JOURNAL OF VIROLOGY, Dec. 1995, p. 8102–8108 Vol. 69, No. 12
0022-538X/95/$04.00+0

Duck Hepatitis B Virus Integrations in LMH Chicken Hepatoma Cells: Identification and Characterization of New Episomally Derived Integrations

SHIH S. GONG,* ANNE D. JENSEN, HAIPING WANG, AND CHARLES E. ROGLER

Marion Bessin Liver Research Center, Department of Medicine, Albert Einstein College of Medicine, Bronx, New York 10461

Received 25 July 1995/Accepted 12 September 1995

While the cytoplasmic phase of the hepadnavirus replication cycle is well understood, very little is known about the nuclear phase. In contrast to retroviruses, proviral integration is not required for hepadnavirus replication; however, some of the viral DNAs in the nucleus are diverted into an integration pathway. Under certain conditions these integrations function as carcinogenic agents. In order to study the integration process, we have utilized LMH-D2 cells, which replicate wild-type duck hepatitis B virus (DHBV), to develop the first protocol to detect and characterize integrations of DHBV originating from episomal viral DNAs. Contrary to expectations, our results showed that stable new integrations are readily detectable in subclones of LMH-D2 cells. Complete characterization of one integration revealed a single-genome-length integrant with the structure of double-stranded linear (DSL) DHBV DNAs which are produced by in situ priming during viral replication. The integration contained a terminal redundancy of 6 bp from the r region of the virus DNA minus strand as well as a direct repeat of 70 bp of cellular DNA. On the basis of the structure of the integrant and the cellular DNA target site, we propose a molecular model for the integration mechanism that has some similarities to that of retroviruses. Identification of DSL hepadnavirus DNA integration suggests the possibility that modified DSL viral DNAs may be the precursors to a class of simple, unrearranged hepadnavirus integrations.

The cytoplasmic phase of hepadnavirus replication occurs in nucleocapsids and has been characterized in great detail (16, 19, 26). The nuclear phase of hepadnavirus replication, on the other hand, is very poorly understood. Results of in vitro infection studies suggest that, after duck hepatitis B virus (DHBV) infects hepatocytes, the open circular (OC) virion DNA is released in the nucleus and is converted to covalently closed circular (CCC) DNA. Subsequently, the CCC DNA molecules serve as templates for production of viral mRNAs and pregenomic RNA (8, 12, 24). The levels of CCC DNA in the nucleus have been shown to correlate with the levels of viral gene expression and replication (9, 20).

Although OC DHBV DNA is the predominant molecule in DHBV virions, a minor fraction of virions contain double-stranded linear (DSL) molecules which are produced when viral plus-strand replication proceeds without translocation of the RNA primer (Fig. 1A and B). This process has been termed "in situ" priming (18). The proportion of virions containing DSL molecules can be varied by introducing mutations which interfere with primer translocation. Thus, when translocation of the plus-strand primer is completely blocked, only virions containing DSL DNAs are produced (13). Until recently, the fate of DSL DNAs in the nucleus was unknown. However, Yang and Summers have now shown that DSL molecules can be converted to CCC DNA molecules in the nucleus (29).

The conversion of OC or DSL virion DNAs to CCC DNA molecules requires several enzymatically catalyzed reactions including removal of the terminal protein and the plus-strand RNA primer, resolution of the terminal redundancy in the minus strand, and filling in and ligation of the gaps and nicks in the plus and minus strands (Fig. 1A to C). These alterations are believed to require enzymatic reactions in the nucleus of the host cell. The balance between reactions which promote completion of CCC DNA molecules versus linearization of viral DNAs ultimately controls the level of productive replication, since molecules which remain in the DSL configuration cannot support viral replication (13, 29).

In contrast to retroviruses, there is no requirement for hepadnaviruses to form proviruses in order to complete the replication cycle (16). However, some proportion of viral DNA molecules which enter the nucleus are diverted into an integration pathway, since integrated viral DNAs have been cloned from chronically infected liver and hepatocellular carcinomas from chronic carriers (4, 5, 11, 14, 15, 17, 21, 22, 27, 28). One of our goals has been to determine the mechanism of viral DNA integration in the nucleus in cells replicating wild-type hepadnaviruses. For these studies we have chosen an avian hepatoma cell line, LMH-D2, which efficiently replicates wild-type DHBV (3, 7). In this report we characterize several DHBV integrations that originated from episomal DHBV DNAs.

