JOURNAL OF VIROLOGY, Dec. 1982, p. 852–863
0022-538X/82/120852-12$02.00/0

Vol. 44, No. 3

Novel Forms of Woodchuck Hepatitis Virus DNA Isolated from Chronically Infected Woodchuck Liver Nuclei

CHARLES E. ROGLER* AND JESSE SUMMERS

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Received 16 February 1982/Accepted 7 August 1982

We cloned several unique forms of woodchuck hepatitis virus, a DNA virus closely related to hepatitis B virus, from a chronically infected woodchuck liver. Each of the three clones contained more than two genome equivalents of viral sequences with extensive rearrangements and no detectable cellular sequences. From the frequency by which they were isolated from a library of recombinant clones, we estimate that they are present in approximately one copy per cell. Of a total of 11 sites at which rearrangements were mapped in the clones, 10 occurred between segments of opposite polarity, and 1 occurred between segments of the same polarity. The possible significance of these findings to the persistence of virus production in infected cells is discussed.

Woodchuck hepatitis virus (WHV) has provided an animal model for the study of hepatic infections by the novel family of DNA viruses, of which hepatitis B virus is the prototype (13). WHV infections are frequently found in woodchucks which are captured from the wild, and a high proportion of these woodchucks develop hepatocellular carcinoma when kept in captivity. Every case of hepatocellular carcinoma observed in captive woodchucks so far has been associated with either portal hepatitis or chronic active hepatitis (11).

Due to the close correlation between chronic active hepatitis and hepatocellular carcinoma, studies have been conducted to investigate the role of WHV DNA in the occurrence of hepatocellular carcinoma. WHV DNA has been found to be integrated in the nuclear DNA of most hepatocellular carcinomas (J. Summers, W. Ogston, G. Jonak, S. Astrin, G. Tyler, and R. Snyder, unpublished data), and the structures of two of these integrated genomes have been determined (8). To investigate further the significance of integrated WHV DNA in hepatitis infections in general, we looked for integrated WHV DNA in chronically infected liver tissue which showed no signs of tumor formation.

The approach we took to investigate the problem was to generate a library of clones containing nuclear DNA from a chronically infected woodchuck liver, using lambda phage (Charon 30) as a vector. Our results yielded two classes of recombinants which contained WHV DNA sequences. The first class of clones are the novel forms of WHV which are described in this paper and which occur at approximately one copy per cell. The second class of clones comprised WHV DNA sequences covalently integrated into liver nuclear DNA. The structure of the integrated clones, which also occur at a frequency of approximately one copy per cell, will be reported in a second paper which is now in preparation.

MATERIALS AND METHODS

Tissues and strains. Woodchuck liver HW197 was obtained from a chronically infected woodchuck (*Marmota monax*) from a colony housed at the Penrose Research Laboratory, Zoological Society of Philadelphia, during a liver biopsy.

Lambda phage, vector Charon 30 (9), was kindly supplied by F. R. Blattner and was grown in *Escherichia coli* LE392 by the method of Enquist and coworkers (5, 6).

DNA extractions, partial digestion. Nuclear DNA was extracted from liver HW197 as previously described (8). Lambda phage Charon 30 DNA and DNA from recombinant clones were isolated by the method of VandeWoude et al. (14), in which phage are purified by pelleting through a 40% glycerol step gradient. *Eco*RI-digested arms of Charon 30 were isolated, free of stuffer fragments, by agarose gel electrophoresis through 1% Sea Plaque agarose (SeaKem) and extraction from the agarose (4). Liver nuclear DNA (HW197) was partially digested before cloning as follows: 50 μg of DNA in *Eco*RI buffer was digested with 16 U of *Eco*RI enzyme (New England Biolabs; 1 U digests 1 μg of λ DNA in 15 min) for 5 min before adding sodium dodecyl sulfate to 0.1%. Labeled lambda *Hin*dIII fragments were included in the digestion mixture as a control for partial digestion. The degree of partial digestion was determined by running a small portion of the sample on an agarose gel, drying the gel on 3MM paper, and autoradiographing. The faint appearance of new lambda bands due to the partial digestion of the labeled *Hin*dIII fragments indicated the degree of *Eco*RI partial digestion. In the sample used for cloning, the digestion had proceeded to approximately 10

to 15% completion. HW197 partial digest DNA fragments between 4.4 and 17 kilobases (kb) were isolated after preparative agarose gel electrophoresis through a 1% Sea Plaque agarose gel. This DNA was used for cloning in the *Eco*RI arms of Charon 30.

Cloning and Southern blot analysis. Partial *Eco*RI-digested HW197 DNA was ligated to the vector arms of Charon 30 and packaged in vitro by slight modifications of the method of Blattner et al. (2) as previously described (8). Recombinants containing WHV sequences were screened by the method of Benton and Davis (1).

Southern blot analysis of genomic DNA and recombinant phage DNA was carried out as previously described (12). The hybridization of Southern blots of recombinant DNA was carried out for only 30 min at 37°C with approximately 1×10^7 cpm of ^{32}P-labeled DNA prepared from cloned WHV DNA (4) using DNA polymerase I (12).

End-labeling and partial digestion. Recombinant phage DNA (10 μg) was digested with *Sal*I and purified by adding 0.1% sodium dodecyl sulfate, 10 mM Na_2 EDTA, and 50 mM NaCl, extracting one time with phenol, and ethanol precipitating. The DNA was taken up in water to a concentration of 1.1 mg/ml. The method of Smith and Birnstiel (10) was used for end-labeling, and the final conditions for the T4 polymerase end-labeling of *Sal*I-digested DNA were as follows (in a 5-μl reaction): 550 μg of DNA per ml, 66 mM potassium acetate, 33 mM Tris acetate (pH 7.8), 10 mM magnesium acetate, 0.5 mM dithiothreitol, 100 μg of bovine serum albumin per ml, 0.1 mM dCTP, 0.1 mM dGTP, 0.1 mM [^{32}P]dTT (Amersham Corp.; specific activity, 410 Ci/mmol), and 0.6 U of T4 polymerase (Bethesda Research Laboratories). The reaction was incubated for 15 min at 11°C, stopped by the addition of 0.1% sodium dodecyl sulfate, 20 mM EDTA, and 0.1 M NaCl, extracted one time with phenol, and precipitated with ethanol. The end-labeled DNA (4.56×10^5 cpm incorporated) was taken up in water, and *Xho*I buffer was added, followed by *Xho*I digestion. Approximately 0.4 μg of end-labeled *Xho*I-digested DNA containing 60,000 cpm was mixed with 2.6 μg of calf thymus DNA and digested with 1 U of the approximate restriction enzyme for partial digestions lasting 10 min. The partial digestion products were separated on 0.8% agarose gels. The gels were then dried on 3MM paper and autoradiographed with Kodak X-Omat film for 16 to 24 h.

Restriction endonuclease analysis of complete digests of cloned DNA in Charon 30 was carried out by standard methods, using agarose gel electrophoresis, visualization, and photography of bands under UV light after staining with ethidium bromide.

