5′ end of plus strand to the same region, i.e., to G or C (2468, 2469) of the

$$\begin{bmatrix} 5'(+) & \overline{GC}TTTTCCATAC \\ 3'(-) & CCGAAAAGGTATG \end{bmatrix}$$

sequence. This result is consistent with the idea that S1 nuclease gradually digested the ends of the double-stranded DNA. It should be noted, for the experiment shown in Fig. 6, that both a dCTP and a dTTP are required by the Klenow fragment DNA polymerase to complete the 3′ end of the

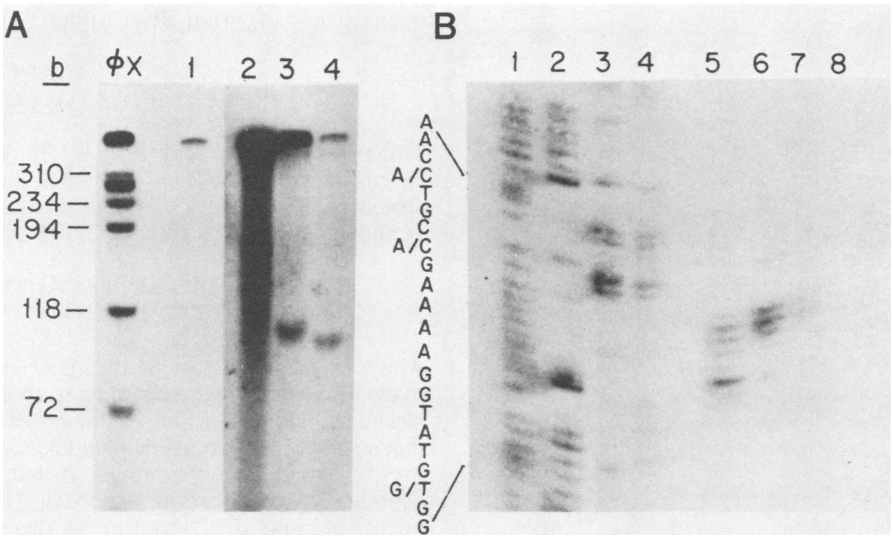

FIG. 5. S1 nuclease mapping of the 5′ end of the plus strand of virion DNA. The *Acc*I to *Bam*HI fragment, which was labeled at its 5′ end at the *Acc*I site, was annealed to virion DNA, and aliquots were incubated at 37°C with S1 nuclease for 3, 10, or 30 min or without nuclease for 30 min (see the text). (A) Samples were denatured by glyoxalation and electrophoresed into an 8% polyacrylamide gel (1). Lane 1, overnight exposure; lanes 2 to 4, 1-month exposure; lanes 1 and 2, no enzyme; lane 3, 10-min digestion; lane 4, 30-min digestion. (B) Samples (lanes 5 to 8) were denatured in alkaline urea buffer at 90°C for 30 s and electrophoresed into a 6% polyacrylamide-urea gel adjacent to the *Acc*I to *Bam*HI fragment derived from cloned DHBV DNA cleaved by the Maxam and Gilbert technique (lanes 1 to 4). Lane 1, dATP and dGTP; lane 2, dGTP; lane 3, dCTP; lane 4, dCTP and dTTP; lane 5, 30-min digestion; lane 6, 10-min digestion; lane 7, 3-min digestion; lane 8, no enzyme.

