JOURNAL OF VIROLOGY, July 1996, p. 4567–4575 Vol. 70, No. 7
0022-538X/96/$04.00+0

Covalently Closed Circular Viral DNA Formed from Two Types of Linear DNA in Woodchuck Hepatitis Virus-Infected Liver

WENGANG YANG,[1] WILLIAM S. MASON,[2] AND JESSE SUMMERS[1]*

Department of Cell Biology, University of New Mexico School of Medicine, Albuquerque, New Mexico 87131,[1] and Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111[2]

Received 17 January 1996/Accepted 5 April 1996

We found that livers from woodchucks chronically infected with woodchuck hepatitis virus (WHV) contained covalently closed circular DNA (cccDNA) molecules with deletions and insertions indicative of their formation from linear viral DNA by nonhomologous recombination, as we previously described for the duck hepatitis B virus (W. Yang and J. Summers, J. Virol. 69:4029–4036, 1995). However, evidence for two different types of linear precursors was obtained by analysis of the recombination joints in WHV cccDNA. Type 1 linear precursors possessed the structural properties that correspond to those of in situ-primed linear DNA molecules, which constitute between 7 and 20% of all viral DNA replicative intermediates synthesized in the liver. Type 2 linear precursors are hypothetical species of linear DNAs with a terminal duplication of the cohesive-end region, between DR1 and DR2. This type of linear DNA has not been previously described and was not detected among the DNA species present in nucleocapsids. A fraction of cccDNAs formed from both type 1 and type 2 linear DNAs are predicted to be functional for further DNA synthesis, and some evidence for the formation of two or more generations of cccDNA from linear DNA was observed.

Hepadnaviruses are small DNA-containing viruses that replicate their DNA genomes through reverse transcription of RNA (17, 24). The RNA genome intermediates (pregenomes) are transcribed from a pool of covalently closed circular viral DNA (cccDNA) in the nuclei of infected cells. Reverse transcription takes place in the cytoplasm through a process that uses a combination of template switches (3, 14, 15, 20–22, 30) to produce relaxed circular double-stranded DNAs, which can be transported to the nucleus to form additional copies of cccDNA (29, 32). Progeny cccDNAs retain the precise sequence of the parental cccDNA.

We recently reported a second, minor pathway of DNA replication via reverse transcription in the avian hepadnavirus duck hepatitis B virus (DHBV). In this pathway, linear double-stranded DNA produced as a result of a failure to prime second-strand (plus-strand) DNA synthesis at the correct location was efficiently converted to cccDNAs by nonhomologous recombination near the two ends of the linear DNA (33). cccDNA molecules resulting from such recombination thus acquired sequence alterations at the site of the recombination and therefore were not identical to the parental cccDNA. Some cccDNA molecules produced from linear DNA were able to produce pregenomes. However, as a result of the mutation created by nonhomologous recombination, most of the pregenomes gave rise only to linear double-stranded DNAs, which recombined to produce further successive generations of cccDNA. We called this process "illegitimate replication," because each such generation of cccDNA differed in sequence from its parent molecule at the site of nonhomologous recombination. We speculated that the low level of virus expression in cells carrying out illegitimate replication might provide such cells a survival advantage during an antigen-specific cellular immune response.

In this paper, we report studies that examined the structures of cccDNA molecules in livers acutely and chronically infected with the mammalian hepadnavirus woodchuck hepatitis virus (WHV) (11, 28). We found that cccDNA molecules commonly suffered deletions that spanned the region corresponding to the predicted ends of two different types of linear double-stranded DNA, consistent with their formation by nonhomologous recombination between the ends of linear DNA precursors. Only a small fraction of such cccDNA molecules would be predicted to be active in further DNA replication; however, evidence for more than one generation of cccDNA formed by nonhomologous recombination of linear DNAs was seen in several sequences, indicating that some virus replication occurs through illegitimate replication.

MATERIALS AND METHODS

Woodchucks. Woodchucks were obtained from Northeast Wildlife and were either naturally or experimentally infected with WHV (11). The animals were housed at the Fox Chase Cancer Center Animal Facility. Liver biopsy samples from acutely or chronically infected woodchucks were immediately frozen and stored at −80°C.

Extraction of viral DNA from WHV-infected woodchuck livers. WHV-infected woodchuck livers were homogenized with a loose-fitting Dounce homogenizer as a 10% (wt/vol) suspension in ice-cold TE 50:1 (50 mM Tris · HCl, 1 mM EDTA [pH 8.0]). For isolation of cccDNA, half of the homogenate was mixed with an equal volume of 4% sodium dodecyl sulfate (SDS) and protein-DNA-detergent complexes were precipitated by using 0.25 volume of 2.5 M KCl and centrifugation (11, 26). The supernatant was extracted twice with phenol, once with phenol-chloroform (1:1), and once with chloroform. Nucleic acids were precipitated with 2 volumes of ethanol.

For isolation of replicative intermediates, the remaining half of the homogenate was adjusted to 0.5% with Nonidet P-40 and the nuclei and cellular debris were removed by centrifugation. Total nucleic acids were isolated from the supernatants after digestion with 0.5 mg of pronase per ml in the presence of 1% SDS, phenol extraction, and ethanol precipitation (9, 26, 27).

Further purification of cccDNA. cccDNA was further purified from contaminating single- and double-stranded DNA by a modification of the alkali lysis procedure for isolation of plasmid DNA (2). cccDNA-containing samples were diluted in an equal volume of 0.1 N NaOH and incubated at 37°C for 30 min. This step irreversibly denatured all double-stranded DNA species that were not covalently closed. After denaturation, the DNA was neutralized by the addition of 3 M potassium acetate (pH 5.0) to a final concentration of 0.6 M. Single-stranded DNA was efficiently removed at this pH by phenol extraction, and the double-stranded cccDNA, which quantitatively remained in the aqueous phase, was recovered by ethanol precipitation.

cccDNAs were gel purified by electrophoresis through 1% low-melting-temperature agarose (SeaPlaque; FMC). A gel slice corresponding to the position of

* Corresponding author. Phone: (505) 277-7979. Fax: (505) 277-7979. Electronic mail address: JSUMMERS@cobra.unm.edu.

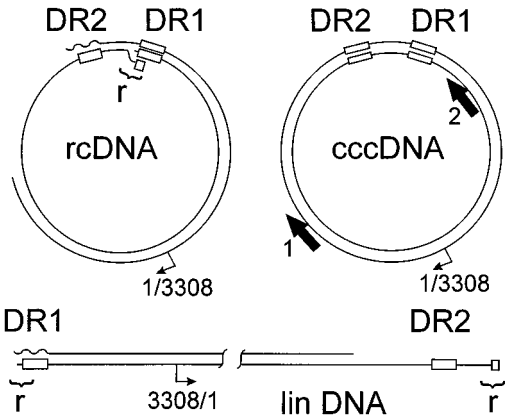

FIG. 1. Structural features of known forms of viral DNA. The three known forms of double-stranded viral DNA, relaxed circular (rc), covalently closed circular (ccc), and in situ-primed linear (lin) are shown. The viral plus strand is on the outside and is oriented 5′ to 3′ in the clockwise direction. The RNA primer for plus-strand synthesis is indicated by the wavy line of the 5′ end. The positions of DR1 (nucleotides 1941 to 1951) and DR2 (nucleotides 1719 to 1729) are indicated by rectangles (not to scale). In relaxed circular and linear DNA, the 10-nucleotide terminal redundancy is indicated by **r**. The 5′ end of the minus strand contains 4 nucleotides of DR1 starting at position 1944 (truncated rectangle), the 3′ end of the minus strand is located at 1935, and the 5′ end of the plus strand is located approximately at 1730 in rcDNA or at 1952 in linDNA. The 3′ ends of the incomplete plus strands are located at many different sites, resulting in partially double-stranded relaxed circular and linear DNAs. The upstream primer 1 (nucleotides 780 to 799) and the downstream primer 2 (nucleotides 2407 to 2391) used to amplify the region spanning DR1 and DR2 in cccDNA are shown.