Electron microscopy. Electron microscopy of cloned DNA and heteroduplex analysis were carried out as previously described (8).

RESULTS

Mapping and cloning of viral DNA from woodchuck liver HW197. The complete nucleotide sequence of the WHV genome was recently published (7). Since variations in the restriction endonuclease map of the virus can occur between isolates obtained at different locations, we decided to map the free viral DNA present in the nuclear DNA from liver HW197 by the Southern blot method. Various enzymes which were predicted to cut the viral sequences one time, according to the published sequence, were mapped with respect to the single *Eco*RI site, using double digests, agarose gel electrophoresis, and hybridization of the Southern blot with WHV DNA cloned in pBR322 (Fig. 1). Two unexpected results were obtained from this mapping experiment: first, the viral DNA of HW197 contained an additional *Xba*I site at approximately 1,180 base pairs (bp) compared with the published WHV sequence. Second, a single *Sac*I site was located at 2,480 bp, very close to the *Bgl*II site (Fig. 1), instead of close to the *Hin*dIII site where it occurs in the prototype. The other restriction endonuclease sites mapped in the places predicted by the published sequence. The circular map of WHV 197 is shown in Fig 2. Integrated viral DNA was not detected in the blot of HW197 DNA (Fig. 1), indicating that integrated DNA was not present in the same genomic location in a majority of the cells. This result is different from that observed with human liver tissue infected with hepatitis B virus, in

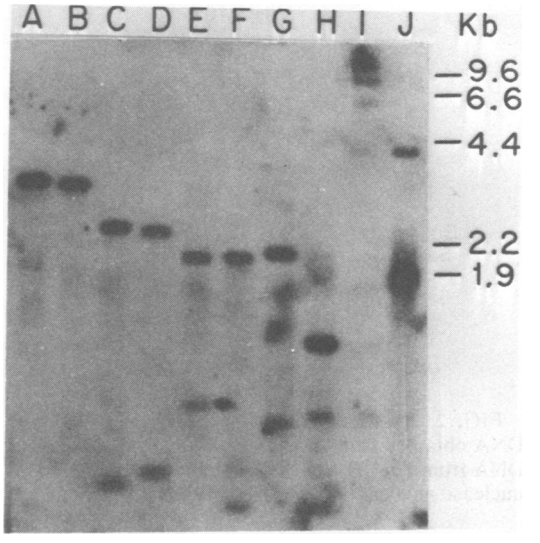


FIG. 1. Restriction endonuclease mapping of vegetative WHV DNA from chronically infected HW197 liver nuclear DNA by electrophoresis through a 1% agarose gel, Southern blotting, and hybridization with cloned WHV DNA. Nuclear DNA (5 μg) digested with appropriate enzymes was run in each lane. Lane A, *Eco*RI, one band hybridized which was 3.3 kb of full-length linear WHV; lane B, *Eco*RI and *Bam*HI; lane C, *Eco*RI and *Bgl*II; lane D, *Eco*RI and *Sac*I; lane E, *Eco*RI and *Hin*dIII; lane F, *Eco*RI and *Xba*I; lane G, *Hin*dIII and *Bam*HI; lane H, *Hin*dIII and *Xba*I; lane I, λ *Hin*dIII digest standards became visible upon longer exposure of autoradiogram; lane J, undigested nuclear DNA.

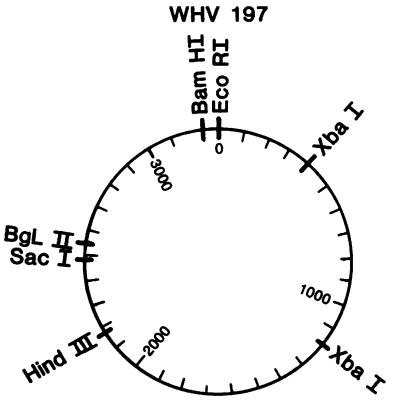


FIG. 2. Restriction endonuclease map of WHV 197 DNA obtained from Southern blot analysis of nuclear DNA from liver HW197 (Fig. 1) and restriction endonuclease analysis of cloned HW197 DNA.

which Southern blots of genomic DNA have revealed specific high-molecular-weight bands which contain hepatitis B virus integrated in genomic DNA (3).

Viral DNA in the genomic DNA preparation from HW197 liver was linearized with *Hin*dIII, isolated by agarose gel electrophoresis and phenol extraction of 3.2- to 3.4-kb DNA fragments, and cloned into the *Hin*dIII site of lambda phage, Charon 30. Restriction mapping of the clone, isolated by screening using the Benton-Davis method (1), confirmed the map initially obtained by the Southern blot method (data not shown).

Rationale and results of cloning partial digests of HW197 DNA. Our initial goal was to screen a genomic library of recombinant clones containing woodchuck DNA for the presence of integrated WHV sequences. We prepared a partial *Eco*RI digest of HW197 DNA and separated fragments between 4.4 and 17 kb by agarose gel electrophoresis. These fragments were ligated into the *Eco*RI arms of Charon 30. This procedure afforded a triple selection against the cloning of free viral DNA present in the HW197 nuclear DNA preparation. First, the minimum size piece of DNA which is clonable at the *Eco*RI site of lambda Charon 30 arms is 4.4 kb, and the free viral DNA is only 3.3 kb. Second, DNA smaller than 4.4 kb was eliminated by extracting partially digested DNA between 4.4 and 17 kb from the agarose gel. Third, the *Eco*RI partial digestion was approximately 15% complete, thereby eliminating 85% of the free viral DNA due to the lack of digestion at the single *Eco*RI.

Approximately 2×10^6 recombinant phage were prepared by in vitro packaging of recombinant DNA and were screened. Assuming that an average of 10 kb of HW197 DNA was present per recombinant, we calculated that this screening encompassed approximately four genome equivalents of HW197 DNA (approximately 5×10^6 kb of DNA per woodchuck genome). A total of five clones were isolated which contained viral sequences. Two of these clones contained viral sequences integrated in genomic DNA, and the structure of these clones will be reported in a separate paper. The other three clones contained what we have termed "novel" forms of WHV, containing exclusively viral sequences. We will present three separate lines of evidence to support the structures which we propose for these novel forms, including (i) restriction endonuclease mapping using partial digests of end-labeled DNA, (ii) heteroduplex analysis, and (iii) confirmation of the maps constructed from the first two lines of evidence by observing fragments produced by single and double restriction endonuclease digestions of the cloned DNA using agarose gel electrophoresis.

Southern blot analysis of clones HW197-1, -3, and -4. The initial *Eco*RI digestion of clones HW197-1, -3, and -4 revealed that each clone contained three *Eco*RI fragments (Fig. 3).