Identification of DHBV integrations in LMH-D2 cells. Our initial approach was to use Southern blotting to characterize the major loci of integrated DHBV DNAs in LMH-D2 cells. The LMH-D2 cell line was derived from LMH chicken hepatoma cells that were stably transfected with a DHBV expression plasmid. The DHBV expression vector contains a greater-than-unit-length DHBV DNA placed under the control of the cytomegalovirus (CMV) immediate-early promoter and the selectable marker neomycin resistance gene (*neo*) (Fig. 2A) (3). These cells constitutively express DHBV RNAs, replicate

* Corresponding author. Mailing address: Marion Bessin Liver Research Center, Department of Medicine, Jack and Pearl Resnick Campus of the Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461. Phone: (718) 430-3651. Fax: (718) 430-8975.

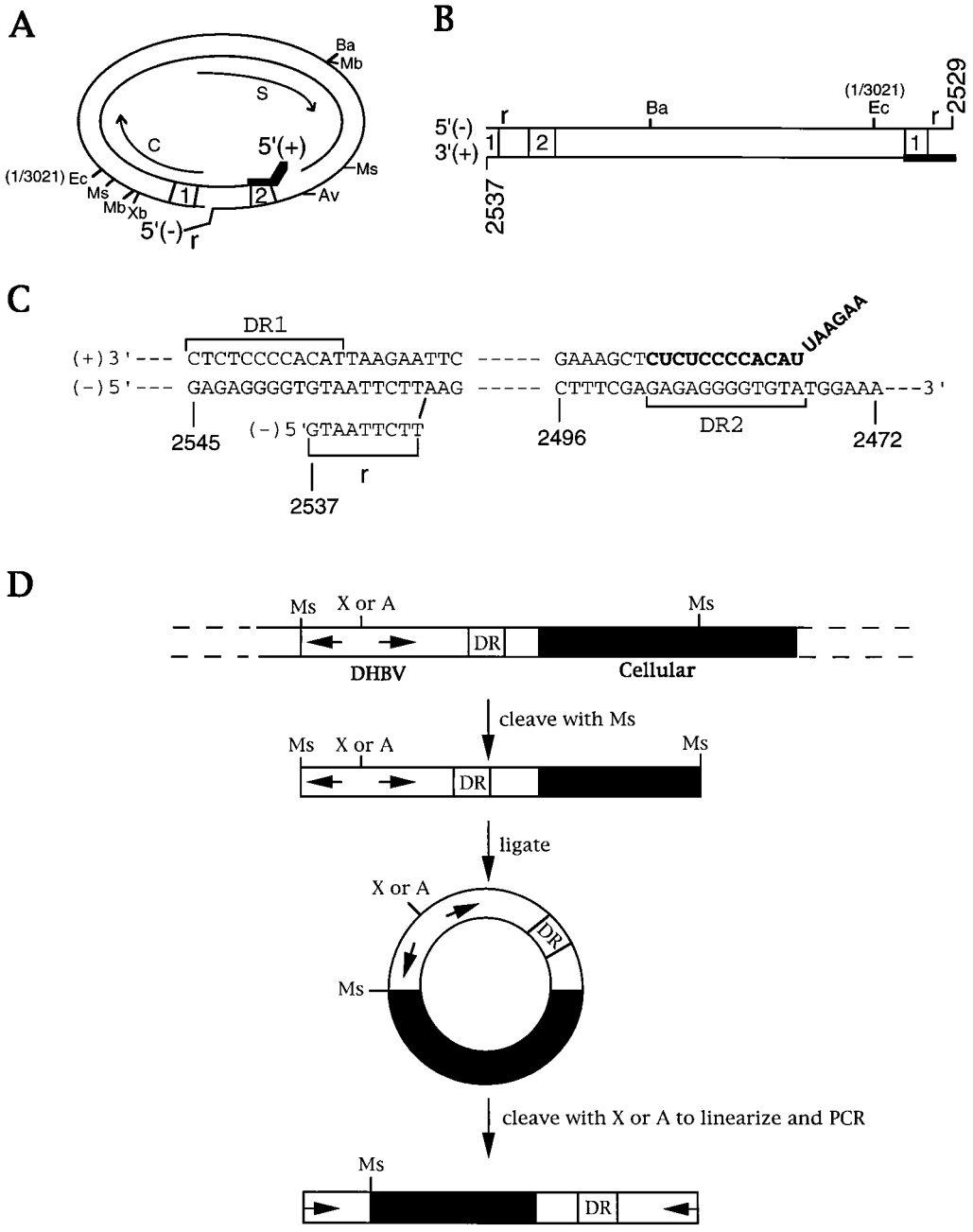

FIG. 1. Schematic representation of the DHBV genomes and the inverse PCR strategy to clone DHBV-cell junctions. (A) OC DHBV virion DNA. The triple-strand region contains the terminal redundancy (r) of the minus strand. Regions of the viral genome containing DR1 and DR2 correspond to the sites of initiation of synthesis of the viral minus and plus strands. The thick line at the 5′ end of the plus strand represents the RNA primer which is partially duplexed. The nucleotide numbers of the DHBV genome are those used by Mandart et al. (10) and start from the unique *Eco*RI (Ec) site. The positions of the unique *Bam*HI (Ba) site and the *Mbo*I (Mb), *Msp*I (Ms), *Xba*I (Xb), and *Ava*I (Av) sites flanking the cohesive overlap region between DR1 and DR2 are indicated. The coding regions for the core (c) and surface (s) genes are also indicated. (B) DSL DNA. The RNA primer (thick line) at the 5′ end of the plus strand is completely duplexed. (C) Sequences of the DRs and the r region of the OC DHBV DNA. The region corresponding to the RNA primer for plus-strand synthesis is boldfaced. (D) Inverse PCR strategy. Open box, DHBV DNA; closed box, cellular DNA. A virus cell DNA junction in the vicinity of the viral DR is diagrammed. The nuclear DNA containing the junction is digested with either (conveniently located) *Mbo*I or *Msp*I (Ms) (see panel A), and the digested DNA is then ligated at 1 ng/μg to circularize individual restriction fragments. The ligated DNA is then digested with either *Ava*I (A) (DHBV nt 2501) or *Xba*I (X) (DHBV nt 2663) to linearize the desired junction DNA before being subjected to two rounds of PCR with nested primer pairs (arrows).

DHBV DNA, and secrete mature infectious DHBV virions. Since LMH-D2 cells contain stable integrations of the transfected DHBV expression plasmid, we wanted to know whether integrated DHBV DNAs not linked to any plasmid vector sequences were present in the LMH-D2 nuclear DNA. Presumably, these integrations would be derived from episomal forms of DHBV in the LMH cells. We hybridized Southern blots of *Pvu*II-digested nuclear DNA of LMH-D2 cells sequentially with probes homologous to DHBV, the Neo gene, or the CMV immediate-early promoter (Fig. 2). *Pvu*II cleaves the

