Changing the Site of Initiation of Plus-Strand DNA Synthesis Inhibits the Subsequent Template Switch during Replication of a Hepadnavirus

DANIEL D. LOEB,* RU TIAN, KAROLYN J. GULYA, AND AMY E. QUALEY

McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, Madison, Wisconsin

Received 20 February 1998/Accepted 15 May 1998

Unique to hepadnavirus reverse transcription is the process of primer translocation, in which the RNA primer for the initiation of plus-strand DNA synthesis is generated at one site on its template, DR1, and is moved to a new site, DR2. For duck hepatitis B virus (DHBV), DR2 is located within 50 nucleotides of the 5' end of the minus-strand DNA template. When the synthesis of plus-strand DNA proceeds to the 5' terminus of the minus strand, the 3' end of the minus strand becomes the template for DNA synthesis. This switch in templates circularizes the nascent genome and is required for the genesis of the relaxed circular form of the DNA and the mature capsid. Maturation of the capsid is a prerequisite for virus egress. We have analyzed a series of DHBV variants in which plus-strand DNA synthesis was initiated from a new position relative to the 5' end of the template. For these variants, the subsequent circularization was inhibited. We found that when the number of nucleotides between the site of initiation of plus-strand DNA synthesis and the 5' end of its template was restored to 54 nucleotides, circularization was substantially restored. These results mean that the process of circularization is influenced by the earlier steps in DNA replication. This sensitivity is consistent with the notion that this region of the nascent genome is in a dynamic structure that is crucial for successful DNA replication.

The process of hepadnavirus genome replication, like other DNA synthetic pathways that employ reverse transcription, is complex (for a review, see reference 3). Underlying this complexity is a process called template switching. At three times during hepadnavirus DNA replication, the strand of DNA undergoing synthesis switches templates: once during minus-strand synthesis (9, 13, 18, 19) and twice during plus-strand synthesis (6, 7, 11, 15, 17, 20). The two template switches during the synthesis of plus-strand DNA, primer translocation and circularization, are required for the genesis of a relaxed circular genome. During primer translocation, the RNA primer, which is initially located at the 3' end of the minus-strand DNA template as a consequence of the final RNase H cleavage (8), switches templates from its site of generation to DR2, which is near the 5' end of the template (6) (Fig. 1, parts 1 and 2). After the elongation of plus-strand DNA from DR2 to the 5' end of the minus-strand DNA, a template switch to the 3' end of the minus strand occurs (7, 11, 15, 20) (Fig. 1, parts 3 and 4). This template switch circularizes the nascent genome and permits resumption of plus-strand DNA synthesis (Fig. 1, part 5). Elongation of circularized plus-strand DNA will ultimately yield relaxed circular DNA, the major form of the genome found in infectious stocks of wild-type virus. Failure to carry out primer translocation can result in the formation of either a duplex linear genome, due to *in situ* priming of plus-strand DNA synthesis (17) (Fig. 1, part 6), or accumulation of a full-length minus strand, due to lack of initiation of plus-strand synthesis (4) (Fig. 1, part 1 or 2). Inhibition of circularization results in the accumulation of the immediate precursor to the circularized genome, a full copy of the minus-

strand DNA with a 50-nucleotide segment of plus-strand DNA (4, 11) (Fig. 1, part 3).

There are sequences distinct from DR1/DR2 and the 5' and 3' copies of *r* that are required for the primer translocation and circularization. These sequences have been mapped to three different regions of the genome. One *cis*-acting sequence is near the 3' end of the minus-strand template, and the region that contains this element is called 3E (4). Viruses containing mutations within region 3E have been shown to have an increase in the level of plus-strand priming *in situ* (17) and/or a deficiency in circularization (4). The second *cis*-acting sequence is located near the middle of the minus-strand template, within a region named M (4, 12). The third *cis*-acting sequence, which is within a region called 5E, is located 3' of DR2 on the minus-strand template (4). Viruses with mutations within either region M or 5E appear to suffer a deficiency in, at least, primer utilization at DR2, leading to a reduction in the level of plus-strand DNA. How the sequences within 5E, M, and 3E contribute to template switching is not understood.

Plus strands of DNA that are primed from DR2 must subsequently circularize to yield relaxed circular genomes. The minus-strand template contains a terminal redundancy of seven or eight nucleotides (nt) (7). These sequences serve as the donor and acceptor templates for the switch. The terminal redundancies, although important, are not sufficient for circularization (11). An element within region 3E also is required for circularization (4). We have found that the process of circularization is sensitive to where plus-strand DNA synthesis initiates on the template. When plus-strand DNA synthesis began at a position more internal on its template than normal, circularization was inhibited. Thus, in addition to the need for specific *cis*-acting elements for circularization, the number of nucleotides between these elements and the *cis*-acting elements required for the preceding step in plus-strand DNA synthesis, primer translocation, is critical for DNA replication.

* Corresponding author. Mailing address: McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, 1400 University Ave., Madison, WI 53706. Phone: (608) 262-1260. Fax: (608) 262-2824. E-mail: loeb@oncology.wisc.edu.

FIG. 1. Model for the synthesis of hepadnavirus plus-strand DNA. The model begins at the completion of minus-strand DNA synthesis. In part 1, the thick black line represents minus-strand DNA with the 3' end indicated. The shaded oval at the 5' end of the minus-strand DNA is the P protein, which is covalently attached. The positions of DR2 and DR1 on the minus-strand DNA are represented by rectangles. The sequence of DR is within each rectangle. The oligoribonucleotide annealed to the 3' end of minus-strand DNA is generated as the consequence of the final RNase H cleavage at the completion of minus-strand DNA synthesis. This moiety is derived from the pregenomic RNA and will serve as the primer for the synthesis of plus-strand DNA. In part 2, for most of the minus-strand templates, the plus-strand primer translocates and anneals to DR2. The duplex is 12 nts. Part 3 shows initiation of plus-strand DNA synthesis and elongation of the 5' end of the minus-strand DNA template. 5'r and 3'r represent the terminal redundancy of the minus strand. The sequence of the redundancy is indicated. In part 4, a template switch circularizes the nascent genome. The r sequence of the nascent plus-strand DNA anneals with the 3'r of the minus-strand template. Part 5 shows that elongation of plus-strand DNA will yield the relaxed circular (RC) form of the genome. Part 6 shows in situ priming of plus-strand DNA. For a minority of templates, the plus-strand primer initiates synthesis at its site of generation, the 3' end of the template. Elongation will ultimately produce the duplex linear (DL) form of the genome.