0.8- to 1.7-kbp linear markers was excised, and the DNA was recovered after melting the agarose at 68°C, cooling it to 37°C, and extraction with phenol. The large gel slice allowed the recovery of cccDNA molecules with substantial deletions as well as full-length cccDNA.

PCR amplification. PCR amplifications were carried out for 30 cycles with *Taq* DNA polymerase, using a kit and methods recommended by the manufacturer (InVitrogen Corp., San Diego, Calif.). Two sets of primers were used, as shown in Fig. 1 and 4. One set of primers was designed to amplify a 1,628-bp region of the genome for the experiment in Fig. 1 and 2. Primers for this reaction consisted of the plus-strand sequence from 780 to 799 (13) and the minus-strand sequence from 2407 to 2391. A second set of primers was used in the experiment in Fig. 4 to amplify a 152-bp region. Primers for this reaction consisted of the plus-strand sequence from 1886 to 1905 and the minus-strand sequence from 2037 to 2017. cccDNA templates were gel purified and linearized by digestion with *Bam*HI before use. Amplification reactions were started by the addition of Mg^{2+}-containing buffer after 2 min at 80°C (hot start) and consisted of a denaturation step at 94°C for 1 min, annealing at 50 or 55°C for 2 min, and elongation at 72°C for 3 min. Amplification products were diluted 100-fold in fresh reaction buffer and subjected to two additional cycles of amplification to eliminate heteroduplex molecules before being cloned. The reamplified DNAs were cloned in the TA cloning vector, pCRII, and transformed into *Escherichia coli* INVαF′. Transformed colonies were identified on 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) plates, and white colonies were picked for analysis.

Consideration of PCR artifacts. Reconstruction amplification reactions were run on purified cloned WHV DNA. We never detected any sequences among the products of such reactions that were different from wild-type sequences other than occasional single base substitutions. Additional controls included with all amplification reactions were designed to detect artifacts that might have arisen from amplification of strands of relaxed circular DNA molecules that might have contaminated the cccDNA fractions. Since many of the amplification products we detected were missing the sequences normally found missing in WHV plus strands of relaxed circular DNA, we wanted to be certain that our amplification conditions did not allow products to be polymerized across the ends of incomplete plus strand templates by, for example, polymerase jumping from the 5′ end to the 3′ end of the same template. Therefore, we used as a template cloned WHV DNA digested with two restriction enzymes, *Hin*dIII and *Nsi*I, which cleaved the plasmid in the WHV sequences at nucleotides 1914 and 2190, between the PCR primer-binding sites. The fragment containing the primer-binding sites was then gel purified and amplified with the primers shown in Figure 1. Amplification of this template compared with a continuous uncleaved template indicated that amplification reactions with the discontinuous template were 3 orders of magnitude less efficient than amplification of the continuous template (results not shown). Moreover, the product that was amplified with large amounts of the discontinuous template corresponded to that expected from uncleaved DNA, which we infer must have been present in the discontinuous-template preparation. No evidence for products with deleted regions was seen. These controls (results not shown) indicated that relaxed circular viral DNA would have to be present at concentrations 1,000-fold higher than cccDNA in our purified samples to generate a significant proportion of the PCR product.

Another source of potential artifacts might include those arising from cloning of heteroduplex DNA molecules formed between two different recombinant strands during the last cycles of amplification. These heteroduplexes could form between the denatured DNAs during the annealing step and result in displacement of the annealed primers. We do not know what the result of cloning such

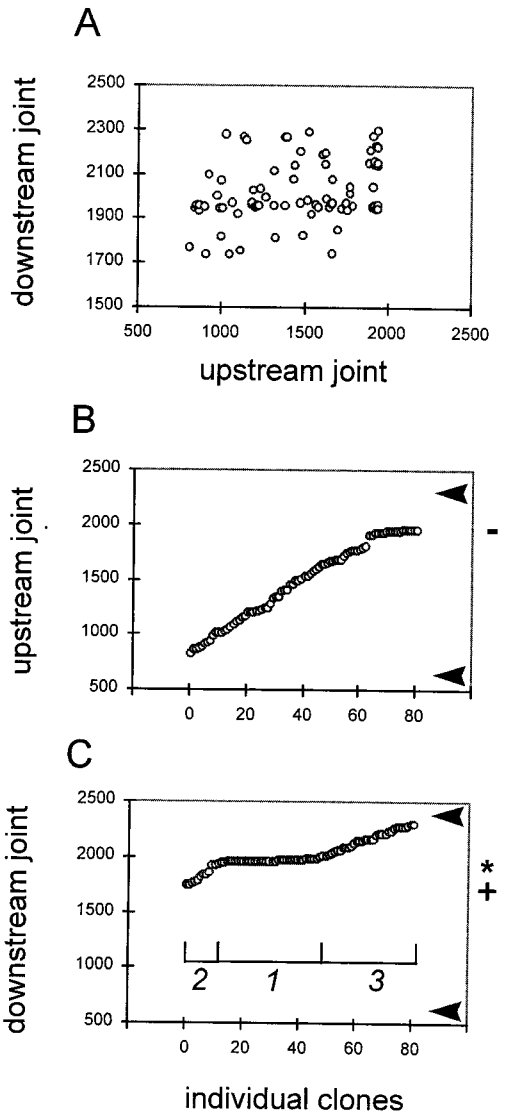

FIG. 2. Positions of upstream and downstream recombination sites for 81 clones amplified from cccDNA. PCR fragments were amplified from gel-purified cccDNA and cloned into pCRII. The sites of recombination causing a single internal deletion were determined by sequencing. (A) Correlation graph of the positions of the upstream sites and the matching downstream sites. No correlation is seen in their locations. (B) Distribution of upstream recombination sites by nucleotide position for individual clones arranged along the *x* axis according to an ascending order of the upstream sites. Approximately even distribution of sites within a region bounded by the upstream primer (lower arrowhead) and 5′ end of the minus strand in DR1 (−) is indicated by the relatively constant slope, which is a function of the mean distance between neighboring sites. (C) Distribution of downstream sites by nucleotide position, arranged in ascending order. Three distinct regions of distribution (*1*, *2*, and *3*) are distinguished by their different densities of sites (regions of different slopes). The positions of the 3′ end of the minus strand (∗) and the 5′ end of the plus strand (+) are indicated.

TABLE 1. Expected and observed numbers of ambiguous recombination joints

No. of ambiguities[a]	Expected frequency of joints[b]	Expected no. of joints	Observed no. of joints[c]
0	0.56	78	51
1	0.28	39	47
2	0.11	15	25
3	0.035	5	10
4	0.010	2	4
≥5	0.005	1	3
Total	1.000	140	140

[a] The number of ambiguities is equal to the length of the nucleotide homology around the site of joining.

[b] The expected frequency of joints with different amounts of ambiguity in the position of joining when formed by a sequence-independent joining mechanism was calculated according to the formula $f = (n + 1)(3/4)^2(1/4)^n$, where f is the frequency of occurrence when there are n nucleotides of homology. The frequency is the product of the probability that a stretch of n nucleotides will have identity between any particular two sites on the two respective parental molecules $[(1/4)^n]$, the probability that the stretch of n nucleotides will be flanked by two nonidentical nucleotide pairs $[(3/4)^2]$, and the number of possible crossover points around the stretch of n nucleotides that result in the same product $(n + 1)$.