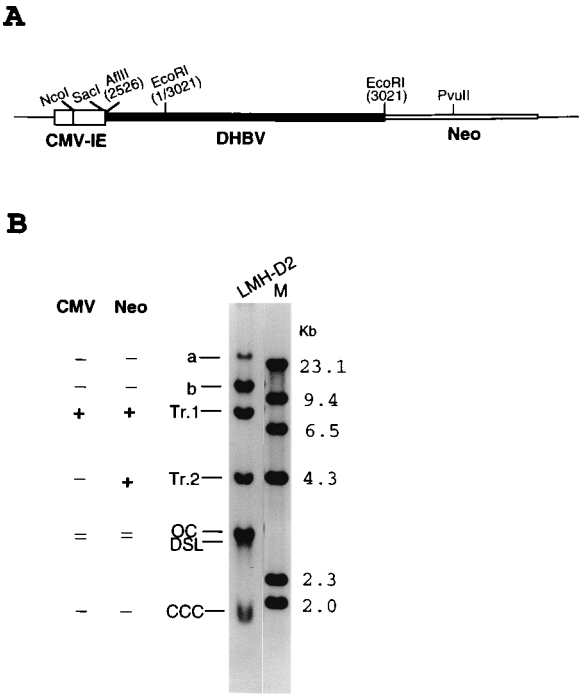

FIG. 2. Analysis of DHBV integrations in the LMH-D2 cell line. (A) Schematic representation of the plasmid used for stable transfection of LMH chicken hepatoma cells to generate the LMH-D2 cell line. A DHBV fragment of greater than unit length was placed under the control of the CMV immediate-early (IE) promoter (3). A 1.8-kb neomycin resistance cassette containing the murine beta-globin promoter to drive the transcription of the neomycin resistance gene (*neo*) was further inserted into the expression plasmid. The unique *Pvu*II site in the plasmid is located within the Neo gene. (B) Total nuclear DNA isolated from the LMH-D2 cells was digested with *Pvu*II, electrophoresed on 1% agarose gel along with a radiolabelled lambda DNA *Hin*dIII digest as size marker (M), transferred to a Zetabind membrane, and probed with entire DHBV DNA. The same filter was also probed sequentially with the 200-bp *Nco*I-*Sac*I fragment of the CMV immediate-early promoter region and the 1.2-kb *Ava*I fragment of the Neo gene. OC, DSL, and CCC forms of the DHBV genomic DNA are indicated. A band of the same migration as band Tr.1 hybridized (+) to both the CMV probe and the Neo probe, and band Tr.2 hybridized to the Neo probe but not to the CMV probe. Bands a and b hybridized to neither the CMV probe nor the Neo probe.

DHBV expression plasmid once in the Neo gene and does not cleave the DHBV DNA (10). As shown in Fig. 2B, nuclear DNA from LMH-D2 cells contained episomal forms of DHBV DNA designated OC, DSL, or CCC plus four high-molecular-weight fragments containing integrated DHBV DNAs. The high-molecular-weight fragments Tr.1 and Tr.2 also hybridized to vector sequences (CMV and/or Neo), and we concluded that these bands were acquired during transfection. Fragments a and b were not linked to either the CMV sequence, the Neo sequence, or other plasmid sequences and were candidates for viral DNA integrations from episomal forms of DHBV.

Cloning and sequence analysis of the b integration in LMH-D2 cells. We designed an inverse PCR (23) strategy, utilizing oligonucleotide primers derived from the region of the DHBV genome flanking the cohesive overlap region to clone virus-cell junction fragments (Fig. 1D). The strategy was designed to preferentially clone virus-cell DNA junctions in the vicinity of the direct repeat (DR) sequences, since they are known to be preferred integration sites for hepadnavirus DNA. In this protocol, we first digested the LMH-D2 nuclear DNA with either *Mbo*I or *Msp*I, both of which cleave the DHBV genome several times but does not cleave the cohesive overlap region. The digested DNAs were then ligated in low concentrations to circularize individual fragments, and the desired junction fragments were linearized by *Xba*I digestion, whose site is located near the DR sequence, for viral integrations proceeding toward the core gene coding region or *Ava*I digestion for viral sequences proceeding toward the surface gene region (Fig. 1A and D). We then used nested primer pairs derived from regions of DHBV DNA on either side of the *Xba*I or *Ava*I site and inside the viral *Msp*I or *Mbo*I site flanking the cohesive overlap as illustrated in Fig. 1D. With this strategy, we amplified two virus-cell junction DNA fragments from the *Msp*I digest and one fragment from the *Mbo*I digest. These fragments were distinct from DNA fragments expected from episomal DHBV DNA or from the DHBV expression plasmid. After cloning and sequencing the fragments, we determined that one *Msp*I fragment (Fig. 3A, fragment 2) contained a DHBV-cell junction at DHBV nt 2532 in the r region adjacent to DR1 and that the viral sequences proceeded toward the core gene region (Fig. 3B and C, right junction). The other two fragments (Fig. 3A, fragments 1 and 3), one containing an *Msp*I site and the other containing an *Mbo*I site, contained identical viral integration sites at DHBV nt 2537, with viral sequences proceeding toward DR2 (Fig. 3B and C, left junction). The overlapping cellular sequences of the two fragments were also identical, suggesting that we had independently cloned the same junction fragment. DHBV nt 2537 is precisely at the 5′ end of the DHBV minus strand, and the DHBV sequences in the junction fragment proceeded toward DR2 (Fig. 3B and C, left junction).