From these findings we infer that the mechanisms of primer translocation and circularization have features in common.

MATERIALS AND METHODS

Molecular clones. All molecular clones were derived from duck hepatitis B virus (DHBV) type 3 (16). The molecular clones used for the expression of the mutant viruses contained approximately 1.5 copies of the DHBV genome (1.5-mer plasmids). These DNA clones contained a *Bam*HI (nt 1658)-to-*Eco*RI (nt 3021) DHBV fragment upstream of a monomeric copy of the genome permuted at the *Eco*RI site. The plasmid vector was pBS(-). Individual mutations were introduced into plasmids containing the *Bam*HI (nt 1658)-to-*Eco*RI (nt 3021) DHBV fragment, which are called 0.5-mer plasmids. The 0.5-mer plasmids were converted to 1.5-mer plasmids by inserting a copy of the genome into the *Nco*I site at nt 2350. All 1.5-mer plasmids contained a frameshift mutation near the 5' end of the P gene, thus making the resultant viruses null for expression of functional P protein. Throughout this paper, the prefix p in a name indicates a DNA plasmid used for expression of a virus (e.g., pG304 expresses virus G304). A virus containing only the frameshift mutation in the P gene served as the comparison in all of the experiments and is called the wild type. The name of this plasmid is p503. To permit replication of the P-protein null viruses, a P-protein expression plasmid named pG308 was used in our transfection experiments. The pregenomic RNA expressed from G308 is defective for DNA synthesis because of a deletion within the RNA packaging signal, epsilon, and a deletion of the 3' copy of DR1.

(i) **Viruses G304, G361, and G404.** Plasmids pG304 and pG361 have a deletion of the copy of DR2 that is encoded within the pregenome. In addition, pG304 and pG361 contained a copy of the DR sequence (TACACCCTCTC), introduced at either the *Msc*I site at nt 2372 or the *Ava*I site at nt 2410, respectively. For virus G361, the *Ava*I termini were filled in with the Klenow fragment of DNA polymerase I. A single copy of the DR sequence was introduced into the 0.5-mer plasmids that were precursors to plasmids pG304 and pG361 by the linker tailing procedure (5). The resultant 0.5-mer plasmids then were converted into 1.5-mers as described above for virus expression. Plasmid pG404 was derived from plasmid pG361. A *Sac*II site was introduced at nt 2453 to nt 2458 by using a site-specific, oligonucleotide-directed mutagenesis procedure, and then the *Sac*II-to-*Afl*III (nt 2526) fragment from this plasmid was removed. This resulted in the deletion of the DHBV sequence between nt 2453 and nt 2526.

(ii) **Viruses 379, G371-6, G371-11, G371-15, G367, and G422.** Plasmid p379 was derived from p503 and has the DHBV sequence from nt 2503 to nt 2515 replaced with the sequence ATAATACGTATCC, which introduces a unique *Sna*BI site. Plasmids pG371-6 and pG371-15 were derived from p379 by introducing the 72-nt *Hae*III fragment from bacteriophage ϕ X174 into the *Sna*BI site.

They differ from each other by the orientation of the insert. pG371-11 was derived from p379 by introducing the 118-nt *Hae*III fragment from ϕ X174 into the *Sna*BI site. pG367 was made by placing the 72-nt *Hae*III fragment from ϕ X174 into the filled-in *Tth*111I site (nt 906) of p503. pG422 was derived from pG371-11 by deleting a 188-nt *Bam*HI fragment (nt 1470 to nt 1658).

(iii) **Viruses G410, G435, G434, G423, and G469.** pG410 was derived from p379 by inserting an 88-nt fragment from pGEM-5Zf(-) into the *Sna*BI site. The 88-nt fragment, which is from the polylinker region of pGEM-5Zf(-), was generated by cleaving with *Aar*II and *Nsi*I. The ends of this fragment were made flush by the action of T4 DNA polymerase. Molecular clones containing both orientations of the insert were isolated and studied. pG435, pG434, pG423, and pG469 were made by deleting 31, 40, 53, and 70 nts, respectively, from the pGEM-5Zf(-) insert within pG410.

(iv) **Virus T247.** pT247 was derived from p379 by inserting a 12-nt *Sma*I linker, TCCCCCGGGGGA, into the *Sna*BI site.

Cell cultures and isolation of viral DNA. LMH cells were used in all experiments. Culturing of cells and DNA transfections were performed as previously described (9). Viral DNA was isolated from cytoplasmic capsids 3 days after transfection as previously described (2). Southern blotting was performed as previously described (17).