[c] The recombination joints used for the analysis were the set of 140 of 157 joints in Fig. 2 and 5 that did not contain extra sequences inserted at the sites of joining.

heteroduplex molecules would be, but potentially complex joints could be generated during repair of such structures in *E. coli* before their replication. As in our previous study (33), we attempted to minimize the presence of such heteroduplexes in the products before cloning by diluting the amplified products to low concentration in fresh amplification mix and carrying out two additional cycles of denaturation, annealing, and elongation. While about 90% of the cloned sequenced recombination joints were apparently the result of a simple joining reaction, about 10% contained some additional nucleotides that may have arisen from this potential artifact.

Another conceivable artifact might be the formation of recombination joints during PCR by annealing of two incompletely elongated strands at sites of significant complementarity. We analyzed the series of 140 simple "nonhomologous" recombination joints reported in Fig. 2A and 5A for evidence of short sequence homology at the sites of joining. Homology was detected by ambiguity in one or more nucleotides in the location of the precise recombination joint. The data, summarized in Table 1, demonstrate a weak bias of only one nucleotide homology over that expected by chance at the sites of joining. This bias is evident from the increase of 27 over the expected 62 of 140 joints that show any ambiguity at all in the position of the joint. In other words, sequence homology appeared to be a factor in the formation of only 27 of 140 joints. This result argues against the artificial joining of sequences by adventitious copriming of partially elongated products, since this small amount of homology could not result in stable base pairing. For these reasons, we believe that the nonhomologous recombination joints we report here were present in the original template sample and were not generated by PCR artifacts.

Colony screening by hybridization. Transformed colonies were inoculated onto nylon filters placed on the surface of Luria-Bertani agar plates with ampicillin selection and incubated until visible colonies formed. The colonies were lysed in situ by incubation with 10% SDS (5 min), 0.5 M NaOH (5 min, twice), 1 M Tris · HCl (pH 7.4) (5 min), and 0.2 M Tris · HCl (pH 5.0) containing 1.5 M NaCl (5 min). The membranes were dried under a heat lamp and then incubated at 68°C for 4 h in hybridization buffer consisting of 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate [pH 7.5]) containing 4× Denhardt's solution and 0.1% SDS.

^{32}P-oligonucleotides used for specific screening were generated by phosphorylation with T4 polynucleotide kinase. Filters were hybridized to the specific probes in hybridization buffer overnight at 37°C, washed with several changes of 1× SSC–0.1% SDS at room temperature for 30 min and then at 37°C for 1 h. Under the conditions used, hybridization did not occur if there were single base insertions, deletions, or mismatches in the target sequence.

Sequencing. Cloned PCR products were sequenced either as plasmid minipreps as previously described (33) or after PCR amplification from single colonies. For the latter procedure, a small amount of each colony to be analyzed was suspended in the PCR mixture before the beginning of the denaturation step. Amplification was carried out as previously described with a biotinylated upstream primer (from 780 to 799). The biotinylated products were adsorbed with an equal volume of streptavidin-coated M-280 Dynabeads (Dynal Corp.) suspended in a solution of 20 mM Tris · HCl (pH 8.0)–2 M NaCl–1 mM EDTA and washed once with the help of a magnetic particle concentrator (Dynal no. 120.04). The nonbiotinylated strand was released from the beads by denaturation in 0.1 N NaOH and removed by washing. Washed beads with specifically bound plus-strand products were used directly in sequencing reactions.

RESULTS

Detection of defective cccDNA in chronic WHV infection. We used PCR to amplify the region of partially purified cccDNA corresponding to the predicted ends of double-stranded linear DNA, as described previously (33) and shown in Fig. 1. The PCR products derived from cccDNA from the liver of a chronically infected woodchuck were cloned into plasmid pCRII, and individual colonies containing WHV inserts were isolated. The inserts from the plasmids in these colonies were reamplified for determination of their sizes and for direct sequencing. Of 171 clones analyzed, 100 contained inserts of WHV sequences that were noticeably smaller than expected whereas the remainder contained inserts of a length indistinguishable from the expected length (results not shown). As shown below, this result does not reflect the true proportion of deletion mutants, which are selectively amplified because of their size. Sequencing of 106 inserts revealed that 81 had suffered a single internal deletion or, in one case, an insertion. No correlation was seen between the sites joined by the deletions with respect to either their locations or their nucleotide sequences, except for a slight preference for homology (Fig. 2A; Table 1).

We examined separately the distribution of the two sites that flanked each deletion. This analysis allowed us to infer the structures of the precursors of the deleted cccDNAs in the following way. Each deletion could be viewed as the result of a nonhomologous recombination between a pair of sites in either a circular or a linear precursor, resulting in circularization of the segment of DNA containing the PCR primer-binding sites and deletion of the intervening or terminal sequences, respectively. In the case of a circular precursor, sites of recombination could have occurred at any location between the two PCR primer-binding sites to generate a circular product, and

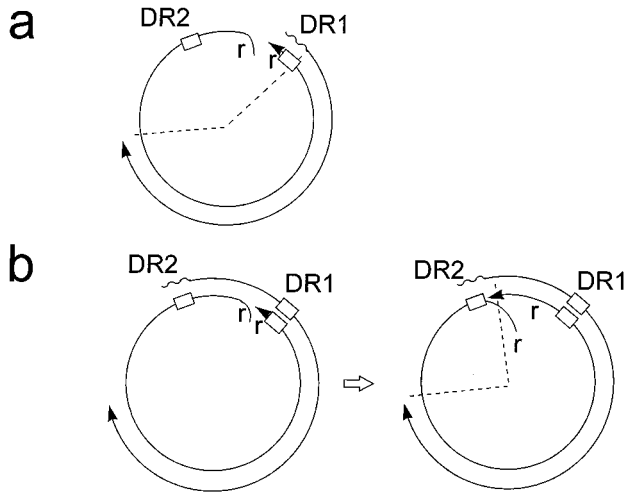

FIG. 3. Model for the formation of cccDNAs with type 1 or type 2 recombination joints. Partially double-stranded linear DNAs produced by in situ priming of plus-strand synthesis (a) or by elongation of the minus strand of relaxed circular DNA (b) undergo nonhomologous recombination at sites near the double-stranded ends to produce cccDNA with characteristic deletions. The position of the 10-nucleotide terminal duplication in the minus strand is indicated (**r**). DR1, DR2, and the RNA primer for plus-strand synthesis are indicated as in Fig. 1.