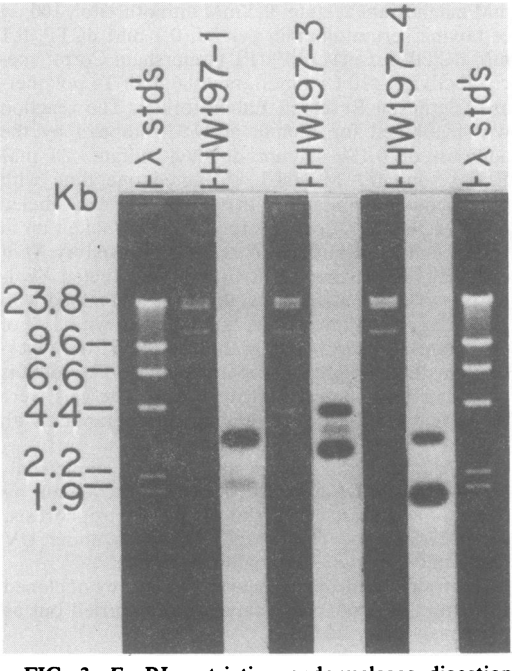


FIG. 3. *Eco*RI restriction endonuclease digestion of clones HW197-1, -3, and -4 and analysis of WHV DNA sequences by Southern blotting and hybridization with WHV [^{32}P]DNA probe. The left lane of each pair is the *Eco*RI restriction fragment pattern of cloned DNA in lambda Charon 30. Vector arms are bands at 15, 23, and 38 kb. The right lane of each pair is the autoradiogram of the Southern blot of each *Eco*RI profile after hybridization with WHV ^{32}P probe DNA.

Southern blot analysis of the clones hybridized with WHV DNA unexpectedly revealed that all of the cloned *Eco*RI fragments contained viral sequences. This indicated that the clones contained more than a single genome length of viral sequences because the integration of a single genome-length fragment would be expected to yield only two *Eco*RI fragments containing viral sequences. Very weak hybridization was observed for the 2.1-kb fragment of HW197-1 and the 3.5-kb fragment of HW197-4.

To determine whether any of these clones contained cellular sequences, we performed the following experiment. Liver DNA from a normal uninfected woodchuck was digested with *Eco*RI, and the fragments were blotted onto nitrocellulose paper after agarose gel electrophoresis. Separate lanes of the blot were hybridized with probes from each of the clones. In four replicate experiments, none of the clones was observed to hybridize to the normal liver DNA, indicating that the clones contained no cellular DNA sequences or at most a very small amount (~0.3 kb or less). Control digests of calf thymus DNA containing one genome equivalent of WHV DNA standard were run alongside the test lanes in several experiments, and they clearly revealed a hybridizing band at 3.3 kb, as was expected for free WHV DNA. Clones HW197-2 and HW197-5, which contain WHV integrated in genomic DNA, each produced strong bands upon hybridization to normal woodchuck DNA (manuscript in preparation).

Partial digests of novel-form clones HW197-1, -3, and -4. The rationale for restriction mapping using end-labeling and partial digestion is outlined in Fig. 4. The success of the method depended on the fact that none of the clones contained either an *Xho*I or a *Sal*I site. Therefore, it was possible to end label each clone at the *Sal*I site in the right arm of Charon 30 (Fig. 4B). This procedure produced three labeled fragments, the clone plus 1 kb of right-arm sequences labeled on the right end, a 1.0-kb internal *Sal*I fragment from the right arm of Charon 30 labeled on both ends, and a 10.2-kb fragment of the right arm, labeled on the left end. *Xho*I digestion of the entire preparation was carried out after end-labeling with T4 DNA polymerase (Fig. 4B) and resulted in the removal of the label from the left end of the 10.2-kb fragment. The final product contained three labeled fragments: (i) the cloned DNA labeled at the right end, (ii) a 1.0-kb internal *Sal*I fragment, and (iii) a 0.5-kb *Sal*I-*Xho*I fragment. The latter two fragments served as standards and did not interfere with the mapping of partial digests, all of which were greater than 1.0 kb.

End-labeled DNA of clone HW197-1 was partially digested with each of the enzymes which cut HW197 one time and one enzyme that cuts twice, and the partial digests were separated on

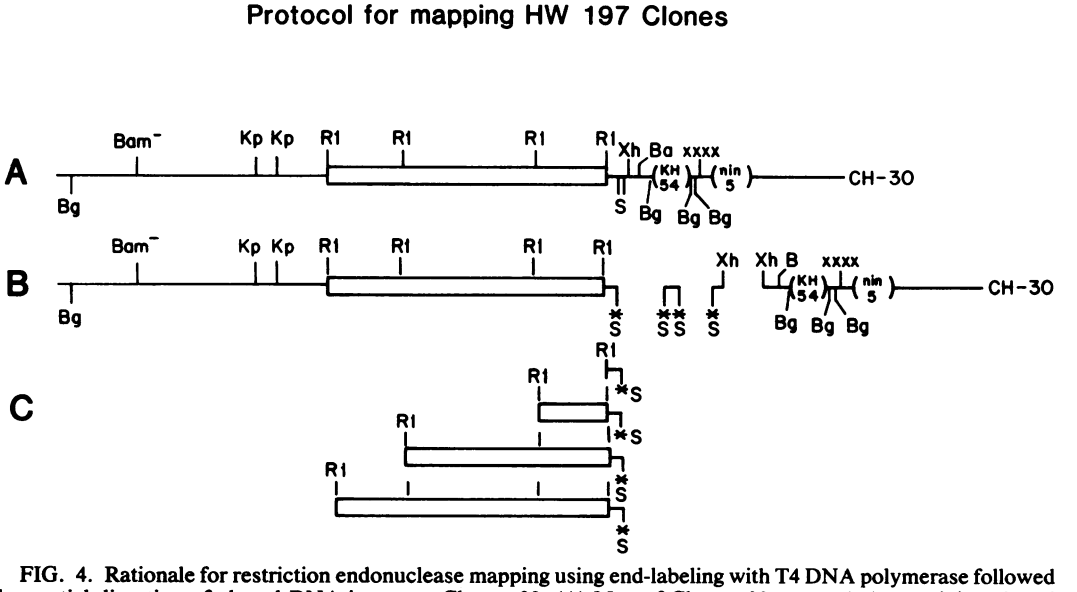


FIG. 4. Rationale for restriction endonuclease mapping using end-labeling with T4 DNA polymerase followed by partial digestion of cloned DNA in vector Charon 30. (A) Map of Charon 30 vector (—) containing cloned DNA (▭) with multiple *Eco*RI sites and no *Sal*I (S) or *Xho*I (Xh) sites. (B) ^{32}P-end-labeled fragments (*) produced after *Sal*I digestion, T4 polymerase end-labeling, and *Xho*I digestion. (C) Labeled fragments produced upon partial digestion of end-labeled DNA with *Eco*RI. The partial digestion products were separated by ararose gel electrophoresis, and the gels were autoradiographed. The autoradiographs of dried gels are shown in Fig. 5.