Primer extension analyses. Typically 100 pg to 1 ng of viral DNA was processed for use in three separate primer extension reactions. First, each viral DNA was mixed with approximately 500 pg to 1 ng of a 0.5-mer plasmid that was digested at two positions. The signal level measured for the plasmid DNA served as an internal standard in our analyses to permit the comparison of levels of 5' ends of viral DNA determined in different primer extension reactions. Next, the samples were treated with alkali to hydrolyze the RNA primer. DNA samples were then recovered by ethanol precipitation for use in three separate primer extension reactions as previously described (11). The oligonucleotide used to detect and measure the level of the plus-strand DNA initiated from a translocated primer and elongated to at least the 5' end of the minus-strand DNA has a sequence complementary to DHBV nts 2520 to 2537 and is called 2537⁻. This primer extension reaction will detect plus strands that have failed to circularize and plus strands that have successfully circularized. To detect and measure the level of plus-strand DNA that has successfully circularized, one of two oligonucleotides was used. Oligonucleotide 2567⁻ has a sequence complementary to DHBV nts 2547 to 2567, while oligonucleotide 2579⁻ has a sequence complementary to DHBV nts 2560 to 2579. Measurement of the level of minus-strand DNA by primer extension was carried out with an oligonucleotide derived from positions 2425 to 2447. The conditions for the primer extension reactions and the electrophoresis of these reactions were as previously described (11). Quantitation of autoradiographic images was performed with a PhosphorImager from Molecular Dynamics. For each primer extension reaction, a value was derived by dividing the level of 5' ends of viral DNA by the level of 5' ends measured from

FIG. 2. Schematic representations of variant viruses and the extents to which they carry out circularization. The diagram at the top represents the full-length minus-strand DNA. P protein (grey oval) is linked to the 5' end of the minus-strand DNA. Below, the 5' end of the minus-strand DNA of the wild-type (WT) virus and each mutant virus is represented. The black rectangle represents DR2, which is at nt 2477. In viruses G304, G361, and G404, DR2 has been deleted and an ectopic copy of DR has been inserted (white rectangle) at either position 2372 or 2411. For the remainder of the variants, the position of DR2 (black rectangle) relative to the 5' end of the minus-strand DNA has changed as a consequence of a non-DHBV sequence insertion (grey line). The size of each insertion is indicated. The name of the variant virus is indicated on left. G371-6 and G371-15 differ in the orientation of the same insert. The average degree of circularization, ($\pm$ the standard deviation) of each variant virus, as a percentage of that of the wild-type virus, is indicated on right, and the number (n) of determinations for each variant virus is indicated on the far right. n.d., not done.

the internal standard, the digested plasmid DNA. This value represented the normalized level of viral DNA detected in a given primer extension reaction. To determine the level of circularization for a virus, the normalized level of plus-strand DNA measured before circularization was divided into the normalized level of plus-strand DNA measured after circularization. The number derived from this calculation is called the relative circularization value. In each experiment, a relative circularization value was determined for the wild-type comparison, virus 503. The relative circularization value for each mutant virus was expressed as a percentage of the relative circularization value of virus 503. The extent of circularization for each mutant virus, except G371-15, was determined multiple times. The average value with its standard deviation for the extent of circularization for the variant viruses is indicated in Fig. 2.

RESULTS

The initial aim of this study was to determine whether the DR sequence (TACACCCCTCTC), when introduced at an ectopic location within the minus-strand template, could serve as an acceptor site for the plus-strand primer during its translocation. To this end, we deleted DR2, the normal acceptor site, and then introduced the 12-nt DR sequence at one of two

nearby positions on the minus-strand template. In a previous study, we had shown that deletion of DR2 resulted in inhibition of primer translocation (10). DR2 is located at DHBV nts 2477 to 2488, which is within 50 nts of the 5' end of the minus-strand DNA (DHBV nt 2537). We introduced into the mutant virus with a deletion of DR2 a copy of the DR sequence either 66 or 104 nts 3' to the normal position of DR2 within the minus-strand DNA template to generate mutant viruses G361 and G304, respectively (Fig. 2). Because the deletion of DR2 and the insertion of copies of DR were located within the P gene, mutant forms of the P protein would be expressed. Because this would complicate the analysis, we ablated the expression of P protein from these variant viruses by introducing a frameshift mutation within the 5' end of the P gene. Therefore, to study the abilities of these variant viruses to translocate the plus-strand primer, we cotransfected their molecular clones with an expression plasmid for the P protein into cells that support viral replication, LMH. Several days later, newly replicated viral DNA was purified from cytoplasmic capsids from the transfected cells. Initially, we performed a Southern blot analysis to evaluate the abilities of the variant viruses to translocate the plus-strand primer to the ectopic site. The presence of relaxed circular DNA would be evidence of primer translocation to the ectopic site. As can be seen in Fig. 3A, the variant viruses, G304 and G361, made very little, if any, relaxed circular DNA. The profiles of replicative intermediates on the Southern blot of the G304 and G361 viruses and the ΔDR2 virus were similar. There was little, if any, relaxed circular DNA and an accumulation of species migrating at the position of the full-length minus-strand DNA. The lack of relaxed circular DNA in the Southern blot analysis was consistent with the plus-strand primers not translocating to the introduced copies of DR. To determine more definitively whether primers were translocating to the ectopic site, we carried out a primer extension analysis to detect plus-strand DNA initiating from the introduced copies of DR. The oligonucleotide used in these analyses annealed to a position on the plus-strand DNA that is synthesized before the template switch that circularizes the nascent genome. If viruses G304 and G361 were defective in primer translocation, then no signal would be detected in the primer extension analysis. To permit the lack of a signal at DR2 to be informative, we incorporated two features into the experimental design. First, the level of minus-strand DNA within each sample of each virus was determined by using primer extension. Lack of detection of a 5' end of plus-strand DNA could only be meaningful if the sample contained an abundant level of minus-strand DNA, the template for the synthesis of plus-strand DNA. Second, an internal standard consisting of cloned viral DNA cleaved with a restriction endonuclease was added to each viral sample prior to the primer extension analysis. The internal standard would demonstrate that the primer extension reactions were successful. The results of primer extension of the minus-strand DNA for viruses G304 and G361 indicated abundant levels of minus-strand DNA and that the 5' ends of these DNAs mapped to the expected positions (Fig. 3B). The results of the primer extension analysis of plus-strand DNA indicated the presence of a 5' end of plus-strand DNA at the position of the introduced DR (nt 2410) for virus G361 (Fig. 3C, lane 3). A lower level of 5' ends of plus-strand DNA was detected at the site of the introduced DR (nt 2372) for virus G304 (Fig. 3C, lane 9). The presence of 5' ends of plus-strand DNA at the location of the introduced DR for virus G361 indicated that the primer translocation had occurred. To determine the extent of primer translocation for virus G361 relative to the wild type, the ratio of the level of 5' ends of plus-strand DNA to the level of 5'