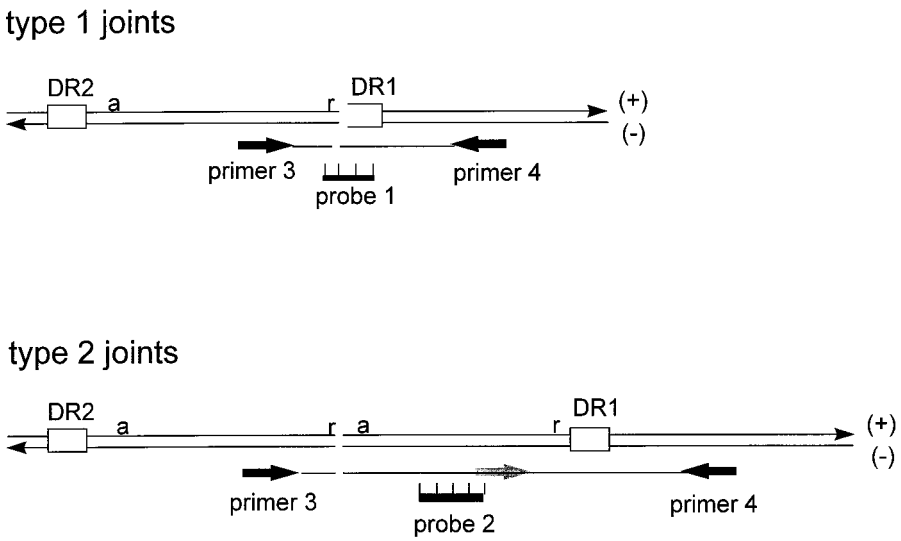

FIG. 4. Strategy for amplification and screening of recombination joints near the ends of the type 1 and 2 linear precursors. Primers 3 (nucleotides 1886 to 1905) and 4 (nucleotides 2037 to 2017) were used to amplify gel-purified cccDNA templates. The PCR products were cloned, and colonies were screened with oligonucleotide probes 1 (nucleotides 1935 to 1949) and 2 (nucleotides 1869 to 1888). Colonies that hybridized to probe 2 contained type 2 recombination joints in which sequences near DR1 were joined to sequences near DR2 to produce a partial repeat of the cohesive end region, as shown in Fig. 5 and 6. Colonies that failed to hybridize to probe 1 suffered deletions or insertions in the **r** region, as shown in Fig. 7. Primer-binding sites and primer orientation are indicated by arrows. DR1, DR2, and the small repeat sequence (**r**) are indicated as in Fig. 1.

therefore upstream (counterclockwise on the genome in Fig. 1) and downstream recombination sites could be found throughout the amplified region. In the case of a linear precursor, circular products would have resulted only from recombination between two sites that were positioned so as to join the two ends of the linear molecule. Thus, the distribution of the upstream sites in the amplified segment would be limited to the region between the upstream primer-binding site and the nearest downstream end whereas the distribution of the downstream recombination sites would be limited to the region between the downstream primer and its nearest upstream end (Fig. 1, lin DNA).

This distribution analysis, shown in Fig. 2B and C, reveals the characteristic limited distribution of the upstream and downstream sites that is predicted from a linear precursor to the circular DNA. For example, the upstream sites (Fig. 2B) were evenly dispersed between the upstream PCR primer-binding site and nucleotide 1944, a position corresponding to the 5′ end of the minus strand (20–22). No sites were found downstream of this position. The 5′ end of the minus strand corresponds to one end of a previously described species of linear double-stranded DNA which is produced by in situ priming of plus-strand synthesis (16, 23). Downstream recombination sites were likewise limited to a specific but distinct region of the amplified segment (Fig. 2C).

The data in Fig. 2C suggest the existence of two types of linear precursors, types 1 and 2, which could be distinguished by their different upstream ends. This inference is drawn from the observation that the distribution of downstream recombination sites consisted of three contiguous regions with different densities of clustering. The region containing the highest density of downstream recombination sites (region 1) extended from nucleotide 1936, close to the upstream end of in situ-primed linear DNA (20–22), to position 1967, 32 nucleotides downstream. Thirty-four independent recombination sites were located in region 1, occurring at an average density of 1 per 0.94 ± 0.92 nucleotide (mean and standard deviation of nearest-neighbor distances). In DHBV, downstream recombination sites in cccDNA derived from in situ-primed linear DNAs were distributed preferentially within a similar region near the upstream end of the linear precursor (33). Therefore, most of the 34 sites located in this region may have occurred by circularization of a linear in situ-primed DNA precursor, as illustrated in Fig. 3a (type 1 precursor). This molecule would have a structure similar to that of the in situ-primed linear double-stranded DNA that has been described in DHBV (23).

Twelve recombination sites, located in a region of less dense clustering (mean density of 1 site per 17 ± 18 nucleotides) between nucleotide 1733 and 1926 (region 2), could not be explained by synthesis from an in situ-primed linear precursor. The upstream limit of this region is very close to the major 5′ end of the plus-strand DNA at DR2 (nucleotide 1730); however, no linear DNA with an upstream end at this position has been previously described. A linear molecule with the predicted end, however, could be generated from relaxed circular DNA by elongation of either strand through the cohesive-end region (the region of base pairing between the 5′ ends of the two strands), resulting in displacement of the cohesive ends and linearization of the molecule. An example involving elongation of the minus strand is illustrated in Fig. 3b (type 2 precursor).

Downstream of region 1, 35 recombination sites were distributed, with a mean density of 1 site per 9 ± 8 nucleotides, along the genome to near the downstream PCR primer-binding site (region 3). Since this region is located downstream from the predicted ends of both type 1 and 2 precursors, it may contain recombination joints generated from either type. Therefore, the relative frequency of cccDNAs synthesized from type 1 and type 2 precursors is uncertain.

Mapping of recombination joints close to the ends of type 1 and type 2 linear precursors. We used a more extensive analysis of recombination joint distributions to confirm the positions of the two ends of the hypothetical type 2 precursor. This analysis involved the mapping of upstream and downstream sites of recombination that were located closer to the ends of the type 2 linear precursor. To isolate a large number of such

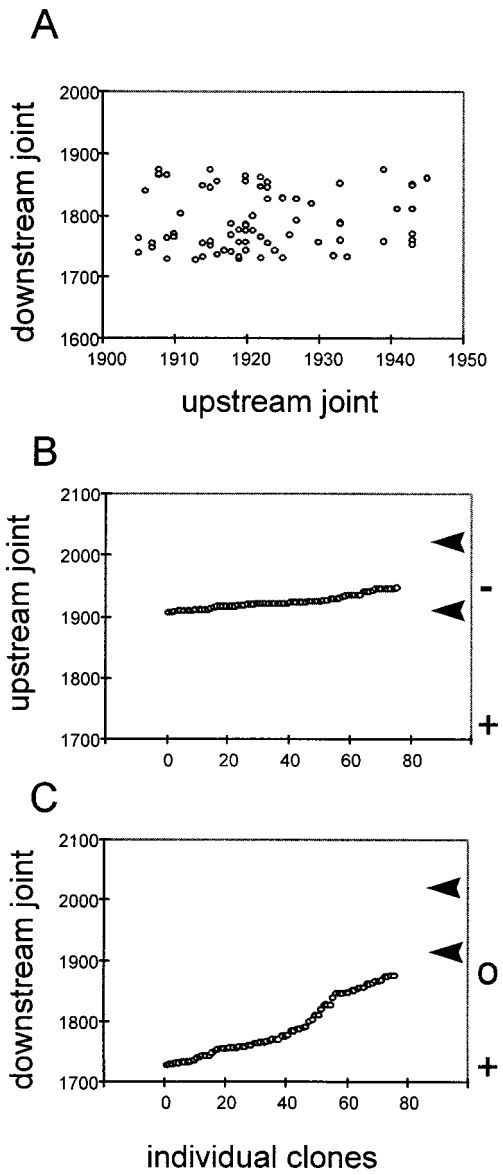

FIG. 5. Distribution of recombination joints from type 2 linear precursors. Regions of gel-purified cccDNA were amplified with primers 3 and 4 and cloned into pCRII. The inserts from 79 colonies hybridizing with the oligonucleotide 2 probe were sequenced, and the upstream and downstream recombination joints were identified. According to the design of the experiment, upstream joints were located within the region between the two primers and all downstream joints were upstream of this region, producing a partial duplication of the cohesive-end region. (A) Correlation plot of the nucleotide positions of the upstream and downstream joints. No correlation was seen. (B) Locations of the upstream recombination sites of individual clones arranged along the *x* axis in ascending order of the position of the recombination sites. The sites are evenly distributed between the position of the upstream primer (lower arrowhead) and the 5′ end of the minus strand (−). (C) Location of the downstream recombination sites. The location of these sites is limited by the position of the 5′ end of the plus strand (+) and the position of the oligonucleotide 2 probe (o), used to screen the colonies.

closely positioned joints, we carried out a selective amplification of a subset of joints originating from type 1 and type 2 precursors by using the strategy depicted in Fig. 4. A region of cccDNA between two closely spaced PCR primers located upstream and downstream of DR1 (Fig. 4, primers 3 and 4) was amplified from gel-purified cccDNA and cloned into pCRII. The recombination joints from type 1 or type 2 linear precursors could produce two different types of PCR products: type 1 precursor joints would generate a PCR product of about 150 nucleotides, with deletions or small duplications around the **r** region, while type 2 joints could produce PCR products of up to 350 nucleotides containing a partial duplication of the cohesive-end region. Bacterial colonies containing type 2 joints were identified by their ability to hybridize to an oligonucleotide probe specific for a region of the cohesive end mapping outside the sequence flanked by the PCR primer-binding sites on normal cccDNA (Fig. 4, probe 2). Of a total of approximately 1,000 WHV-containing clones, 76 hybridized to probe 2 and were sequenced.