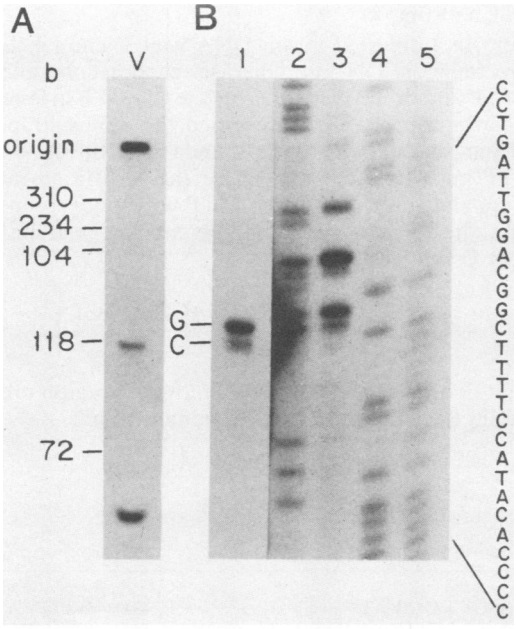

FIG. 6. Location of the 5′ end of the plus strand in relationship to the *Acc*I site. Virion DNA (~10 ng) was digested with 1 U of *Acc*I in 20 μl of restriction buffer at 37°C for 1 h. [^{32}P]dTTP (50 pmol), 200 pmol of dCTP, and 1 U of Klenow fragment was added to the sample, and it was incubated at 22°C for 30 min to label the 3′ ends. The reaction was stopped by the addition of 6 μl of 0.1 M EDTA, 0.5 μl of 10% SDS, and 2 μg of yeast tRNA, and the sample was concentrated by ethanol precipitation (14). (A) The virion DNA was glyoxalated and electrophoresed into an 8% polyacrylamide gel (1). (B) The virion DNA was denatured in urea buffer and electrophoresed into a 6% polyacrylamide-urea gel, adjacent to the *Acc*I to *Bam*HI fragment, which was derived from cloned DNA, 3′ end labeled at the *Acc*I site, and cleaved by the modified technique of Maxam and Gilbert (14, 19). Lane 1, *Acc*I-digested, 3′-labeled virion DNA; lane 2, dATP and dGTP; lane 3, dGTP; lane 4, dCTP; lane 5, dCTP and dTTP.

*Acc*I endonuclease site at 2577

$$\begin{bmatrix} 5'\ \mathrm{GT}\downarrow\mathrm{CT}\quad\mathrm{AC}\ 3' \\ \mathrm{CA}\quad\mathrm{GA}\uparrow\mathrm{TG} \end{bmatrix}$$

and in the reaction, only dTTP, which was polymerized after dCTP, was radioactive. The two bands observed in Fig. 6B should therefore reflect heterogeneity at the 5′ end of the labeled strand, suggesting that the majority of virion DNA contained the 5′ end of the plus strand at one site (dG, nucleotide 2468), but that a minor population contained a 5′ end of the plus strand at a second site, dC, nucleotide 2469. As the 5′ end of the plus strand of virion DNA cannot be labeled with ^{32}P by T4 polynucleotidyl kinase (10), it is also possible that the results are explained by a novel 5′ structure. Primer extension of the plus strand of virion DNA with AMV DNA polymerase (Fig. 4) did not result in dideoxynucleotide-truncated fragments labeled with [^{32}P]dTTP that were less than 57 b in length as measured from the *Ava*I site, indicating that the first [^{32}P]dTTP was incorporated at nucleotide 2471 (GGCTTTTCCATAC). This latter result, together with those above, suggests that at present, we can limit the 3′ and 5′ ends of the elongated plus strand to the GCTT sequence of

$$\begin{bmatrix} 5'(+)\ \overline{\mathrm{GCTT}}\mathrm{TTCCATAC} \\ 3'(-)\ \mathrm{CCGAAAAGGTATG} \end{bmatrix}$$

(Table 2 and Fig. 8).