Since the viral sequences in the two junction fragments proceeded into opposite regions of the DHBV genome, we predicted that the two junctions represented two ends of a single DHBV integration. To test this hypothesis, we synthesized oligonucleotide primers from the left and right cellular sequences and used them in a PCR with LMH-D2 DNA. We amplified a fragment slightly larger than 3 kb (fragment 8 in Fig. 3A), and sequence analysis of the entire fragment confirmed that our hypothesis was correct (Fig. 3A and B). The integration contained an unrearranged, complete DHBV genome with a 6-bp DR at the ends of the DHBV integrant and with junctional sequences identical to the two junction fragments isolated previously (Fig. 3B). Thus, we have isolated a complete DHBV integrant that contains a DHBV genome with a 6-bp terminal redundancy. The structure of the integrant suggests that it was a primary integration that was produced from either an OC virion DNA which was cleaved at staggered positions in the r region or a DSL virion DNA.

In order to isolate the fragments shown in Fig. 3A, for fragments 1 to 3, 9, and 10, which were isolated by inverse PCR, two noncontiguous segments of DNA (dashes) were coamplified through ligation of the *Msp*I (Ms) ends (fragments 1 and 2), the *Mbo*I (Mb) ends (fragment 3), or *Bam*HI (Ba) ends (fragments 9 and 10). The sequences of the oligonucleotide primers used for isolation of the fragments were as follows: fragment 1, DHBV nucleotides (nt) 1750 to 1731 and 2420 to 2439 followed by nested primers of nt 1710 to 1690 and 2441 to 2460; fragment 2, DHBV nt 2620 to 2600 and 2845 to 2864 followed by nt 2592 to 2573 and 2870 to 2890; fragment 3, DHBV nt 2252 to 2233 and 2420 to 2439 followed by primers of nt 2226 to 2207 and 2441 to 2460. Fragments 4 to 7 were amplified with one primer from the cellular sequences in junction fragment 1 (fragments 4 and 5) or 2 (fragments 6 and 7) and one primer deduced from the circular DHBV genome. The sequences of the primers were as follows: fragment 4, 5′-CGTGTCACCTTGCATCTCAACATC-3′ (located in fragment 1) and DHBV nt 1812 to 1833; fragment 5, the same