FIG. 3. Translocation of the plus-strand primer to an ectopic site and subsequent inhibition of circularization. (A) Southern blot analysis indicates that viruses G304 and G361 synthesize reduced levels of relaxed circular DNA. Viral DNA was isolated from cytoplasmic capsids from transfected cells, RC, DL, and SS represent the mobilities of the relaxed circular, duplex linear, and minus-strand single-stranded forms of the genome, respectively. The hybridization probe was genomic length and minus-strand specific. WT, wild type. (B) Determination of the level and position of the 5' end of minus-strand DNA for viruses G304 and G361. The oligonucleotides used in primer extension and in the sequence reaction was derived from DHBV coordinates 2425 to 2447. The band representing the 5' end of minus-strand DNA for viruses G304 and G361 is 12 nts shorter than that of the wild type because DR2 has been deleted. (C) Translocation of the plus-strand primer to the ectopic site for virus G361. Primer extension analysis was done with oligonucleotide 2537⁺ to measure the level of plus-strand DNA synthesized before circularization. The 5' ends of G361 DNA (lane 8) and G304 DNA (lane 9) yield primer extension products that are 66 and 104 nts longer, respectively, than the wild-type DNA (lane 10). The position of the internal standard present in each lane is indicated on the left. A DHBV sequencing ladder was generated with the same oligonucleotide that was used in the primer extension reactions, 2537⁺. (D) Virus G361 is inhibited for circularization. Primer extension analysis was done with oligonucleotide 2567⁺ to measure the level of plus-strand DNA synthesized shortly after circularization. The positions of the primer extension products that represent the 5' end of the wild-type virus (lane 13), virus G361 (lane 11), and the internal standards are indicated on left. A DHBV sequencing ladder was generated with the same oligonucleotide that was used in primer extension reactions, 2567⁺.

ends of minus-strand DNA was calculated. The value for virus G361 was $52\% \pm 5\%$ ($n = 4$) of the value for the wild-type virus. This comparison indicates that the extent of primer translocation for virus G361 was within 50% of the extent of the wild-type virus.

Although a relatively high level of plus-strand DNA was initiated from the introduced DR for virus G361, mature relaxed circular DNA was not synthesized at a comparable level. Clearly, a problem during the elongation of plus-strand DNA was occurring. An inability to carry out circularization could account for the lack of relaxed circular DNA for virus G361. To determine whether the circularization step was inhibited, we performed an additional primer extension analysis. This

strategy is depicted in Fig. 4. In this analysis, we used an oligonucleotide that anneals to plus-strand DNA 30 nts after circularization. If circularization were inhibited, a lower level of 5' termini of plus-strand DNA would be detected. The analysis of virus G361 indicated a much lower level of 5' ends of plus-strand DNA originating from the ectopic DR than was seen in the previous primer extension analysis (Fig. 3D, lane 11), which meant that the circularization step was inhibited. To quantify the extent of this inhibition, the ratio of the level of 5' ends of plus-strand DNA detected after circularization to the level detected before circularization was calculated. The value determined for virus G361 was 8% of the value determined for our wild-type DHBV comparison, on average (Fig. 2).

FIG. 4. Strategy for measuring circularization. Shown are two replicative intermediates, one competent (A) and one inhibited (B) for circularization. Each replicative intermediate has a full-length minus-strand DNA (black line) with the P protein (grey oval) linked to the 5' end. The RNA primer for plus-strand synthesis is annealed to the DR2 site on the minus strand. Plus-strand DNA is represented as a grey line. (A) Successful circularization. Plus-strand DNA that initiated from DR2 has circularized and elongated. Oligonucleotides B and A, represented as short arrows, will anneal to plus-strand DNA before and after the circularization point, respectively. In a primer extension analysis, both oligonucleotides will produce a signal mapping to DR2. (B) Failure to circularize. Primer extension with oligonucleotide B will yield a signal at DR2, while that with oligonucleotide A will not.

Insertion of non-DHBV sequences within the 49 nt between the 5' ends of plus- and minus-strand DNAs inhibited circularization. The number of nucleotides from the acceptor site of the plus-strand primer to the 5' end of minus-strand DNA is 49 for DHBV. For virus G361, this number increased to 115. We considered three explanations for why virus G361 was inhibited for circularization. (i) Circularization is optimal when the number of nucleotides within this region of the genome is 49. This model predicts that inserting non-DHBV sequences of various lengths between DR2 and DR1 would inhibit circularization. (ii) A relationship between two determinants is met when this region contains 49 nts, and this relationship can be fulfilled by sizes other than 49 nts. This explanation predicts that some insertions between DR2 and DR1 would inhibit circularization while other insertions would not. With a large enough data set, the periodicity, if any, of this relationship would be determined. (iii) Inhibition of circularization for virus G361 is unrelated to the number of nucleotides between the 5' ends of plus- and minus-strand DNAs. According to this explanation, insertions between DR2 and DR1 would not necessarily inhibit circularization.