S1 nuclease digestion (Fig. 3) located the 5′ end of the

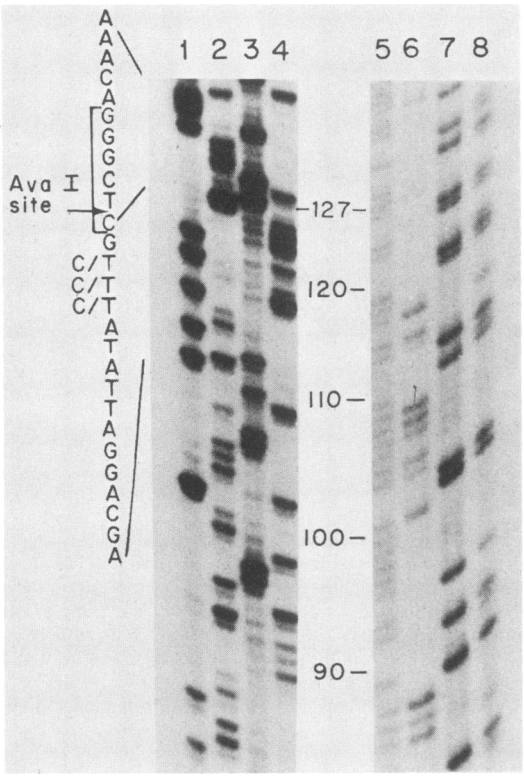

FIG. 7. Location of the site of initiation of minus-strand DNA synthesis. DHBV replication complexes from infected liver were incubated at 37°C for 30 min in four 300-μl reactions containing a 260 nM concentration each of [α-^{32}P]dTTP, dCTP, dGTP, and dATP, 20 mM Tris-hydrochloride (pH 7.4), 7 mM $MgSO_4$, 50 mM NaCl, 0.1% (vol/vol) 2-mercaptoethanol, 100 μg of actinomycin D per ml, and a 26 μM concentration of dideoxynucleotides (20). To minimize premature truncation, all four cold deoxynucleotides (0.5 mM) were added, and samples were incubated for an additional 15 min. Samples were then treated with 0.5% SDS and pronase (0.5 mg, 1 h, 37°C), deproteinized with phenol-chloroform (1:1), and ethanol precipitated with 1 μg of yeast tRNA. Samples (lanes 1 to 4) were denatured in alkaline urea buffer and electrophoresed into a 6% polyacrylamide gel adjacent to a cleaved fragment (14, 19) of 5'-end-labeled cloned DNA of known sequence (lanes 5 to 8). Lane 1, ddATP; lane 2, ddGTP; lane 3, ddCTP; lane 4, ddTTP; lane 5, dATP and dGTP; lane 6, dGTP; lane 7, dCTP; lane 8, dCTP and dTTP.

minus strand of virion DNA to the TGGGG region of the

$$\begin{bmatrix} 5'(+)\ \text{AAGAATTACACCCCTCT} \\ 3'(-)\ \text{TTCTTAATGTGGGG} \end{bmatrix}$$

sequence (Table 2 and Fig. 8). In contrast, the primer extension technique (Fig. 4) mapped the 5' end of the minus strand ca. 123 b from the *Ava*I cleavage site of the minus strand (Fig. 1B), i.e., to G (2537) of

$$\begin{bmatrix} 5'(+)\ \text{AAGAATTACACCCCTCT} \\ 3'(-)\ \text{TTCTTAATG} \end{bmatrix},$$

which is slightly closer to the *Ava*I site than the estimate made by S1 mapping. According to the rationale of the plus-strand mapping given above, this result is somewhat surprising, since primer extension would be expected to progress to the 5' terminus of the minus strand, giving an accurate estimate of its location, whereas S1 nuclease digestion might, because of a low rate of digestion at the ends of double-stranded molecules, locate the 5' end slightly downstream of its real position. This anomaly may be the result of a fragment of the 5'-terminal protein resistant to protease digestion interfering with one or both assays (34). At present, our data suggest that the 5' end of the minus strand of virion DNA maps within GTGGGG of

$$\begin{bmatrix} 5'(+)\ \text{AAGAATTACACCCCTCT} \\ 3'(-)\ \text{TTCTTAATGTGGGG} \end{bmatrix}$$

(Table 2 and Fig. 8).

When *Acc*I-digested virion DNA was 3' end labeled in a Klenow fragment DNA polymerase reaction containing only [^{32}P]dTTP and dCTP (Fig. 6), a species ca. 47 b in length was labeled which probably represented the minus-strand fragment from the *Acc*I site to the 3' end (Fig. 1B, Fig. 8). Since [^{32}P]TTP must be incorporated for the 47-b fragment to be detected, and since dATP and dGTP were not present in the reaction, the 3' end of the 47-b fragment may be at the T after the A or G (underlined) of

$$\begin{bmatrix} 5'(+)\ \text{AAGAATTACACCCCTCT} \\ 3'(-)\ \text{TTCTTAATG} \end{bmatrix}$$

The first G (underlined) falls within the 6-b region predicted to contain the 5' terminus of the minus strand.

TABLE 2. Location of 5' ends of plus and minus strands of DHBV

Method	Figure	Plus strand 5' CGGCTTTTCC[a] 3'	Minus strand 5' GGGGTGTAAT[a] 3'
S1 endonuclease digestion	3 (minus)		+++++
	5 (plus)	++++++	
Primer extension	4	++++	+
*Acc*I-digested, 3'-end-labeled virion DNA	6	++	+ +
Site of initiation of reverse transcription	7		+

[a] Sequences encompassing the 5' ends of plus and minus strands were 2466 to 2475 and 2533 to 2542, respectively.