As shown in Fig. 5A, the position of the upstream recombination joint did not correlate with the position of the downstream joint, as was the case in the combination of joints shown in Fig. 2A. As in Fig. 2A, recombination joints also did not show a preference for joining at particular nucleotides except for a slight preference for homology (one nucleotide) at the sites of joining (Table 1). The individual distributions of the upstream and downstream joints were determined (Fig. 5B and C, respectively). Of 76 upstream joints, 75 were located in the region between the upstream PCR primer and position 1944, which corresponds to the 5′ end of the minus strand in DR1, with one upstream joint occurring at position 1945. All downstream joints were evenly distributed between the oligonucleotide-binding site used for screening and nucleotide 1726, close to the predicted major 5′ end of the plus strand at DR2. These data support the hypothesis that the fixed ends of a type 2 precursor would be located at the positions predicted for a linear double-stranded DNA molecule generated by displacement synthesis through the cohesive-end region of relaxed circular DNA.

Of the 76 clones sequenced, 4 clones that contained evidence of at least two recombination events were observed. The structures of these four clones are shown in Fig. 6. In clones a

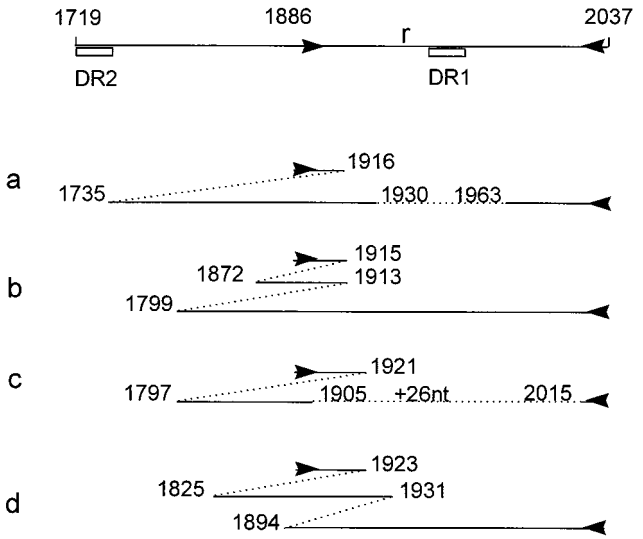

FIG. 6. Type 2 recombination joints showing evidence of two generations of cccDNA formation from linear DNA. The positions of the PCR primers (arrowheads), DR1, DR2, the repeat sequence **r**, and the sites of recombination are shown. Clones a and c each contain one type 2 precursor recombination joint (partial duplication in the cohesive end region) and one recombination joint from either a type 1 or type 2 precursor (deletion around the **r** sequence). Clones b and d each contain two joints from type 2 precursors. The presence of two separate recombination joints in these four clones can be explained by two or more generations of cccDNA synthesis from linear DNA. nt, nucleotides.

wild type	1921	TTCTGTAACCATGTATCTTTTTCACCTGTG	1950
a	1921	TTCTGTAACCATGTATCTTT TCACCTGTG	1950
b	1921	TT CTTTTTCACCTGTG	1950
c	1921	TTCT TCACCTGTG	1950
d (2)	1921	TTCTGTAACCATGTATCTT TTTTCACCTGTG	1950
e	1921	TTCTGTAACCATGTATCTTTTTC CACCTGTG	1950
f (2)	1921	TTCTGTAACCATGTATCTT CTTTTTCACCTGTG	1950
g	1921	TTCTGTAACCATGTATCTTTTTC CTTTTTCACCTGTG	1950
h	1st generation [hypothetical]		
	1921	TTCTGTAACCATGTATCTT CTTTTTCACCTGTG	1950
	2nd generation		
	1921	TTCTGTAACCATGTATCTTCTT CTTCTTTTTCACCTGTG	1950

FIG. 7. Sequence of clones with small deletions or insertions in the **r** region. For each clone, the nucleotide sequence upstream and downstream of the recombination site (indicated by a discontinuity in the sequence) is shown. In clones a to c, the recombinations produced deletions. In clones d to g, both sites of recombination occurred within the **r** sequence (underlined sequence), resulting in duplications. In clone h, two generations of cccDNA formation from linear DNA, designated 1st generation and 2nd generation, are hypothesized, resulting in a reduplication of the sequence CTT to generate four copies. All sequences contain the minus-strand DNA origin at nucleotide 1944 and therefore may be competent for further DNA synthesis.

and c, a type 2 precursor recombination joint (the downstream recombination site in the cohesive-end region) was combined with a deletion that could have been generated by circularization of either a type 1 or type 2 precursor, since the second downstream site is 3′ of the cohesive-end region. In clones b and d, both downstream recombination sites were located in the cohesive-end region. The presence of two independent recombination joints in the same cccDNA molecule can be explained by a process in which cccDNA was formed through two or more generations of linear DNA intermediates (33).

Using the same library of cloned PCR products, we next examined the positions of recombination joints located close to the ends of the linear type 1 precursors. We screened for these joints by colony hybridization to an oligonucleotide probe (probe 1, Fig. 4) that hybridized to 15 nucleotides covering the 10-nucleotide **r** region present at both of the ends of the putative type 1 linear precursor. We predicted that cccDNA produced from relaxed circular DNA or from type II precursors would retain this exact sequence intact and hybridize to the probe whereas cccDNAs formed from type 1 precursors would contain alterations of this sequence as a result of nonhomologous recombination and would not hybridize to the probe. Nine independent clones containing seven different examples of the expected types of alterations were identified from a total of about 300 WHV-positive clones screened (Fig. 7). Three alterations consisted of small deletions around the **r** region, and six consisted of 1-, 3-, or 7-bp duplications within the **r** sequence. The presence of these duplications indicates that the linear precursor of each of these cccDNAs was terminally redundant by at least 7 nucleotides, consistent with the structure of in situ primed linear DNA, which is terminally redundant by 10 nucleotides.

It is noteworthy that all nine clones contained the sequences known to be required for pregenome transcription and minus-strand synthesis. However, insertions and deletions at the 5′ end of the pregenome, resulting from transcription through these recombination joints, would be expected to favor the production of linear DNA from the pregenome transcripts by preventing plus-strand primer translocation. Production of in situ primed linear DNA (type I precursor) from these cccDNAs might result in a second generation of cccDNA from linear DNA, a process we previously described in DHBV as illegitimate replication. Only one example suggestive of such a second-generation cccDNA from in situ-primed linear DNA was found in a total of over 200 altered joint regions sequenced during this project (Fig. 7, clone h). This sequence could be explained by formation of a 3-nucleotide duplication during the first generation of cccDNA (Fig. 7, 1st generation) and a reduplication of this duplication during the second generation (Fig. 7, 2nd generation).