To begin to discriminate among these possibilities, we analyzed three additional variants of DHBV that contained an insertion of either 72 (viruses G371-6 and G371-15) or 118 (virus G371-11) nts between DHBV coordinates 2509 and 2510, which is between DR2 (nt 2488) and the 5' end of the minus-strand DNA (nt 2537) (Fig. 2). The insertions were derived by *Hae*III endonuclease digestion of bacteriophage ϕ X174 DNA. These insertions were introduced into a molecular clone, p379, in which the sequence between positions 2503 and 2515 was changed and a *Sna*BI site was created. DHBVs containing this 13-nt substitution made wild-type proportions

of relaxed circular DNA as determined by Southern blotting, indicating that the 13-nt substitution had not inhibited circularization (data not shown). The ϕ X174 DNA was inserted into the *Sna*BI site. Molecular clones containing both orientations of the 72-nt fragment (named pG371-6 and pG371-15) were studied. Because the insertions were within the 3' end of the P gene and therefore would yield mutant forms of the P protein, a strategy similar to that described for viruses G304 and G361 was used to express only the wild-type form of the P protein in the experiment. Southern blot analysis of viral DNA isolated from cytoplasmic capsids of cells expressing the variant viruses showed very little, if any, relaxed circular DNA (Fig. 5A). This result indicated that a step(s) early in the process of plus-strand DNA synthesis, such as circularization or primer translocation, was inhibited. To determine whether circularization was inhibited, we employed our primer extension strategy to measure the levels of plus-strand DNA before and after circularization. But first, we determined the position of the 5' end of the minus-strand DNA. This was necessary because the position of the 5' end of the minus strand contributes to determining the number of nucleotides between the 5' ends of the two DNA strands. The results of this analysis indicated that the 5' end of the minus-strand DNA for the variant viruses was at the normal position at nt 2537 (data not shown). Next, primer extension analysis of the plus-strand DNA was performed. By using the oligonucleotide that anneals to plus-strand DNA immediately before the circularization, we detected 5' ends of DNA at DR2 (Fig. 5, lanes 1, 3, 5, and 7), indicating that primer translocation had occurred. When the oligonucleotide that anneals to plus-strand DNA after circularization was used, the levels of 5' ends at DR2 were much lower (Fig. 5, lanes 2, 4, 6, and 8). Viruses G371-6 (72 nts, orientation 1), G371-15 (72 nts, orientation 2), and G371-11 (118 nts) carried out the template switch to circularize the genome at 2, 5, and 12% of the level of wild-type DHBV, respectively (Fig. 2). The insertion of ϕ X174 DNA made the respective genomes greater than the wild type in size, which could contribute to the inhibition of circularization. To control for this variable, two additional analyses were performed. First, the 72-nt fragment of ϕ X174 DNA was inserted at another position of the genome (nt 902). This variant virus, named G367, made normal proportions of relaxed circular DNA in a Southern blot analysis (data not shown). Second, a derivative of virus G371-11 was analyzed. This variant, named G422, contains a 188-nt sequence deletion (from nt 1470 to nt 1658) that was previously shown to be dispensable for the synthesis of plus-strand DNA (4). On the basis of primer extension, virus G422 was inhibited for relaxed circular DNA synthesis to a similar degree as virus G371-11 (data not shown). From these results, we conclude that viruses G371-6, G371-11, and G371-15 were inhibited for circularization because the number of nucleotides between the 5' ends of plus- and minus-strand DNAs had been increased to 121 or 167 nts. Based upon the results from the analysis of four variants (G361, G371-6, G371-11, and G371-15), we concluded that adding 66 or more nts between the 5' ends of plus- and minus-strand DNAs inhibited circularization.

We wondered whether any increase in the number of nucleotides within this region of the genome would inhibit circularization. If not, would the degree of inhibition correlate with the size of the insertion? To address these issues, we analyzed a series of variant viruses with progressively smaller insertions derived from the same parental sequence. An 88-nt fragment from the polylinker region of plasmid pGEM-5Zf(+) was inserted into the *Sna*BI site of virus 379. This new variant virus was named G410. A series derivatives of G410 containing

FIG. 5. Insertion of a foreign sequence between DR2 and DR1 leads to inhibition of circularization. Viruses G371-6, G371-11, and G371-15 contain 72-, 118-, and 72-nt insertions of bacteriophage ϕ X174 DNA, respectively. (A) Southern blot analysis indicates that variant viruses synthesize reduced levels of relaxed circular DNA. Viral DNA was isolated from cytoplasmic capsids from transfected cells. RC, DL, and SS represent the mobilities of the relaxed circular, duplex linear, and minus-strand single-stranded forms of the genome, respectively. The hybridization probe was genomic length and minus strand specific. wt, wild type. (B) Viruses G371-6, G371-11, and G371-15 are partially inhibited for circularization. Primer extension was used to measure the levels of plus-strand DNA synthesized before and after circularization. To the left of the band that represents the internal standard is an asterisk. To the left of the bands that represent the 5' end of plus-strand DNA synthesized immediately before the circularization point is the letter B. To the right of the bands that represent the 5' end of plus-strand DNA synthesized after circularization is the letter A. The sequence ladders labeled A and B at the bottom were synthesized with the same oligonucleotides used in the respective primer extension reactions.

progressively larger deletions were made by using the unique restriction endonuclease sites within the inserted polylinker sequence. This set of variant viruses had 88 (G410), 57 (G435), 48 (G434), 35 (G423), and 18 (G469) nts inserted between DR2 and DR1. A variant with a 12-nt insertion (T247) was also made by inserting a *Sma*I linker into the *Sna*BI site. Southern blot analysis of this set of viruses indicated that each variant, with the exception of virus G434 (48 nts added), made levels of DNA similar to the level of the wild type (Fig. 6A). The reason for the reduced level of DNA for virus G434 (48 nts added) is not clear. In addition, the Southern blotting results indicated a reduction in the proportion of relaxed circular DNA. This reduction was most striking for the mutant viruses with the largest insertions. Next, we performed our primer extension analyses to quantify the extent of inhibition of circularization for each variant. First, we sought to determine whether the 5' terminus of minus-strand DNA was at the normal position. Primer extension analysis of minus-strand DNA of each variant indicated that the 5' end of the minus strand was predominantly, if not solely, at the wild-type position, nt 2537 (data not shown). Next, primer extension analysis with the oligonucleotide that anneals to the plus-strand DNA before circularization indicated the presence of 5' ends at DR2 (Fig. 6B), indicating that the plus-strand primer was translocating. By using primer extension analysis to detect plus-strand DNA that initiated from DR2 and circularized (Fig. 6C), we calculated the degree of circularization by mutant virus G410 (88 nts added) to be 8% of that of the wild-type virus (Fig. 2). This level was comparable to the level measured for G361 (9%; 66 nts added), G371-6 (2%; 72 nts added), and G371-11 (12%; 118 nts added). Virus G435, which contains a 57-nt insert, circularized at 12.5% of the level of the wild type, also comparable to what was previously measured (Fig. 2). With G434, which contains a 48-nt insertion, it was difficult to obtain a value for efficiency of circularization, as a consequence of its reduced level of DNA. The two attempts to measure the efficiency of circularization of virus G434 that were successful yielded an average value of 13% of the wild-type level (Fig. 2). The remaining viruses were not severely inhibited in their circularization. Circularization occurred at 50, 62, and 93% of the wild-type level for viruses G423 (35 nts added), G469 (18 nts added), and T247 (12 nts added), respectively (Fig. 2).