FIG. 8. Sequence of cohesive overlap region of DHBV. The plus-strand sequence is listed above the minus-strand sequence. The (−) strand origin marks the site of the 5′ end of minus-strand virion DNA and the initiation site of reverse transcription. Twelve base-pair repeats are overlined; a 10-bp hairpin is underlined; the asterisk marks the first ATG in the (5 + 8) open reading frame. "@" marks the termination codon of region 6. We determined the sequence of the region shown by the procedure of Maxam and Gilbert, using cloned DHBV DNA.

The 5′ ends of the minus and plus strands of DHBV map to nucleotides 2537 to 2542 and 2468 to 2471, respectively, on the complete sequence of DHBV (Fig. 8; 13). Thus, the cohesive overlap region of DHBV spans 69 ± 4 b (Fig. 8). This region contains a 10-bp inverted repeat in approximately the middle and 12-bp direct repeats near each end. Although estimates of the size of the cohesive overlap region of HBV, 200 to 300 bp, are much larger (3, 4, 26, 27), it has been postulated from comparison of the sequences and restriction enzyme data that a 10- to 12-bp inverted repeat lies near or within the cohesive overlap regions of both HBV and WHV (3, 4, 6, 7, 21, 26, 27).

Mandart et al. (13) have located three open reading frames in the DHBV sequence in positions analogous to those found in HBV and WHV (6, 7, 35): (i) a long open reading frame, encompassing ca. 80% of the genome, which corresponds to region 6 of HBV and WHV, (ii) an open reading frame ca. 1,200 b long which in HBV and WHV encodes the surface antigen and a pre-S region, and (iii) an open reading frame containing a fusion of the two open reading frames, 5 and 8, of HBV and WHV. For WHV and HBV, region 8 codes for the core antigen and region 5 encodes a product that has not yet been identified. The 5′ end of the minus strand interrupts the fused (5 + 8) open reading frame of DHBV which contains the first ATG at nucleotide 2520 (13, Fig. 8). Although the protein covalently bound to the DNA may interfere with transcription across the 5′ and 3′ ends, transcription of such a message could occur on the covalently closed circular molecules (25, 39) or some other form of viral DNA present in infected liver.

We have sequenced the cohesive overlap region, ca. 400 bp, from a covalently closed circular DHBV molecule that had been isolated from infected duck liver nuclei and cloned into the *Eco*RI site of λ bacteriophage Charon 27. Mandart et al. (13) recently sequenced virion DNA of DHBV which had been cloned into Charon 16A λ bacteriophage. We observed neither deletions nor additions of nucleotides between these two sequences.

The site of initiation of minus-strand DNA synthesis was located ca. 127 nucleotides from the *Ava*I cleavage site of the minus strand (2541). The location of the site of initiation was approximate, since sequencing from the first base was not possible, inasmuch as the first 20 to 30 b were not synthesized in the in vitro system and a possible remnant of the protein covalently bound to the 5′ end of nascent minus strands may have slightly altered the electrophoretic mobility of these DNAs on the sequencing gel. With these reservations, the data mapped the origin of reverse transcription to within two to four nucleotides of the 5′ end of the minus strand of virion DNA. Although the mechanism of initiation of reverse transcription in hepatitis viruses is not known, we have previously shown that a protein is covalently bound to nascent minus-strand DNA intermediates as well as to virion DNA (20). Thus, a protein appears to become covalently bound to nucleotides very early in the synthesis of the minus strand. These results are consistent with the hypothesis that the protein acts as the primer for DNA synthesis and remains covalently bound throughout virus assembly, analogous to the protein primers observed in adenoviruses and bacteriophage φ29 (2, 12, 22).

Although the initiation site of synthesis for the plus strand of DHBV has not been determined, it has been shown that retroviruses initiate plus-strand synthesis at the 5′ end of the LTR (36, 37). This plus-strand fragment, called "strong stop plus," extends the length of the LTR and is made before the elongation of the minus strand is completed. To investigate whether a similar species of plus strand is synthesized in DHBV, DNA in replication complexes from infected liver

was elongated in an endogenous polymerase reaction (20, 31). These complexes incorporated radiolabeled deoxynucleotides into an approximately 68-b species, i.e, the length of the cohesive overlap (data not shown). Synthesis of this 68-b species was inhibited by actinomycin D (data not shown), suggesting that the synthesis occurred in a DNA-dependent reaction and, therefore, that this species is a plus-strand fragment (31). These data are consistent with the notion that the 68-b species is a strong stop plus species of DHBV.