Existence of recombination joints with inserted sequences. The majority of recombination joints appeared to be the result of direct joining of two sites in type 1 or type 2 linear precursors. However, a significant number of joints contained insertions of as few as 1 or as many as 246 nucleotides of DNA derived from a different region of the viral genome or from unidentified sources (results not shown). Specifically, 11 of 81 joints described in Fig. 2 and 7 of 76 joints described in Fig. 5 contained such extra nucleotides between the two recombination sites listed. Usually, these inserted sequences consisted of 1 to 6 nucleotides whose origin could not be determined. Larger segments of DNA that could be identified as being of viral origin were seen as well, inserted in either direct or inverted orientations. Although the distribution of the recombination sites flanking these joints was consistent with recombination between two sites on a corresponding linear DNA precursor, we do not account for the presence of these extra sequences with our current models.

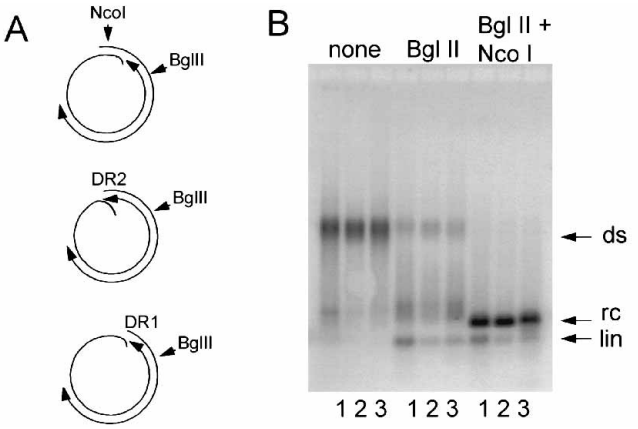

FIG. 8. Assay for circular and linear DNA replicative intermediates in WHV-infected liver. Replicative intermediates were extracted from the livers of three chronically infected woodchucks and digested with either *Bgl*II or *Bgl*II plus *Nco*I. (A) Strategy for assay of relaxed circular (top diagram) or linear (middle and bottom diagrams) DNAs by digestion to release characteristic double-stranded fragments derived from each species. (B) The undigested and digested DNAs were subjected to agarose gel electrophoresis and Southern blotting. The filter was probed with a riboprobe specific for plus-strand DNA in the region between the *Nco*I and *Bgl*II sites. Arrows indicate the position of digestion products derived from relaxed circular (rc) or in situ-primed linear DNA (lin). The position of partially double-stranded undigested DNA (ds) is also indicated.

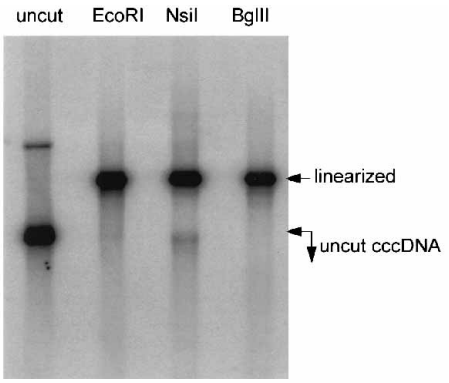

FIG. 9. Southern blot assay for deleted cccDNAs in chronically infected woodchuck liver. cccDNA extracted from the liver of woodchuck 1 in Fig. 8 was digested with *Eco*RI, *Nsi*I, or *Bgl*II and analyzed by agarose gel electrophoresis and Southern blot hybridization. cccDNAs with deletions (approximately 10% of the total) are detected as uncleaved DNA in the *Nsi*I digestion. Each digestion was internally controlled by determining the extent of cleavage of a heterologous DNA with a second hybridization probe (not shown). The controls indicated that each digestion was more than 99% complete.

Linear viral DNAs in infected liver. The structure of the altered cccDNA molecules we detected was consistent with their formation from linear partially double-stranded DNA in which one end of the DNA was located either at DR1 or at DR2. If such DNA existed in the liver of an infected woodchuck, it would not be distinguishable from circular, partially single-stranded viral DNA by electrophoretic mobility shift assay, since both forms would migrate in a common heterogeneous "smear," between fully single-stranded and fully double-stranded viral DNA in agarose gels. To detect these putative species of DNA, we used the strategy shown in Fig. 8A. We digested the replicative intermediate fractions from three different WHV-infected livers with *Bgl*II (nucleotide 2531) to produce a fragment derived from the naturally double-stranded end of either species of linear DNA. In the case of linear DNAs with one end located at DR1 (nucleotide 1935), the fragment would be 596 nucleotides in length, whereas the fragment derived from an end located at DR2 (1730) would be 801 nucleotides long. In addition to *Bgl*II digestion, we included double digestion with *Bgl*II and *Nco*I, which cuts at nucleotide 1741, just downstream of DR2 in the cohesive-end region of relaxed circular DNA. The quantity of the 790-bp fragment produced by the double digestion served as a measure of the relative amount of relaxed circular DNA. Digested DNA was analyzed by agarose gel electrophoresis and hybridization to a probe specific for these fragments (Fig. 8B). Digestion of the DNA with *Bgl*II alone generated the 596- but not the 801-nucleotide fragment, indicating that the replicative intermediate fraction contained in situ-primed linear DNAs (type I precursor) but no linear DNAs with ends located at DR2 (type 2 precursor). Doubly digested DNA produced the expected 790-nucleotide fragment derived from relaxed circular DNA. Quantitative measurements by phosphorimaging determined that the fractions of in situ-primed linear DNA in the three samples differed from each other, with a high of 20% to a low of 7%. The failure to find linear DNA with ends at DR2 may indicate that if this species did exist, it was not found in the cytoplasm of infected cells, from which replicative intermediates were isolated for our studies.

Frequency of cccDNAs with deletions in chronically infected livers. During the course of these experiments, we found that slight alterations in the PCR amplification conditions could generate large changes in the fraction of products derived from cccDNAs with deletions. Therefore, we carried out an experiment to estimate directly the frequency of such cccDNAs in chronically infected livers. Of 81 deletions, 57 (70%) characterized in the experiment described in Fig. 2 included the *Nsi*I site at nucleotide 1914. We therefore used the absence of the *Nsi*I site as an independent marker for deleted cccDNA molecules in a Southern blot analysis of the cccDNA purified from a chronically infected woodchuck (Fig. 9). We compared the fraction of molecules lacking the *Nsi*I site with those lacking the *Eco*RI and *Bgl*II sites, located in regions of the cccDNA that were not expected to suffer deletions. The results of this assay, shown in Fig. 8, revealed that approximately 10% of all cccDNA molecules detected were lacking the *Nsi*I site, compared with less than 1% that lacked either the *Eco*RI or *Bgl*II sites. These molecules resistant to *Nsi*I cleavage presumably lacked the *Nsi*I site because of deletions through the region encompassing the ends of in situ-primed linear DNA. This interpretation is supported by the presence of a heterogeneous smear of faster-migrating uncleaved cccDNA molecules in the *Nsi*I-digested sample. The results indicate that about 10% of cccDNA molecules in this sample suffered deletions characteristic of formation of cccDNA from in situ-primed linear DNA. This fraction compares with the linear fraction of total WHV replicative intermediates of around 20% isolated from this liver.

DISCUSSION

In this study, we examined the structures of cccDNAs present in WHV-infected livers for evidence of illegitimate replication. The defining characteristic of illegitimate replication is the synthesis of cccDNA from linear viral DNA by nonhomologous recombination (33). The expected feature of this type of synthesis, the presence of nonhomologous recombination joints in cccDNA, was clearly shown. The large majority of such recombination events resulted in the deletion of viral sequences. In addition, analysis of the distribution of a large number of recombination sites strongly suggested that the deletions did not arise through random recombination events within a circular DNA precursor but through recombination between two sites positioned to join the ends of one of two hypothetical linear DNA precursors. While the evidence for linear DNAs acting as cccDNA precursors is circumstantial, our hypotheses adequately explain almost all of the data. Remaining unexplained is the mechanism by which either viral or heterologous sequences were occasionally included between the upstream and downstream recombination joints.