Restoring the number of nucleotides between the 5' ends of plus- and minus-strand DNAs to 54 from 115 partially restored circularization. Virus G361, which has 115 nts between the 5' ends of its two strands, was inhibited for circularization. We asked whether circularization could be restored by deleting sequences within virus G361 to restore the number of nucleotides between the 5' termini of the two strands to a value closer to that found in the wild-type virus. A derivative of G361, named G404, that contained a deletion of the sequence between DHBV coordinates 2453 and 2526 was made (Fig. 2). The deletion within virus G404 not only restored the number of nucleotides to a value closer to that of the wild type (54 versus 49) but also resulted in the sequence between the respective 5' termini being different than the sequence found in the wild-type virus. Analysis of viral DNA replicated within cytoplasmic capsids indicated that virus G404 carried out circularization to a greater degree than virus G361. This conclusion is evident from a Southern blot analyses (Fig. 7A, lanes 2 and 3). Whereas virus G361 synthesizes little, if any, relaxed circular DNA, virus G404 made a readily detectable level of relaxed circular DNA, indicating that the circularization function had been partially regained. Quantitation of the level of the relaxed circular, duplex linear, and single-stranded forms of viral DNA indicated that the proportion of relaxed circular

FIG. 6. Insertion of sequences of different lengths between DR2 and DR1 leads to different degrees of inhibition of circularization. Virus G410 has an 88-nt fragment, which contains multiple restriction endonuclease recognition sites, inserted between DR2 and DR1. Viruses G435, G434, G423, and G469 were derived from G410 by deleting sequences within the 88-nts and therefore contain net insertions, relative to the wild type (wt), of 57, 48, 35, and 18 nts, respectively. (A) Southern blot analysis of the variant viruses. Viral DNA was isolated from cytoplasmic capsids from transfected cells. RC, DL, and SS represent the mobilities of the relaxed circular, duplex linear, and minus-strand single-stranded forms of the genome, respectively. The hybridization probe was genomic length and minus strand specific. (B) Primer extension analysis with oligonucleotide 2537⁻ to measure the level of plus-strand DNA synthesized before circularization. Viruses are referred to by the sizes of their insertions. The position of the band that represents the internal standard is labeled IS. The sequence ladder was generated with oligonucleotide 2537⁻. (C) Primer extension analysis with oligonucleotide 2579⁻ to measure the level of plus-strand DNA synthesized after circularization. The sequence ladder was generated with oligonucleotide 2579⁻.

DNA of these major forms was less for virus G404 than for the wild-type virus but greater than that for virus G361. In these comparisons, the relaxed circular form was, on average, 71% of the total of the three major forms for the wild-type virus, while this value was 29% for virus G404. This value was less than 1% for virus G361. These results indicated that although substantial restoration of the circularization function was measured for virus G404, the restoration was not complete. Primer extension analysis of virus G404 indicated that the minus-strand DNA 5' end was located at nt 2537, which is the wild-type position (Fig. 7B). Primer extension analysis with the oligonucleotide that annealed to plus-strand DNA after circularization showed a predominant 5' end at the site of the introduced DR sequence for virus G404 (Fig. 7C, lane 7). The ratio of the level of 5' ends of plus-strand DNA measured with an oligonucleotide that anneals to a position that is synthesized after circularization to the level of 5' ends of minus-strand DNA was calculated for virus G404 and the wild type. The ratio for virus G404 was then expressed as a percentage of the ratio determined for the wild type. The average percentage for G404 with respect to the wild type, from four independent comparisons, was $28\% \pm 6\%$. The measurements from the primer extension and Southern blot analyses are consistent with each other. Overall, these results indicated that when the number of nucleotides between the 5' ends of the plus and minus strands

was reduced from 115 to 54, substantial restoration of the circularization function resulted.

DISCUSSION

We have shown for DHBV that when the acceptor site for primer translocation, and therefore the site of initiation of plus-strand DNA synthesis, is moved to a position that is more internal than normal on the minus strand, the subsequent template switch that circularizes the nascent genome is inhibited. We found that, for a variant virus inhibited in this manner, circularization could be partially regained by restoring the number of nucleotides in the nascent plus-strand DNA to a value similar to that of the wild type (54 versus 49). It should be noted that as a consequence of restoring the number of nucleotides to 54, the sequence of the plus strand synthesized before circularization in this variant virus would be different than in the wild type. Overall, these results indicate that the circularization process is sensitive to the position of the 5' end of plus-strand DNA relative to the site of the template switch, the minus-strand terminal redundancy, *r*. These results lead to the new recognition of an additional *cis*-acting requirement for circularization and provide a greater appreciation of the complex nature of the mechanism of this step of replication.