Although the mechanism of priming of the plus and minus strands has not been demonstrated, the similar replication schemes of retroviruses and hepatitis viruses prompted us to investigate whether tRNAs, which act as primers for minus-strand DNA in retroviruses (36) and cauliflower mosaic virus (23) could be primers of DHBV DNA synthesis. In a preliminary sequence search, using software from the Los Alamos data bank, we failed to find complementarity (> 9 of 16) to the 3′-terminal nucleotides of any of 23 mammalian or avian tRNA species (9) adjacent to the presumptive site of initiation of either the plus or minus strand of DHBV. Recent studies of retroviruses suggest that the RNase H activity associated with retroviral reverse transcriptase degrades the RNA template during minus-strand elongation and yields RNA oligomers, one of which occurs just upstream from the first start site for plus-strand synthesis (29). This RNA oligomer, a polypurine sequence, can act as primer for the synthesis of the strong stop plus species. Cauliflower mosaic virus, which appears to have a similar mechanism of replication, also exhibits a polypurine stretch just upstream of presumptive sites of plus-strand initiation (23). In DHBV, no polypurine sequence analogous to that reported in retroviruses (36) or cauliflower mosaic virus (23) was observed near the 5′ end of the plus strand, the presumptive site of initiation in DHBV (Fig. 8). Although HBV and WHV exhibit a polypurine stretch in the plus strand (23), detailed analysis is required to determine whether this area is located adjacent to the site of initiation of plus-strand synthesis.

ACKNOWLEDGMENTS

We thank J. M. Taylor for helpful discussions, F. Galibert and his colleagues (13) for sharing unpublished data on the sequence of DHBV, and the sequencing facility, especially Lynn Gatmaitan, for all their help. The technical assistance of Carol Aldrich and Laura Coates and the secretarial expertise of Helene McGinley are greatly appreciated.

This work was supported in part by Public Health Service grants AI-18641, AI-15166, CA-09035, CA-06927, and RR-05539 from the National Institutes of Health, by grant MV-133 from the American Cancer Society, and by an appropriation from the Commonwealth of Pennsylvania.