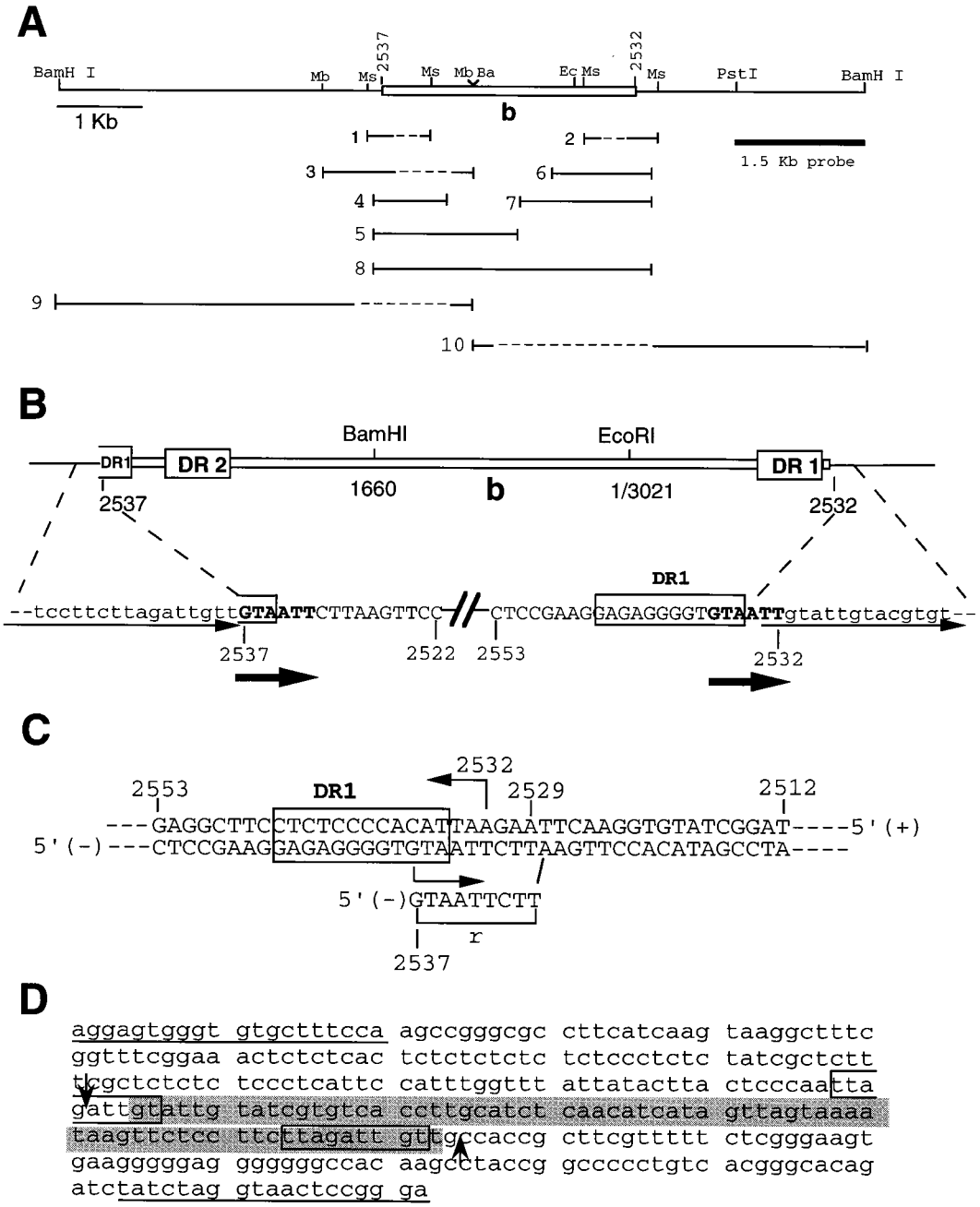

FIG. 3. Features of the chromosomal locus containing the DHBV b integration. (A) Restriction map of the integration b locus. Open bar, DHBV DNA; thin lines, cellular DNA. The DHBV junctional nucleotide numbers are indicated. Fragments (numbered) containing the corresponding regions of DNA were isolated by either inverse PCR (fragments 1 to 3, 9, and 10) (see the legend to Fig. 1) or PCR (fragments 4 to 8). (B) Junctional sequences of the b integration. Integration b containing a complete complement of the DHBV genome is shown. DHBV (uppercase) and cellular (lowercase) sequences at the junctions are shown. The DHBV sequence is that of the minus strand. The terminal redundancy at each end of the b integration is boldfaced. Thin arrows, parts of the 70-bp repeat sequence flanking the integration (see panel D). (C) Locations in the DHBV genome of the left and right viral junctions from the b integration are indicated (arrows). The sequence near the terminal redundancy of the minus strand of the OC DHBV DNA is shown. (D) Sequence of the cellular DNA at the site of integration. The sequence of a fragment amplified by oligonucleotides (underlined) derived from the cellular sequences of the junction fragments 1 and 2 (see panel A) and amplified from the untransfected LMH cells is shown. The 70-bp DR flanking the b integration is shaded. Nine-base-pair DRs located at the ends of the 70-bp sequence are boxed. Arrows, possible sites of two staggered nicks occurring across the 70-bp sequence during integration of the DHBV sequences (see Fig. 5).