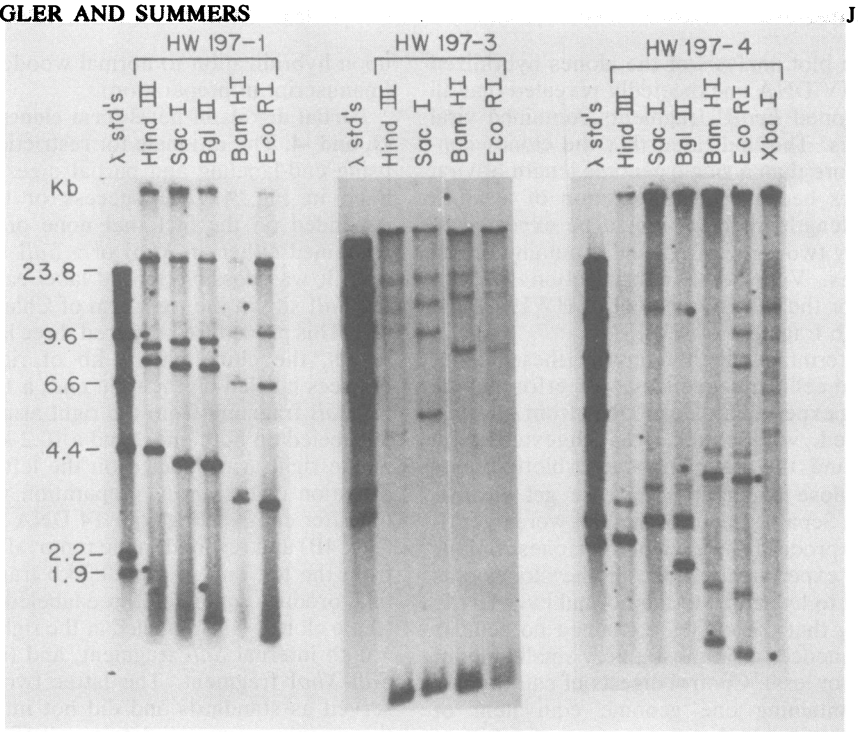


FIG. 5. Autoradiograms of dried agarose gels containing the ^{32}P-end-labeled partial digestion products of clones HW197-1, -3, and -4 digested with the indicated enzymes and separated by agarose gel electrophoresis. The band at 23.8 kb in all of the lanes is undigested labeled DNA. The largest *Eco*RI band in the right lane represents the entire piece of cloned DNA plus 1 kb of the right arm of Charon 30 vector. The restriction sites on the clone are determined by reading down the ladder of bands in order from the largest to the smallest band. The smallest *Eco*RI is the 1.0-kb labeled fragment from the right arm of Charon 30 (see protocol, Fig. 4).

an agarose gel (Fig. 5). The enzyme digests were organized on the gel in the order by which the restriction sites appear in the virus, beginning with the *Hin*dIII site and proceeding through the *Eco*RI and *Xba*I sites (*Xba*I is a two-cut enzyme and not all of the data for *Xba*I mapping are included). This procedure allowed us to determine the orientation of viral sequences in the clones. In Fig. 5, HW197-1, the largest *Eco*RI-generated fragment (the 23.8-kb band is undigested material from partial digestion), represents the extreme left-hand *Eco*RI site of the cloned DNA (clones inserted in Charon 30 at the *Eco*RI site), and the order of internal restriction sites in the clone can be determined by proceeding down from the largest *Eco*RI fragment (which represents the entire clone plus 1.0 kb of the right arm of Charon 30), reading the restriction sites in the "ladder" of fragments to the smallest *Eco*RI fragment, which is the right-hand *Eco*RI site located on the 1.0-kb fragment of the right arm of Charon 30. By this analysis method, it became evident that the restriction sites on HW197-1 appeared to be colinear with viral restriction sites from the left-hand *Eco*RI site through the *Hin*dIII site (bands between 9.6 and 8 kb; Fig. 5), following which there is an inverted copy of the same region of the viral genome through the viral *Hin*dIII site to the *Eco*RI site (bands between 8 and 6.0 kb, Fig. 5). These sequences are followed by a direct repeat of the same sequences, as is revealed by the set of bands between 4.4 and 2.2 kb (Fig. 5), which occur in the same orientation as the preceding set of bands. When one calculates the sizes of the partial digest fragments, the map of restriction sites obtained reveals that most of the sites are spaced the same distances apart from each other as they appear in the intact cloned viral DNA. This indicated to us that they represented stretches of viral DNA and that there was a large inverted repeat in viral sequences around the *Hin*dIII site followed by a direct repeat of viral sequences. A new *Bgl*II site, not present in HW197 viral DNA, was observed 400 bp from the right-hand *Eco*RI site (Fig. 5; *Bgl*II HW197-1). This site lies in viral sequences since it is flanked by *Xba*I sites that are in positions relative to the *Eco*RI site expected from the viral map (Fig. 2); therefore, we conclude that this site must have arisen by mutation. This mutant *Bgl*II site is not present in other copies of this