The identities of the putative type 1 and type 2 precursors are not definitely known. The structure of predicted type I precursor corresponds closely to the structure of a previously described hepadnavirus DNA species produced by in situ priming (23) of plus-strand DNA (Fig. 1), and a species with the same structure was shown to be abundant (7 to 20%) in the replicative intermediate fraction from infected liver, in which in situ primed DNA is expected to occur. Moreover, in DHBV, mutations that enhance the formation of this species at the expense of relaxed circular DNA increased the frequency of cccDNAs formed by nonhomologous recombination accordingly. For these reasons, we strongly favor in situ-primed linear DNA as the type 1 precursor. No previously described species of linear DNA corresponds to the structure of the putative type 2 precursors.

Much of the cccDNA produced from linear viral species during WHV replication differed from that produced from linear DHBV DNA in the average size of the region that was

deleted during nonhomologous recombination. This difference was true of cccDNAs generated from both type 1 and type 2 precursors. The differences can be attributed to the fact that the deletions encompassing the 5′ terminus of the minus strand (upstream joints) were much more extensive than those found in DHBV cccDNA, in which the 5′ end of the minus strand was conserved in about 40% of the nonhomologous recombination events (33). This difference can be rationalized by postulating a mechanism for nonhomologous recombination that preferentially deleted the single-stranded portion of the 5′ end of minus-strand DNA, which is much more extensive in mammalian hepadnavirus replicative intermediates than it is in avian hepadnaviruses (18, 25, 28).

The hypothetical type 2 precursors, not detected in this study, were postulated to be linear molecules with terminal duplications of the region of the genome that lies between DR1 and DR2, the cohesive-end region. Moreover, the extensive deletions in cccDNAs synthesized from type 2 precursors were consistent with most type 2 precursors having a large single-stranded region at the 5′ end of the minus strand. Such molecules might have been derived from relaxed circular DNA by extension of the 3′ end of the minus strand to the 5′ end of the plus strand at DR2 (Fig. 3). This extension would necessarily have displaced the 5′ end of the minus strand from base pairing with the plus strand, resulting in a partially double-stranded linear molecule whose ends could be joined by nonhomologous recombination. The products of such recombination would be analogous to those produced from in situ-primed linear DNA but could be recognized by downstream recombination joints occurring between DR1 and DR2 (region 2, Fig. 2C). Of 81 unselected clones that we sequenced, 12 contained such distinctive joints; however, since type 2 precursors might circularize at recombination sites downstream of DR1, a larger number of clones may have been derived from such precursors.

We obtained further support for the occurrence of a terminally redundant linear precursor for these molecules when we analyzed the subset of cccDNA products formed by recombination close to the predicted ends of the proposed precursor (Fig. 5). The positions of the recombination sites producing such molecules indicated that the two ends of the putative linear precursor mapped very close to the positions of the 5′ ends of the two strands in relaxed circular DNA.

We did not detect the putative type 2 linear precursor in the replicative intermediate fraction of viral DNA. However, such molecules might have been formed in the nucleus from relaxed circular DNA after release from the capsid during cccDNA amplification and therefore would not have been detected by the method we used, in which nuclei were removed prior to DNA extraction. However, we also failed to detect such molecules in a nuclear fraction consisting of non-protein-bound DNAs (data not shown). The meaning of this result is not clear, however, since after uncoating, linear DNAs may be quickly converted to cccDNAs. Such putative linear DNAs could represent authentic intermediates in cccDNA formation from rcDNA, or they could be the result of aberrant events.

An alternative model for the origin of type 1 and type 2 precursors might be suggested by the work of Wang and Rogler (31), who detected preferred cleavage of WHV DNA by purified topoisomerase I in vitro at a cluster of sites around DR1 and DR2 in relaxed circular DNA to generate linear DNAs with a reactive molecule of topoisomerase I covalently bound at one of the ends. They proposed that such reactive molecules may be intermediates in illegitimate recombination with cellular DNAs and other viral DNAs, and in support of this proposal, they demonstrated by sensitive PCR methods that such recombination events could occur in vitro, albeit at unknown efficiency. The Wang and Rogler model has been proposed by Kew et al. (12) as a mechanism to account for certain WHV genomes that they cloned from virus particles and that contained deletions similar to those we report here. We believe, however, that several features of our data are inconsistent with reactive linear molecules produced by this mechanism acting as precursors of the cccDNAs we describe. First, the cleavage site at DR1, which would generate the type 1 linear precursor, is not positioned to allow topoisomerase-mediated circularization of the linear DNA, since the topoisomerase I would be bound to the 3′ end of the cleaved plus strand, a sequence which is almost always deleted during the nonhomologous recombination. Second, cleavage at DR2 in the minus strand would not be capable of generating type 2 linear precursors, since there would be no terminal duplication of the cohesive-end region. Third, the downstream recombination sites involved in the circularization of type 1 and type 2 precursors do not cluster around the restricted preferred cleavage sites mapped by Wang and Rogler but are much more widely dispersed downstream of these sites. In fact, cleavage at all but a handful of these sites would not be expected to linearize the relaxed circular DNA. Fourth, although Wang and Rogler showed that topoisomerase I-mediated illegitimate recombination between relaxed circular DNA and heterologous double-stranded DNA occurred at regions of short sequence homology (3 to 5 bp in their experiments), our recombinations showed little evidence of homology over that expected by chance (Table 1). We therefore believe that our data are inconsistent with a significant role for topoisomerase I acting on relaxed circular DNA in the generation of these recombination joints.

In spite of the preference for extensive deletions associated with the formation of cccDNA from linear DNA by nonhomologous recombination, a small fraction of the cccDNA molecules contained all known sequences required in *cis* for further participation in DNA synthesis (1, 10, 21, 22, 30). Two classes of potentially functional molecules were observed in which nonhomologous recombination produced small deletions or insertions either (i) in and around the **r** sequence or (ii) within the cohesive-end region. In the former case, an intact pregenome promoter would be expected to produce a pregenome with a corresponding deletion or insertion within the **r** sequence that would not interfere with transcription or packaging but would be expected to result in the preferred synthesis of linear DNA (16, 23), initiating one or more cycles of illegitimate replication. One example of a possible second-generation cccDNA molecule of this class, which contained four copies of a trinucleotide sequence of the **r** region was observed (Fig. 7, clone h).

In the case of cccDNA containing small deletions or duplications in the cohesive-end region, it is not known to what extent the molecules we detected would be functional for further DNA replication. Many of these molecules presumably retained an intact upstream copy of the pregenome promoter followed by various lengths of sequence derived from the cohesive-end region preceding DR1, the packaging signal sequence ε (10), and all four essential open reading frames intact. It is difficult to predict from theory whether progeny DNA could be synthesized from such molecules, since most of these molecules retained an intact precore translation start codon, which would be expected to interfere with pregenome packaging (4, 19). Nevertheless, we found several examples of such molecules that contained sequence evidence of more than one generation of replication through type 2 linear DNA. This conclusion is inferred from the presence of more than one recombination joint within the cohesive-end region, as seen in

four separate sequences, suggesting sequential rounds of cccDNA formation from linear precursors. Taken together, our studies provide modest evidence that illegitimate replication of WHV DNA through two kinds of linear DNA intermediates occurs at low levels in WHV-infected woodchuck liver. The fraction of cells in the liver which carried out this mode of DNA replication could not be estimated from these experiments.