cellular primer as for fragment 4 and DHBV nt 1031 to 1051; fragment 6, 5′-TATGATGTTGAGATGCAAGATGAC-3′ (located in fragment 2) and DHBV nt 271 to 251; fragment 7, the same cellular primer as for fragment 6 and DHBV nt 963 to 943; fragment 8, 5′-CGTGTCACCTTGCATCTCAA CATC-3′ and 5′-TATGATGTTGAGATGCAAGGTGAC-3′ (cellular sequences of fragments 1 and 2, respectively); fragment 9, 5′-TACAGTATGTAGGGAGCTGGA-3′ (located in fragment 3) and DHBV nt 1750 to 1731 (located near the *Bam*HI site), followed by the nested primers 5′-GGTAGGAC CATCACGATTGAT-3′ and DHBV nt 1710 to 1690; fragment 10, 5′-GTCATCTTGCATCTC (located in fragment 2)

and DHBV nt 1440 to 1460, followed by 5′-AGATTGTTGC CACCGCTTGGT-3′ and DHBV nt 1633to 1653.

Analysis of the b integration chromosomal insertion site. In order to determine whether we had cloned integration a or b (Fig. 2), we hybridized a Southern blot of LMH-D2 genomic DNA with flanking sequence probes immediately adjacent to the integrated DHBV DNA. Southern blot analysis indicated that both cellular junctional sequences immediately adjacent to the integration contained repetitive DNA elements (data not shown). We isolated a unique sequence probe from the integration site by first "walking" further away from the DHBV integration using a second inverse PCR protocol (see above). This protocol produced two fragments extending to the next *Bam*HI site in the cellular DNA (fragments 9 and 10 in Fig. 3A). The 1.5-kb *Pst*I-*Bam*HI fragment (Fig. 3A) in the right-hand flanking sequence contained unique cellular sequences and was used as a probe to identify the integration in LMH-D2 cells. We found that it hybridized to the *Pvu*II fragment which we previously identified as integration b (data not shown). (Further analysis of the integration site in an LMH-D2 subclone which lost the b integration will be published separately.)

Analysis of the cellular flanking sequences of integration b revealed a 70-bp DR immediately adjacent to the ends of the viral integrant (Fig. 3B). Segments of the repeated sequence are underlined by arrows in Fig. 3B, and the entire sequence is shaded in Fig. 3D. In order to determine whether this repeat occurred as a result of DHBV integration, we cloned the normal cellular sequence spanning the integration site. To do this, we used cellular DNA primers homologous to the left and right cellular flanking sequences and amplified a 320-bp fragment from the parental LMH cell line (see the legend to Fig. 3D). Sequence analysis of several independent clones of the PCR product showed that they contained only a single copy of the 70-bp sequence (Fig. 3D). Thus, we conclude that integration of the viral genome resulted in duplication of a 70-bp cellular sequence at the integration site. Interestingly, a 9-bp DR (Fig. 3D, boxed sequences) was also found at the ends of the 70-bp cellular DNA sequence.

Cloning of additional new viral-cell junction fragments from LMH-D2 cells. In order to determine whether the viral sites in the b integration might represent common integration sites for DHBV DNA in LMH-D2 cells, we searched for additional new integrations in subclones of the LMH-D2 cell line. We hypothesized that stable DHBV integrations that occurred in a subpopulation of LMH-D2 cells could be detected by analyzing the DNA from subclones grown from single cells. Single-cell suspensions of LMH-D2 cells were plated out at low cell density, and single colonies were picked and grown into subclones. Southern blot analysis of three generations of subclones generated in this manner demonstrated that new integrations do arise in LMH-D2 cells, and a lineage analysis of the subclones will be published separately (3a). Southern blot analysis of one second-generation subclone, P1(21)-2, and one third-generation subclone, P1(21)-9-8, which identified three new integrations in LMH-D2 cells is shown in Fig. 4A.