We have found that perturbing the genome between the

FIG. 7. Deletion of 61 nts from virus G361 partially restores circularization. The 5' end of plus-strand DNA is 115 nts from the 5' end of minus-strand DNA for virus G361. Virus G404 has 54 nts between the 5' ends of the two strands of DNA as a consequence of a 61-nt deletion introduced into virus G361. (A) Southern blot analysis indicates a higher proportion of relaxed circular DNA for G404 than for G361. Viral DNA was isolated from cytoplasmic capsids from transfected cells. RC, DL, and SS represent the mobilities of the relaxed circular, duplex linear, and minus-strand single-stranded forms of the genome, respectively. The hybridization probe was genomic length and minus strand specific. WT, wild type. (B) Primer extension analysis to determine the position of the 5' end of minus-strand DNA and measure its level. The oligonucleotide used was derived from positions 2363 to 2382. The band which represents the internal standard is labeled IS. To the left of each band representing the 5' end of minus-strand DNA for each virus is an asterisk. The primer extension product of virus G361 is 4 nt longer than the wild type due to the nature of its mutation (see Materials and Methods). The primer extension product of virus G404 is smaller than the wild type due to a 61-nt deletion. (C) Primer extension analysis with oligonucleotide 2567⁻ to determine the position and measure the level of the 5' end of plus-strand DNA. Oligonucleotide 2567⁻ was used to generate the sequence ladder. The bands representing the 5' end of the plus-strand DNA are labeled with asterisks.

termini of the two strands, either by introducing small (<100-nt) insertions or changing the position of the acceptor site for the plus-strand primer, inhibited circularization. The phenotypes seen with the viruses described in this study are not a reflection of the inability of the virus to tolerate similar perturbations elsewhere in the genome. Comparable mutations elsewhere in the genome do not lead to inhibition of circularization. For example, in this study, a 72-nt insertion within the middle of the genome did not influence circularization. In an earlier study (4), deletions averaging 300 nts in length, that were placed in multiple locations within the genome did not perturb circularization. Our results are consistent with the idea that the region of the genome between DR2 and the 5' end of minus-strand DNA is in a dynamic structure that is important for circularization. This structure is perturbed in our mutants.

Primer translocation can proceed to ectopic acceptor sites.

We have shown that when the normal acceptor site, DR2, is deleted, the plus-strand primer can efficiently translocate to an introduced copy of the DR sequence 66 nts away from the natural position of DR2 (virus G361). The level of primer translocation to the ectopic site in this variant virus was within 50% of the level of primer translocation for the wild-type virus. For virus G304, in which the ectopic copy of DR was 104 nts from the site of the deleted DR2, primer translocation was also detected, albeit at a lower level. We interpret these results to indicate that the mechanism of primer translocation can tolerate variation in the position of the acceptor site on the template. The range of positions on the minus strand that can serve as acceptor sites for primer translocation remains to be determined. It is possible that the difference in the abilities of viruses G304 and G361 to support primer translocation reflects a requirement for an optimal number of nucleotides between the acceptor site and other *cis*-acting sequences for primer translocation, such as the element in the 5E region (4). The region 5E element, which is adjacent to, but not contiguous with, DR2, appears to play an essential role in primer translocation.

Circularization is sensitive to the relative position of the 5' end of plus-strand DNA. We showed that for seven different variant viruses with insertions of 48 nts or greater between the positions of the 5' termini of the respective strands of DNA, circularization was reduced to 13% of the value of the wild-type virus or less. Insertions of 18 and 35 nts also led to inhibition of circularization, but it was less severe. An insertion of 12 nts had little effect on circularization. These results are consistent with the notion that the degree of inhibition of circularization is proportional to the number of nucleotides inserted until the increase is sufficiently large (e.g., 48 nts) and further increases in the number of nucleotides do not lead to a further reduction in the extent of circularization. Normally during the synthesis of plus-strand DNA, the minus-strand DNA template is copied to its 5' end before the template switch (11). Because the oligonucleotide used in the primer extension analysis to detect plus-strand DNA before circularization was complementary to the 18 nts synthesized immediately prior to the template switch, it appears that the defect is manifested precisely at the step of circularization and not during the process of synthesizing plus-strand DNA immediately before the switch.

When the number of nucleotides between the primer acceptor site and the 5' end of the minus strand was reduced from 115 (virus G361) to 54 (G404), restoration of circularization, albeit partial, was observed. As a consequence of the deletion in virus G404, the sequence of the nascent plus strand was different than the wild-type sequence. Although this result emphasizes the role of the number of nucleotides in this region

of the genome in the process of circularization, the observation that the restoration was incomplete suggests that there are additional requirements within this region.

How does the number of nucleotides within this region of the genome contribute to circularization? We offer the following model. During the process of primer translocation, the primer donor site, which is at the 3' end of minus-strand DNA, is juxtaposed with the primer acceptor site, DR2. The 3' end of minus-strand DNA also contains the 3' copy of *r*, which subsequently is the acceptor template for circularization. Upon completion of primer translocation, the tertiary structure of the nascent replication complex is poised to carry out the requisite steps that will, with the satisfaction of some additional requirements, ultimately lead to successful circularization. These steps include reorganization of the nascent replication complex such that the 5' and 3' ends of the minus strand are juxtaposed. For the variant viruses that were inhibited for circularization, the positioning of the 3' end of the minus strand with the acceptor site for primer translocation changes the tertiary relationship between two determinants that play a subsequent role in circularization and leads to inhibition of circularization. A candidate for one of these determinants is the *cis*-acting sequence near the 3' end of the minus-strand template, within region 3E, that is required for circularization (4). Another candidate determinant is the P protein which is covalently linked to the 5' terminus of minus-strand DNA, which is also the location of the 5' copy of *r*.