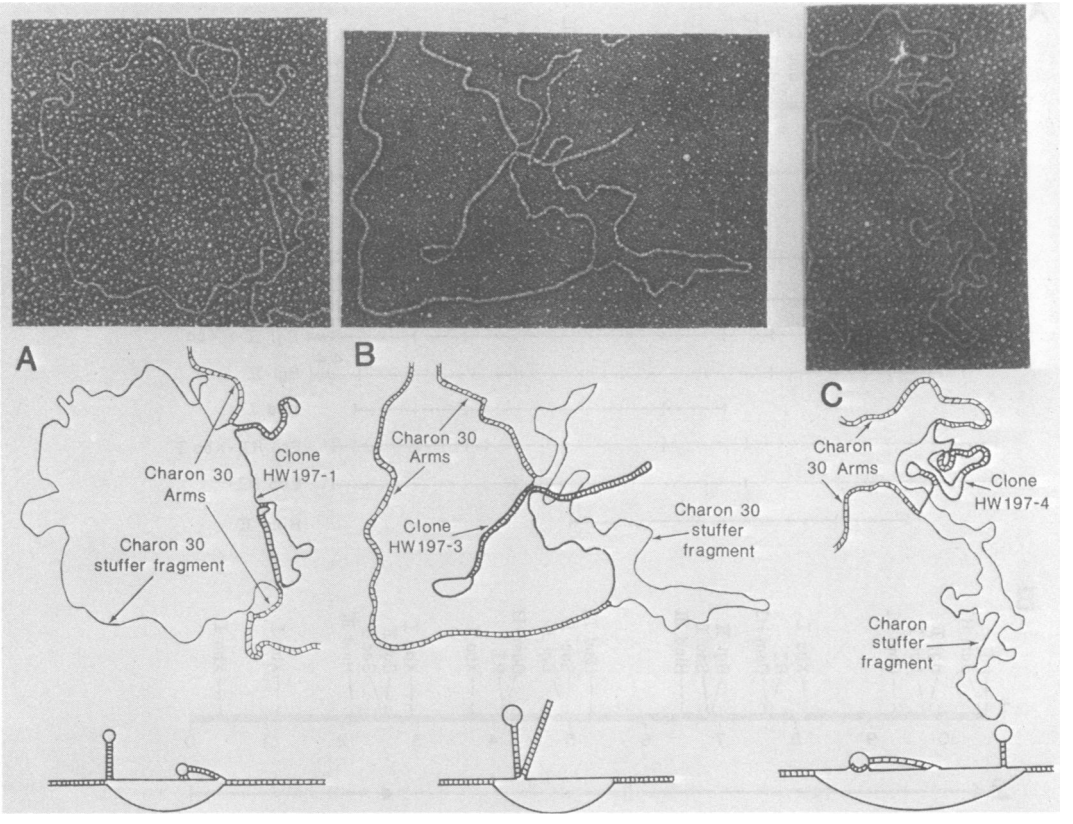
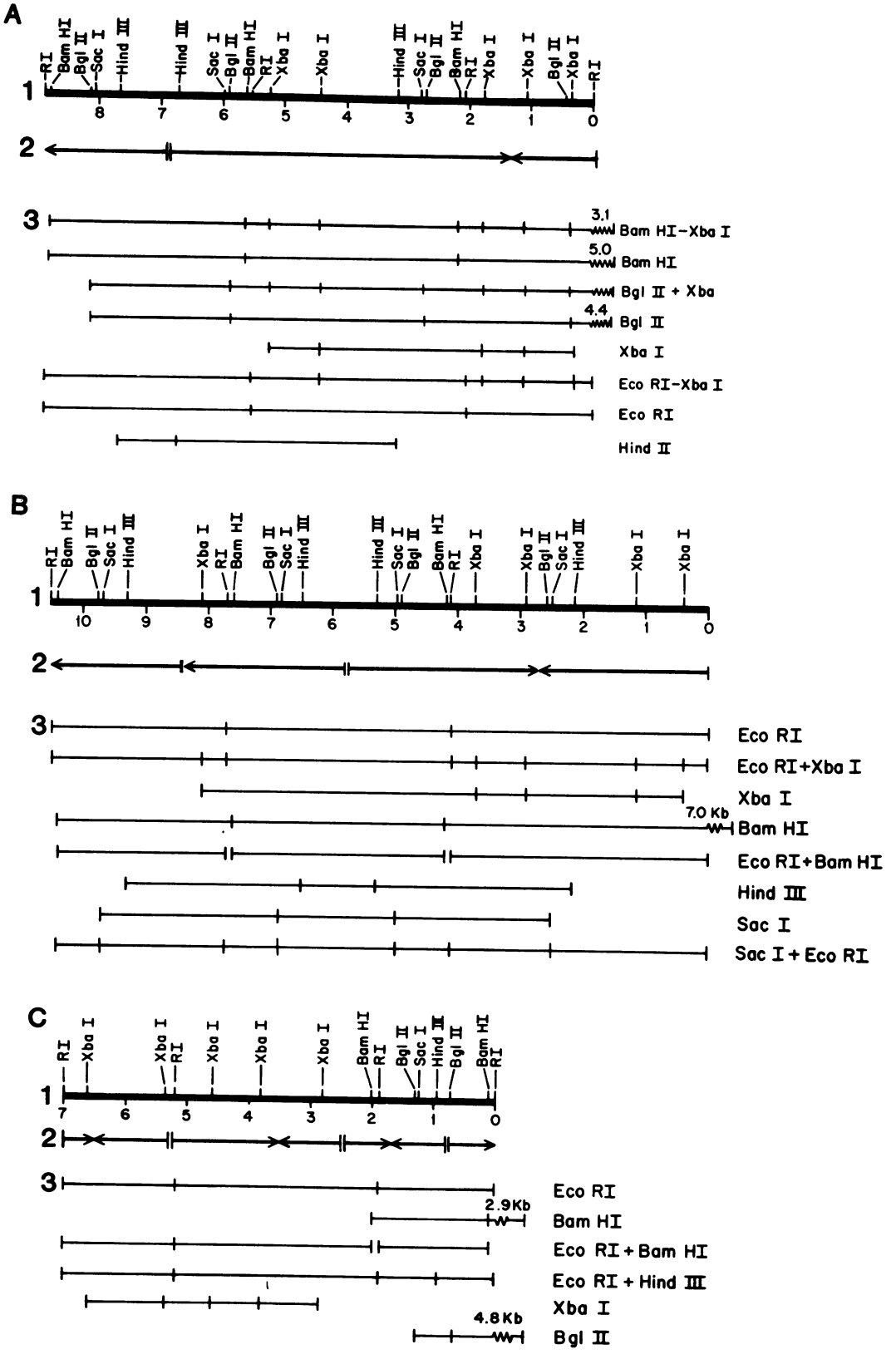


FIG. 6. Heteroduplex analysis of clones HW197-1 (A), HW197-3 (B), and HW197-4 (C). Charon 30 DNA was hybridized to each of the HW197 clones, and the hybrids were observed under the electron microscope. The interpretation of each figure is presented below the electron micrograph: ▭, double-stranded DNA; ———, single-stranded DNA.

segment of the viral genome which are present in the cloned novel form.

An analysis similar to the one carried out for HW197-1 was also carried out for clones HW197-3 and HW197-4. The partial digest of HW197-3 shown in Fig. 5 again revealed a set of directly repeated viral sequences (bands between 13 and 8 kb; Fig. 5) followed by an inverted repeat of the same sequences (bands between 6.6 and 4.4 kb; Fig. 5). These sequences were followed by a second inversion which was accompanied by the deletion of the *Eco*RI and *Bam*HI sites because the *Eco*RI and *Bam*HI bands, which would be expected if it was an inversion without a deletion, are missing from the last set of viral bands (bands between 4.4 and 2.2 kb; Fig. 5). The labeled 0.5-kb *Sal*I fragment from the Charon 30 right arm is also visible on this gel.

Partial digestion analysis of end-labeled HW197-4 DNA is also shown in Fig. 5. The interpretation of this gel was complicated by the fact that the *Sal*I digestion, carried out before labeling, was incomplete. Therefore, each partial digestion fragment of the cloned DNA is represented by two bands which are 0.5 kb apart (0.5 kb is the distance between the *Sal*I sites in the right arm of Charon 30). Each set of fragments is represented by a heavy band and a light band, which is the *Sal*I partial digestion product. Taking this into account, it is evident that this clone does not contain the long stretches of colinear viral DNA between the *Hin*dIII and *Eco*RI digestion sites. Several of the sites present on the restriction map which was constructed from these data did not lie at the position expected when the viral DNA was present as an uninterrupted viral DNA sequence. It was necessary to postulate that this clone contained six sites at which inversions in the viral sequences were accompanied by deletions of adjacent viral sequences to explain the position of the various restriction endonuclease sites. Electron microscopic analysis of heteroduplex structures was necessary to formulate this postulation, and these data will be presented in the next section.