LITERATURE CITED

1. **Carmichael, G. G., and G. K. McMaster.** 1980. The analysis of nucleic acids in gels using glyoxal and acridine orange. Methods Enzymol. **65**:380–390.
2. **Challberg, M. D., J. M. Ostrove, and T. J. Kelly, Jr.** 1982. Initiation of adenovirus DNA replication: detection of covalent complexes between nucleotide and the 80-kilodalton terminal protein. J. Virol. **41**:265–270.
3. **Charnay, P., C. Pourcel, A. Louise, A. Fritsch, and P. Toillais.** 1979. Cloning in *Escherichia coli* and physical structure of hepatitis B virion DNA. Proc. Natl. Acad. Sci. U.S.A. **76**:2222–2226.
4. **Delius, H., N. M. Gough, C. H. Cameron, and K. Murray.** 1983 Structure of the hepatitis B virus genome. J. Virol. **47**:337–343.
5. **Favalovo, J., R. Freisman, and R. Kamen.** 1980. Transcription maps of polyoma virus specific RNA: analysis by two dimensional nuclease S1 gel mapping. Methods Enzymol. **65**:718–749.
6. **Galibert, F., E. Mandart, F. Fitonssi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus (subtype ayw) cloned in *E. coli*. Nature (London) **281**:646–650.
7. **Galibert, F., T. Nan Chen, and E. Mandart.** 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. J. Virol. **41**:51–65.
8. **Ganem, D., L. Greenbaum, and H. E. Varmus.** 1982. Virion DNA of ground squirrel hepatitis virus: structural analysis and molecular cloning. J. Virol. **44**:374–383.
9. **Gauss, D. H., and M. Sprinzl.** 1983. Compilation of tRNA sequences. Nucleic Acids Res. **1**:r1–r54.
10. **Gerlich, W. H., and W. S. Robinson.** 1980. Hepatitis B virus contains protein attached to the 5′ terminus of its complete DNA strand. Cell **21**:801–809.
11. **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.
12. **Lichy, J. H., M. S. Horwitz, and J. Hurwitz.** 1981. Formation of a covalent complex between the 80,000 dalton adenovirus terminal protein and 5′-dCMP *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **78**:2678–2682.
13. **Mandart, E., A. Kay, and F. Galibert.** 1984. Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with woodchuck and human hepatitis B virus sequences. J. Virol. **49**:782–792.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Marion, P. L., L. S. Oshiro, D. C. Regnery, G. H. Scullard, and W. S. Robinson.** 1980. A virus of beechey ground squirrels which is related to hepatitis B virus of man. Proc. Natl. Acad. Sci. U.S.A. **77**:2941–2945.
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. U.S.A. **79**:3997–4001.
17. **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.
18. **Mason, W. S., J. M. Taylor, G. Seal, and J. W. Summers.** 1982. An HBV-like virus of domestic ducks, p. 107–116. *In* W. Szmuness, H. S. Alter, and J. E. Maynard (ed.), Viral hepatitis. The Franklin Press, Philadelphia, Pa.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol. **65**:499–559.
20. **Molnar-Kimber, K. L., J. Summers, J. M. Taylor, and W. S. Mason.** 1983. Protein covalently bound to minus-strand DNA intermediates of duck hepatitis B virus. J. Virol. **45**:165–172.
21. **Pasek, M., T. Goto, W. Gilbert, B. Zink, H. Schaller, P. MacKay, G. Leadbetter, and K. Murray.** 1979. Hepatitis B virus genes and their expression in *E. coli*. Nature (London) **282**:575–579.
22. **Penalva, M. A., and M. Salas.** 1982. Initiation of phage phi 29 DNA replication in vitro: formation of a covalent complex between the terminal protein, p3, and 5′-dAMP. Proc. Natl. Acad. Sci. U.S.A. **79**:5522–5526.
23. **Pfeiffer, P., and T. Hohn.** 1983. Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. Cell **33**:781–789.
24. **Rimm, D. L., D. Horness, J. Kucera, and F. R. Blattner.** 1980. Construction of coliphage lambda Charon vectors with *Bam*HI cloning sites. Gene **12**:301–309.
25. **Ruiz-Opazo, N., P. R. Chakrabarty, and D. A. Shafritz.** 1982. Evidence of supercoiled hepatitis B DNA in chimpanzee liver and serum dane particles: possible implications in persistent HBV infection. Cell **29**:129–138.
26. **Sattler, F., and W. S. Robinson.** 1979. Hepatitis B viral DNA molecules have cohesive ends. J. Virol. **32**:226–233.
27. **Siddiqui, A., F. Sattler, and W. S. Robinson.** 1979. Restriction endonuclease cleavage map and location of unique features of the DNA of hepatitis B virus, subtype ad w_2. Proc. Natl. Acad.

Sci. U.S.A. **76**:4664–4668.
28. **Smith, A. J.** 1980. DNA sequence analysis by primed synthesis. Methods Enzymol. **65**:560–580.
29. **Smith, J. K., A. Cywinski, and J. M. Taylor.** 1984. Initiation of plus-strand DNA synthesis during reverse transcription of an avian retrovirus genome. J. Virol. **49**:200–204.
30. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
31. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of a B-like virus by reverse transcription of an RNA intermediate. Cell **29**:403–415.
32. **Summers, J., A. O'Connell, and I. Millman.** 1975. Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. Proc. Natl. Acad. Sci. U.S.A. **72**:4597–4601.
33. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc. Natl. Acad. Sci. U.S.A. **75**:4533–4537.
34. **Tamanoi, F., and B. W. Stillman.** 1982. Function of the adenovirus terminal protein in the initiation of DNA replication. Proc. Natl. Acad. Sci. U.S.A. **79**:2221–2225.
35. **Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter.** 1981. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes, p. 50-70. *In* B. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.
36. **Varmus, H. E.** 1983. Retroviruses, p. 411–503. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.
37. **Varmus, H. E., S. Heasley, H.-J. Kung, H. Oppermann, 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.
38. **Wahl, E. M., M. Stem, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. U.S.A. **76**:3683–3687.
39. **Weiser, B., D. Ganem, C. Seeger, and H. E. Varmus.** 1983. Closed circular viral DNA and asymmetrical heterogeneous forms in livers from animals infected with ground squirrel hepatitis virus. J. Virol. **48**:1–9.