The mechanism of circularization is more complex than initially anticipated. The findings presented in this article bring to three the number of *cis*-acting requirements for circularization: (i) the terminal redundancy, *r*, found at the ends of the minus-strand template; (ii) a *cis*-acting sequence near or at the 3' end of the minus-strand template that is distinct from the terminal redundancy, named 3E; and (iii) the number of nucleotides between the acceptor site of primer translocation and the 5' copy of *r* on the minus-strand template. The contribution, if any, of proteins to the mechanism of circularization is unclear. As the array of *cis*-acting requirements increases, it becomes more reasonable to posit that proteins, viral and/or cellular, contribute to the process of circularization by interacting with *cis*-acting sequences. Our results also provide an explanation for why the plus strands that are initiated in situ, at DR1, do not circularize upon elongation to the 5' end of the minus-strand template.

On the basis of our results, we infer that the events associated with the initiation of synthesis of plus-strand DNA (e.g., primer translocation) and circularization share a mechanistic link. These results are consistent with the notion that the capsid particle and its contents are in a dynamic tertiary structure that is integral to the execution of each step of DNA replication. As DNA replication progresses, the structure of the complex changes in an ordered fashion, with each subsequent step dependent upon the successful execution of the preceding step.

Comparison to retroviral reverse transcription. Do other reverse transcription pathways have a requirement during plus-strand synthesis similar to that described here for hepadnaviruses? The corresponding steps of the retrovirus pathway have been examined, and the answer appears to be no. Bowman and colleagues (1) studied the replication of a spleen necrosis virus vector that had the normal site of plus-strand initiation replaced with a site more internal on the template (analogous to virus G361). In that study, the variant spleen necrosis virus replicated at levels similar to that of the wild-type comparison. In another study (14), derivatives of Moloney murine leukemia

virus that had 3,100 nts inserted into U3 (analogous to virus G410) replicated at high levels. Based on these comparisons, it appears that the hepadnavirus and retrovirus pathways are dissimilar in this requirement. This might not be surprising, considering that the preceding step in hepadnavirus reverse transcription, primer translocation, does not occur in the retrovirus pathway.

ACKNOWLEDGMENTS

We thank Ashok Aiyar, Paul Lambert, Karlyn Mueller-Hill, and Bill Sugden for critical review of the manuscript.

This work was supported by NIH research grants GM50263 and CA22443 and core grant CA07175. D.D.L. is the recipient of an American Cancer Society Junior Faculty Research Award (JFRA-651).

REFERENCES

1. Bowman, E. H., V. K. Pathak, and W.-S. Hu. 1996. Efficient initiation and strand transfer of polypurine tract-primed plus-strand DNA prevent transfer of internally initiated plus-strand DNAs. *J. Virol.* **70**:1687-1694.
2. Calvert, J., and J. Summers. 1994. Two regions of an avian hepadnavirus RNA pregenome are required in *cis* for encapsidation. *J. Virol.* **68**:2084-2090.
3. Ganem, D. 1996. *Hepadnaviridae* and their replication, p. 2703-2737. In B. N. Fields, D. M. Knipe, P. M. Howley, et al. (ed.), *Virology*, 3rd ed. Lippincott-Raven Publishers.
4. Havert, M. B., and D. D. Loeb. 1997. *cis*-acting sequences in addition to donor and acceptor sites are required for template switching during synthesis of plus-strand DNA for duck hepatitis B virus. *J. Virol.* **71**:5336-5344.
5. Lathe, R., M. P. Kieny, S. Skory, and J. P. Lecocq. 1984. Linker tailing: unphosphorylated linker oligonucleotides for joining DNA termini. *DNA* **3**:174-182.
6. Lien, J.-M., C. A. 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.
7. Lien, J.-M., D. J. Petcu, C. E. Aldrich, and W. S. Mason. 1987. Initiation and termination of duck hepatitis B virus DNA synthesis during virus maturation. *J. Virol.* **61**:3832-3840.
8. Loeb, D. D., R. C. Hirsch, and D. Ganem. 1991. Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses; implication for other reverse transcribing elements. *EMBO J.* **10**:3533-3540.
9. Loeb, D. D., and R. Tian. 1995. Transfer of the minus strand of DNA during hepadnavirus replication is not invariable but prefers a specific location. *J. Virol.* **69**:6886-6891.
10. Loeb, D. D., R. Tian, and K. J. Gulya. 1996. Mutations within DR2 independently reduce the amount of both minus- and plus-strand DNA synthesized during duck hepatitis B virus replication. *J. Virol.* **70**:8684-8690.
11. Loeb, D. D., K. J. Gulya, and R. Tian. 1997. Sequence identity of the terminal redundancies on the minus-strand DNA template is necessary but not sufficient for the template switch during hepadnavirus plus-strand DNA synthesis. *J. Virol.* **71**:152-160.
12. Mueller-Hill, K., and D. D. Loeb. 1996. Previously unsuspected *cis*-acting sequences for DNA replication revealed by characterization of a chimeric heron/duck hepatitis B virus. *J. Virol.* **70**:8310-8317.
13. Nassal, M., and A. Rieger. 1996. A bulged region of the hepatitis B virus RNA encapsidation signal contains the replication origin for discontinuous first-strand DNA synthesis. *J. Virol.* **70**:2764-2773.
14. Reddy, S., J. V. DeGregori, H. von Melchner, and H. E. Ruley. 1991. Retrovirus promoter-trap vector to induce *lacZ* gene fusions in mammalian cells. *J. Virol.* **65**:1507-1515.
15. Seeger, C., D. Ganem, and H. E. Varmus. 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. *Science* **232**:477-484.
16. Sprengel, R., C. Kuhn, H. Will, and H. Schaller. 1985. Comparative sequence analysis of duck and human hepatitis B virus genomes. *J. Med. Virol.* **15**:323-333.
17. Staprans, S., D. D. Loeb, and D. Ganem. 1991. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. *J. Virol.* **65**:1255-1262.
18. Tavis, J. T., S. Perri, and D. Ganem. 1994. Hepadnavirus reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer. *J. Virol.* **68**:3536-3543.
19. Wang, G.-H., and C. Seeger. 1993. Novel mechanism for reverse transcription in hepatitis B viruses. *J. Virol.* **67**:6507-6512.
20. 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.