A
RI
Bam HI
Bgl II
Sac I
Hind III
Hind III
Sac I
Bgl II
Bam HI
RI
Xba I
Xba I
Hind III
Sac I
Bgl II
Bam HI
RI
Xba I
Xba I
Bgl II
Xba I
RI
1
8
7
6
5
4
3
2
1
0
2
3
3.1
Bam HI-Xba I
5.0
Bam HI
Bgl II + Xba
4.4
Bgl II
Xba I
Eco RI-Xba I
Eco RI
Hind II
B
RI
Bam HI
Bgl II
Sac I
Hind III
Xba I
RI
Bam HI
Bgl II
Sac I
Hind III
Hind III
Sac I
Bgl II
Bam HI
RI
Xba I
Xba I
Bgl II
Sac I
Hind III
Xba I
Xba I
1
10
9
8
7
6
5
4
3
2
1
0
2
3
Eco RI
Eco RI+Xba I
Xba I
7.0 Kb
Bam HI
Eco RI+Bam HI
Hind III
Sac I
Sac I + Eco RI
C
RI
Xba I
Xba I
RI
Xba I
Xba I
Xba I
Bam HI
RI
Bgl II
Sac I
Hind III
Bgl II
Bam HI
RI
1
7
6
5
4
3
2
1
0
2
3
Eco RI
2.9Kb
Bam HI
Eco RI + Bam HI
Eco RI + Hind III
Xba I
4.8Kb
Bgl II

It should be noted that the left and right ends of each of the novel form clones contained *Eco*RI sites which were accompanied by either a *Bam*HI site approximately 80 bp away or an *Xba*I site approximately 400 bp away. This is the result that would be expected if the left or right *Eco*RI sites were in viral sequences, and the presence of viral sequences at the ends of the clones is taken as additional evidence supporting the conclusion that these clones contain exclusively viral sequences.

Heteroduplex analysis of clones HW197-1, -3 and -4. The restriction maps obtained from the partial digest data led us to conclude that there were inverted repeat viral sequences in each of the clones. To test our conclusions by an independent method, we examined the secondary structures of DNA heteroduplexes of each clone with Charon 30. Under the conditions which we chose for the experiments, inverted repeat sequences would be expected to reanneal with themselves and appear as hairpin-like structures and loops within the clone.

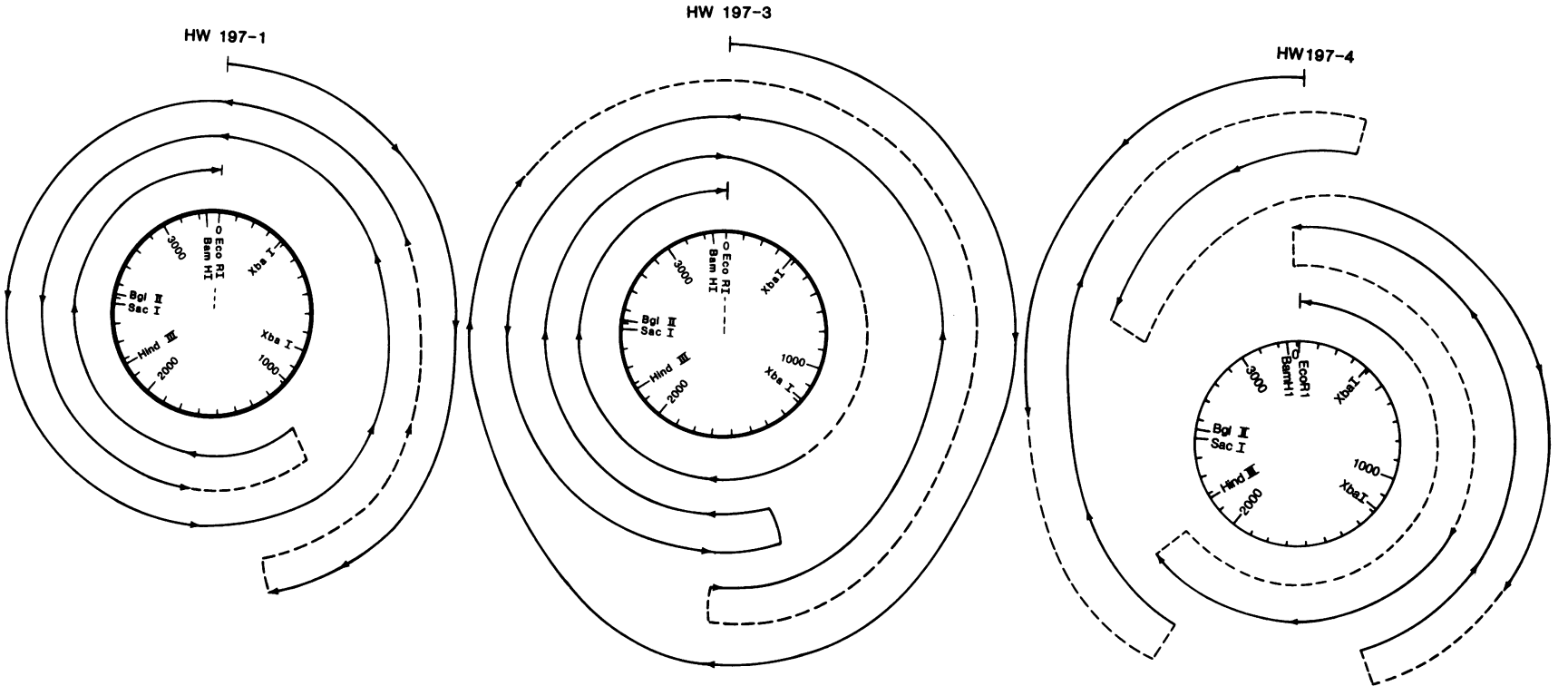
The characteristic heteroduplex structures observed for each clone are presented in Fig. 6A through C along with an interpretation of the structures below it. HW197-1 contained "lollipop" structures located at the extreme right and left ends of the clone. These structures resulted from inverted duplications in which some of the viral sequences were deleted to one side of the inversion site. Upon reannealing, the deleted sequences present on only one side of the inversion did not hybridize and appear as a single-stranded loop (Fig. 6A). The length of the inverted sequences was determined, using the internal Charon 30 stuffer fragment as a molecular weight standard. When the length of the inverted repeats was compared with the restriction map, the inversion sites were located between the *Hin*dIII sites on the left side of the clone and the *Xba*I sites on the right side of the clone, within the region in which they were expected from the restriction map. The actual sequences involved in the inverted repeat structures have been denoted by arrows below the final map of clone HW197-1 shown in Fig. 7A, line 2.

The same process of analysis, comparing observed secondary structures with the restriction map from partial digests, was done for clones HW197-3 and -4. Two inverted duplications were also observed in clone HW197-3. In this case, however, one of the resulting hairpin structures did not have a single-stranded loop at the end, indicating a nearly symmetrical arrangement of the inverted sequences. The portions of HW197-3 participating in the hairpins are shown in Fig. 7B, line 2, below the restriction map.