The deletions that we have described in cccDNAs from chronically infected woodchuck liver bear a striking similarity to deletions that have been reported to occur in hepadnavirus genomes found in virus particles in the blood of infected humans (5–8) and woodchucks (12). These similarities suggest that similar processes of defective genome formation may be occurring in chronically infected human and woodchuck liver. However, these previous reports raise the question whether the cccDNAs we observed were derived from WHV genomes that suffered deletions generated during the process of virion DNA synthesis or whether altered viral genomes reported in previous studies were derived from RNA pregenomes transcribed from preexisting deleted cccDNAs similar to those we describe here. No evidence addresses this question directly. We have used analogies between the structures of the WHV cccDNAs with deletions and those known to be derived from linear DHBV DNA to argue that the cccDNA population described here was derived from linear DNA molecules produced by conventional mechanisms. If this is the case, a subset of these cccDNAs may be capable of being expressed in such a way as to generate progeny viral genomes with corresponding deletions. This hypothesis is consistent with the findings that although sequences that are strictly required for minus-strand DNA synthesis (DR1, ε) were frequently deleted in the cccDNA we described, those sequences were always present in the viral genomes found in blood borne particles (5, 6, 8). This fact argues that the genomes found in virus particles were derived from a functional subset of the available sequences in the pool of defective cccDNA.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants CA-42542, AI-18641, and CA-57425.

REFERENCES

1. **Bartenschlager, R., M. Junker-Niepmann, and H. Schaller.** 1990. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. J. Virol. **64:**5324–5332.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
3. **Condreay, L. D., T. T. Wu, C. E. Aldrich, M. A. Delaney, J. Summers, C. Seeger, and W. S. Mason.** 1992. Replication of DHBV genomes with mutations at the sites of initiation of minus- and plus-strand DNA synthesis. Virology **188:**208–216.
4. **Enders, G. H., D. Ganem, and H. E. Varmus.** 1987. 5′-terminal sequences influence the segregation of ground squirrel hepatitis virus RNAs into polyribosomes and viral core particles. J. Virol. **61:**35–41.
5. **Feitelson, M. A., L. X. Duan, J. Guo, and B. S. Blumberg.** 1995. X region deletion mutants associated with surface antigen-positive hepatitis B virus infections. Gastroenterology **108:**1810–1819.
6. **Feitelson, M. A., L. X. Duan, J. Guo, B. Sun, J. Woo, K. Steensma, N. Horiike, and B. S. Blumberg.** 1995. X region deletion variants of hepatitis B virus in surface antigen-negative infections and non-A, non-B hepatitis. J. Infect. Dis. **172:**713–722.
7. **Feitelson, M., L. X. Duan, N. Horiike, and M. Clayton.** 1991. Hepatitis B X open reading frame deletion mutants isolated from atypical hepatitis B virus infections. Hepatology **13:**S58–S60.
8. **Feitelson, M., L. Lega, J. Guo, M. Resti, M. E. Rossi, C. Azzari, B. S. Blumberg, and A. Vierucci.** 1994. Pathogenesis of posttransfusion viral hepatitis in children with beta-thalessemia. Hepatology **19:**558–568.
9. **Horwich, A. L., K. Furtak, J. C. Pugh, and J. Summers.** 1990. Synthesis of hepadnavirus particles containing replication-defective duck hepatitis B virus genomes in cultured Huh-7 cells. J. Virol. **64:**642–650.
10. **Junker-Niepmann, M., R. Bartenschlager, and H. Schaller.** 1990. A short *cis*-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J. **9:**3389–3396.
11. **Kajino, K., A. R. Jilbert, J. Saputelli, C. E. Aldrich, J. Cullen, and W. S. Mason.** 1994. Woodchuck hepatitis virus infections: very rapid recovery after a prolonged viremia and infection of virtually every hepatocyte. J. Virol. **68:**5792–5803.
12. **Kew, M. C., R. H. Miller, H. S. Chen, B. C. Tennant, and R. H. Purcell.** 1993. Mutant woodchuck hepatitis virus genomes from virions resemble rearranged hepadnaviral integrants in hepatocellular carcinoma. Proc. Natl. Acad. Sci. USA **90:**10211–10215.
13. **Kodama, K., N. Ogasawara, H. Yoshikawa, and S. Murakami.** 1985. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: evolutional relationship between hepadnaviruses. J. Virol. **56:**978–986.
14. **Lien, J. M., C. Aldrich, and W. S. Mason.** 1986. Evidence that a capped oligonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J. Virol. **57:**229–236.
15. **Lien, J. M., D. Petcu, C. Aldrich, and W. S. Mason.** 1987. Initiation and termination of duck hepatitis B virus DNA synthesis during virus maturation. J. Virol. **61:**3832–3840.
16. **Loeb, D., R. Hirsch, and D. Ganem.** 1991. Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. EMBO J. **10:** 3533–3540.
17. **Mason, W. S., C. Aldrich, J. Summers, and J. M. Taylor.** 1982. Asymmetric replication of duck hepatitis B virus DNA in liver cells (free minus strand DNA). Proc. Natl. Acad. Sci. USA **79:**3997–4001.
18. **Mason, W. S., G. Seal, and J. Summers.** 1980. A virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. J. Virol. **36:**829–836.
19. **Nassal, M., M. Junker-Niepmann, and H. Schaller.** 1990. Translational inactivation of RNA function: discrimination against a subset of genomic transcripts during HBV nucleocapsid assembly. Cell **63:**1357–1363.
20. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science **232:**477–484.
21. **Seeger, C., and J. Maragos.** 1989. Molecular analysis of the function of direct repeats and a polypurine tract for plus-strand DNA priming in woodchuck hepatitis virus. J. Virol. **63:**1907–1915.
22. **Seeger, C., and J. Maragos.** 1990. Identification and characterization of the woodchuck hepatitis virus origin of DNA replication. J. Virol. **64:**16–23.
23. **Staprans, S., D. Loeb, and D. Ganem.** 1991. Mutations affecting hepadnavirus plus-strand synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. J. Virol. **65:**1255–1262.
24. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell **29:**403–415.
25. **Summers, J., A. O'Connell, and I. Millman.** 1975. Genome of hepatitis B virus: restriction enzyme cleavage and structure of the DNA extracted from Dane particles. Proc. Natl. Acad. Sci. USA **72:**4597–4601.
26. **Summers, J., P. Smith, and A. L. Horwich.** 1990. Hepadnaviral envelope proteins regulate amplification of covalently closed circular DNA. J. Virol. **64:**2819–2824.
27. **Summers, J., P. Smith, M. Huang, and M. Yu.** 1991. Regulatory and morphogenetic effects of mutations in the envelope proteins of an avian hepadnavirus. J. Virol. **65:**1310–1317.
28. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc. Natl. Acad. Sci. USA **75:**4533–4537.
29. **Tuttleman, J., C. Pourcel, and J. Summers.** 1986. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. Cell **47:**451–460.
30. **Wang, G., and C. Seeger.** 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J. Virol. **67:**6507–6512.
31. **Wang, H.-P., and C. E. Rogler.** 1991. Topoisomerase I-mediated integration of hepadnaviral DNA in vitro. J. Virol. **65:**2381–2392.
32. **Wu, T.-T., L. Coates, C. E. Aldrich, J. Summers, and W. S. Mason.** 1990. In hepatocytes infected with duck hepatitis B virus, the template for viral RNA synthesis is amplified by an intracellular pathway. Virology **175:**255–261.
33. **Yang, W., and J. Summers.** 1995. Illegitimate replication of linear hepadnaviral DNA through nonhomologous recombination. J. Virol. **69:**4029–4036.