The secondary structure of HW197-4 was very complex, as we had predicted. Two different figures were observed under the electron microscope. The most prevalent form observed is the one illustrated in Fig. 6C. An important aspect of this figure is the fact that the single-stranded loop of the large hairpin hybridizes specifically to a section of the clone to its left. The restriction map predicts a possible inversion around the *Xba*I sites located between 2.8 and 3.8 kb which could form a large hairpin. The map also predicts that a homologous section of DNA in the opposite orientation is located at the *Xba*I site at 5.35 kb. We take the fact that the loop of the large hairpin often hybridized to the segment of DNA at 5.3 kb from the right end as confirmation of the restriction map in which the *Xba*I site at 5.35 kb is homologous to the *Xba*I site at 3.8 kb, but that the two sequences are in opposite orientation, thus allowing the hybridization of the hairpin loop. A similar result was also observed for clone HW197-1. The single-stranded loops on each hairpin showed sequence homology with segments in the center of the clone which are in opposite orientation and were often observed to be hybridized to the center of the clone. This is evident for one of the hairpin loops in Fig. 6A. The other secondary structure of HW197-4 has not been included for the sake of brevity but supports the proposed restriction map (see the diagram under Fig. 7C, line 2, for a summary of the orientation of the viral sequences in HW197-4).

Restriction endonuclease mapping using complete digests. As a final step in the characterization of the novel forms, single and double digests of the cloned DNA were analyzed by agarose gel electrophoresis. These experiments enabled us to confirm the restriction map by observing the fragments predicted by the partial digest data (Fig. 7A through C). It was important to confirm the order of some sites on the restriction map to establish the orientation of the viral sequences, particularly for clone HW197-4. One example in which this was done was to test the orientation

FIG. 7. Restriction endonuclease maps of the HW197 clones. (A) HW197-1; (B) HW197-3; (C) HW197-4. Line 1 in each section, Final restriction endonuclease map. Line 2 in each section, Arrows defining the orientation of segments of the clone with regard to the 5′→ 3′ direction of the minus strands of viral DNA. Line 3 in each section, Diagram of the restriction endonuclease fragments observed upon single and double digestion of the cloned DNA in Charon 30. These data were used for final confirmation of the above maps obtained from partial digestions.


HW 197-1
Eco RI
Bam HI
Xba I
Bgl II
Sac I
Hind III
3000
2000
1000
HW 197-3
Eco RI
Bam HI
XbaI
Bgl II
Sac I
Hind III
3000
2000
1000
Xba I
HW197-4
EcoR1
BamH1
XbaI
Bgl II
Sac I
Hind III
3000
2000
1000
XbaI

of the *Bam*HI sites on the right-hand side of clone HW197-4. If the *Bam*HI site at 2 kb was to the right of the *Eco*RI site, it would have been a simple model to propose a single inversion around the *Hin*dIII site at 1.0 kb, and the rest of the viral sequences through 4.0 kb would be colinear with the viral genome. However, the opposite orientation of the *Bam*HI site required at least two more inversions to explain the occurrence of the *Bgl*II and *Sac*I sites at 1.3 kb and the *Xba*I site at 2.8 kb. The size of the restriction fragments observed after the digestion of HW197-4 with *Bam*HI only or *Bam*HI plus *Eco*RI unequivocally established the position of the *Bam*HI site to the left of the *Eco*RI site at 1.9 kb because *Bam*HI alone cuts out a piece of a size identical to that of the 1.9-kb *Eco*RI fragment, which could not be the case if the *Bam*HI site were to the right of the *Eco*RI site. Therefore, we proposed inversions at 0.7, 1.6, and 2.5 kb. The secondary structures observed for HW197-4 confirmed that these inversions occur at the expected positions.

Another case in which restriction digestion confirmed an unexpected restriction site was in clone HW197-1. A *Bgl*II site was observed in the partial *Bgl*II digest which did not coincide to an expected *Bgl*II site in the viral sequences (Fig. 5). It was necessary to confirm the presence of this *Bgl*II site and map it with respect to the viral restriction sites present in its vicinity. *Bgl*II digestion confirmed the position of the *Bgl*II site at 0.45 kb from the right side (Fig. 7A, line 3), and *Bgl*II-*Xba*I double digestion showed that approximately 50 bp were cut off the *Bgl*II fragments by *Xba*I, thus locating a *Xba*I site 0.4 kb from the *Eco*RI site. This is the position predicted for the *Xba*I site if the sequences are viral. The *Bam*HI-*Xba*I double digest provided a second confirmation of the position of the *Xba*I site at 0.4 kb from the right arm of Charon 30.

Various single and double digests of each of the clones were analyzed by agarose gel electrophoresis. The fragments observed for each digest were reconciled exactly to the maps which had already been constructed. These results are summarized as line drawings of the restriction maps, showing the fragments observed for digests of each clone with different enzymes (Fig. 7A through C).

DISCUSSION

The original cloning experiment that we conducted was designed to isolate cellular DNA fragments containing integrated viral sequences from a chronically infected liver. Three of the clones that we isolated after screening 2×10^6 recombinant phage contained the novel forms of WHV which we have described in this paper. In four separate experiments, we looked for the presence of cellular DNA sequences in these clones by hybridizing them to Southern blots of *Eco*RI-digested normal uninfected liver DNA. We did not observe hybridization to any cellular sequence in these experiments. We also established the fact that the extreme right and left *Eco*RI sites of each clone are viral *Eco*RI sites because they are located next to either a *Bam*HI site, which is approximately 80 bp away, or next to an *Xba*I site, which is approximately 400 bp away, as is expected for viral *Eco*RI sites. Therefore, none of the clones appears to consist of anything other than viral sequences.

We established the structure of the DNA sequences of these novel forms by three independent lines of evidence. The initial restriction maps were obtained by partial digestion of end-labeled cloned DNA with enzymes that cut the virus. Three enzymes that do not cut the virus, *Xho*I, *Sal*I, and *Kpn*I, also did not cut the clones, providing further circumstantial evidence for the lack of nonviral sequences. Certain secondary structures were predicted from the restriction maps obtained from partial digestions, and these structures were confirmed to be present with heteroduplex analysis. The presence of secondary structures in the denatured DNA also explains the faint hybridization of some of the inserts seen in the Southern blot in Fig. 3. Finally, we used restriction endonuclease single and double digests to confirm the presence of the restriction fragments predicted by the map obtained with partial digests and electron microscopic analysis.

At what stage did the rearrangements in the viral sequences occur? It is unlikely that these alterations occurred during the propagation of the recombinant phage since they are strictly confined to the region of the viral insert, whereas no rearrangements were observed in the lambda cloning arms. Moreover, similar rearrangements were observed in recombinant clones of WHV DNA derived from viral genomes integrated into the cellular DNA of two hepatomas (8). The latter rearrangements could be confirmed to be present in the tumor DNAs by Southern blot analysis. Thus, it is likely that the rearrangements observed were already present in the DNA in the infected tissue.

The origin of these novel forms is not clear.

FIG. 8. Structure of the novel forms in relationship to the intact circular viral map. The solid lines represent sections of the clone which are colinear with the viral map, and the arrows indicate progression from the left- to right-hand ends of the clone. Sections of the maps which appear as dashed lines connect sections of the genome which are adjacent in the clones. Symbols: ↦, Left end of clone; ↤, right end of clone; ———, viral DNA segments in the clone; ----, connections between segments of viral DNA present in the clone.

The rearranged cloned DNA segments were generated by a partial digestion with *Eco*RI and may have been derived by excision from still larger segments of DNA. Such DNA molecules may have been of chromosomal origin, i.e., viral segments integrated into the host cell genome, or they may have been extrachromosomal, consisting of only viral sequences. Alternatively, each of the cloned segments may have been generated by a single cleavage in a circular viral form containing multiple rearrangements. In the latter case, viral sequences at the two RI ends of the cloned segment would have been adjacent in the uncleaved circular molecule and, therefore, in the same orientation. Since this is the case in the three clones isolated, the latter possibility cannot be excluded.

Figure 8 is presented for the purpose of illustrating the complex structure of the novel forms in relationship to the viral genome. It can be seen that two major features of the cloned novel forms are the presence of multiple copies of viral sequences within the same molecule and the preponderance of reversals of polarity of the viral sequences at the sites at which rearrangements are found. Of 11 such sites, 10 occurred between segments of opposite polarity. Moreover, the same viral segment is present as many as four times in the same molecule (e.g., the *Hin*dIII site in clone HW197-3). These features suggest a process involving either rearrangements and duplications occurring in the same molecule, for example, during DNA replication, or multiple recombinational events among different molecules of viral DNA.

Not all classes of virus-specific DNA in infected cells are likely to participate in such a process of rearrangement. The major pathway for DNA synthesis by hepatitis B virus-like viruses appears to proceed through the production of multiple RNA copies of the viral genome, presumably by the transcription of duplex DNA. These genomic RNAs are assembled into viral cores and subsequently reverse-transcribed into viral DNA, as a step in virus maturation (12). Thus, at least two classes of viral DNA must exist in infected cells. One class consists of progeny virion DNA in the process of being synthesized de novo inside assembled viral cores. These molecules are not likely to be transcriptionally active or necessary for the persistence of the infection at the cellular level, nor are these packaged DNAs likely to be the precursors of the novel forms since they are presumably unable to undergo DNA replication or to recombine with other DNA molecules. The majority of virus-specific DNA in infected cells consists of these packaged forms. A second class consisting of duplex viral DNA must serve as the template for the transcription of genomic and mRNAs. This class of molecules must be continually present for the production of viral particles by the infected cell. The structure of the transcriptionally active viral DNA is not known; however, closed circular viral DNA has been detected as the major species of duplex DNA in the nucleus of infected cells, at a frequency of about 50 copies per cell or about 2% of the total viral DNA in the infected cell (unpublished data). Transcriptionally active viral DNAs in the nucleus would seem to be likely targets for the process that generates the novel forms since they would be accessible to the replicative and recombinational machinery of the cell. Rearrangements introduced into these molecules would be expected to destroy their ability to participate in replicative functions, and this process may account, in part, for the gradual decline in virus production that occurs in chronic infections. Although novel forms were estimated to be rare (average of about one copy per cell), their frequency may become significant in any one cell relative to the number of transcriptionally active forms. Thus, novel forms may be the end products of a process that gradually attenuates viral replication by destroying forms of viral DNA that are essential for continued virus production. Such a process, in the extreme, could lead to the resolution of the infection at the cellular level.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants AI-15166, RR-05539, and CA-06927 from the National Cancer Institute and an appropriation from the Commonwealth of Pennsylvania.

LITERATURE CITED

1. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
2. **Blattner, F. R., A. E. Blechl, K. Denniston-Thompson, H. E. Faber, J. E. Richards, J. L. Slightom, P. W. Tucker, and O. Smithies.** 1978. Cloning human fetal γ globin and mouse α-type globin DNA: preparation and screening of shotgun collections. Science **202**:1279–1284.
3. **Brechot, C., M. Hadchouel, J. Scotto, M. Fonck, F. Potet, G. Vyas, and P. Tiollais.** 1981. State of hepatitis B virus DNA in hepatocytes of patients with hepatitis B surface antigen-positive and -negative liver diseases. Proc. Natl. Acad. Sci. U.S.A. **78**:3906–3910.
4. **Cummings, I. W., J. K. Browne, W. A. Salser, G. V. Tyler, R. L. Snyder, J. M. Smolec, and J. Summers.** 1980. Isolation, characterization and comparison of recombinant DNAs derived from genomes of human hepatitis B virus and woodchuck hepatitis virus. Proc. Natl. Acad. Sci. U.S.A. **77**:1842–1846.
5. **Enquist, L., and N. Sternberg.** 1979. In vitro packaging of λ Dam vectors and their use in cloning DNA fragments. Methods Enzymol. **68**:281–298.
6. **Enquist, L. W., M. J. Madden, P. Schiop-Stansly, and G. F. VandeWoude.** 1979. Cloning of herpes simplex type I DNA fragments in bacteriophage lambda vector. Science **203**:541–544.
7. **Galibert, F., T. N. Chen, and E. Mandart.** 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence.

J. Virol. 41:51–65.
8. **Ogston, C. W., G. J. Jonak, C. E. Rogler, S. A. Astrin, and J. Summers.** 1982. Cloning and structural analysis of integrated woodchuck hepatitis virus sequences from hepatocellular carcinomas of woodchucks. Cell 29:385–394.
9. **Rimm, D. L., J. K. Horness, and F. R. Blattner.** 1980. Construction of coliphage lambda Charon vectors with Bam H1 cloning sites. Gene 12:301–309.
10. **Smith, H. O, and M. L. Birnstiel.** 1976. A simple method for DNA restriction site mapping. Nucleic Acids Res. 3:2387–2398.
11. **Snyder, R. L., and J. Summers.** 1980. Woodchuck hepatitis virus and hepatocellular carcinoma, p. 447–457. *In* Cold Spring Harbor Conference on Cell Proliferation VII, Viruses in naturally occurring tumors. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
12. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell 29:403–415.
13. **Summers, J., J. M. Smolec, and R. Snyder.** 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc. Natl. Acad. Sci. U.S.A. 75:4533–4537.
14. **VandeWoude, G. F., M. Oskarsson, L. W. Enquist, S. Nomura, M. Sullivan, and P. J. Fischinger.** 1979. Cloning of integrated Moloney sarcoma proviral DNA sequences in bacteriophage λ. Proc. Natl. Acad. Sci. U.S.A. 76:4464–4468.