We cloned DHBV-cell DNA junction fragments from these subclones, using the same inverse PCR strategy as earlier, and the sequence analysis of four virus-cell junctions is shown in Fig. 4B. One of the viral junctions was within the terminal redundancy of the minus strand at nt 2533, with the viral sequence proceeding toward the core gene. This viral junction was only 1 bp away from the corresponding junction in the b integration. Two additional viral junctions were located between DR1 and DR2, at nt 2503 and 2507, with DHBV sequences proceeding toward DR2 (Fig. 4B and C). A fourth

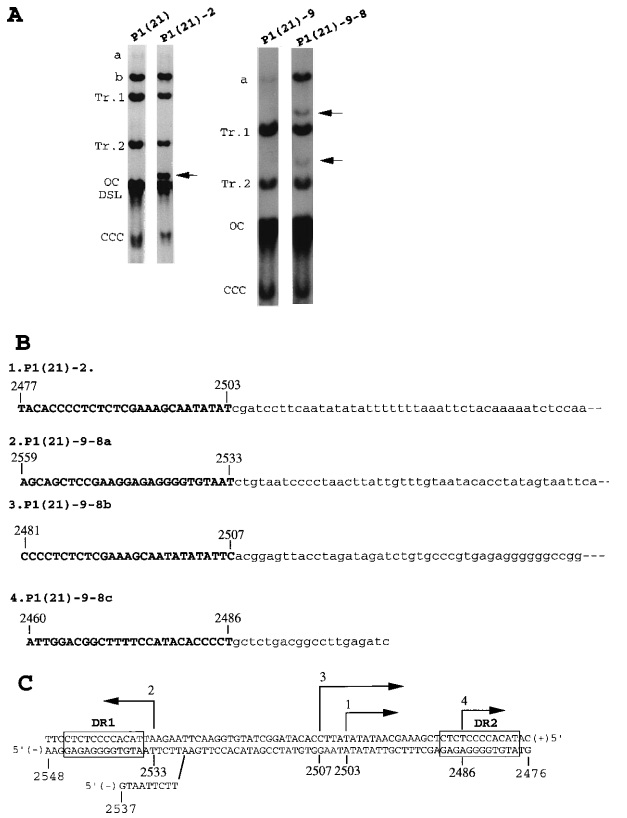

FIG. 4. New DHBV integrations detected and cloned from subclones of LMH-D2 cells. (A) LMH-D2 single-cell subclones P1(21)-2 and P1(21)-9-8 were derived, respectively, from clones P1(21) and P1(21)-9, which were in turn derived from the LMH-D2 parental cell population by single-cell cloning. Nuclear DNA was restricted with *Pvu*II and probed with total DHBV DNA in Southern analysis. The bands are labelled as in Fig. 2. Arrows, new DHBV integrations in the subclones. (B) Four DHBV-cell DNA junction fragments were isolated from subclones P1(21)-2 and P1(21)-9-8 by inverse PCR using DHBV primers as described for Fig. 3. Viral sequences are uppercased; cellular sequences are lowercased. (C) Locations in the DHBV genome and orientation of the viral junctions of the integrations shown in panel B. The sequence of the cohesive overlap region in the OC DHBV DNA is shown. The arrows indicate the positions of the corresponding viral junctions shown in panel B and the directions in which the viral sequences of the integrations proceed in the genome.

viral junction mapped within DR2, with the viral sequence proceeding toward the surface gene coding region (Fig. 4B and C). The DNA sequences of all of the new integrations were confirmed by rederiving them with a unique new cellular primer from each one plus a different DHBV primer and then cloning and sequencing (data not shown).

To unequivocally demonstrate that the virus-cell junctions represented new integrations, the pairs of viral and cellular primers used to amplify the fragments containing the new junctions were used in PCRs with the original LMH-D2 cell DNA and the first-, second-, and third-generation subclones. As shown in Fig. 5, a 930-bp fragment containing the viral junction 1 shown in Fig. 4B was amplified from DNA of the second-generation subclone P1(21)-2 (lane 2) and from DNAs of 12 third-generation single-cell subclones (lanes 3 to 14) derived from P1(21)-2. All 12 third-generation subclones contained the new integration detected in P1(21)-2 (Fig. 4A) by Southern analysis (3a). In contrast, the 930-bp fragment was not detected in the parental first-generation subclone P1(21) (Fig. 5, lane 1) or in another second-generation subclone,

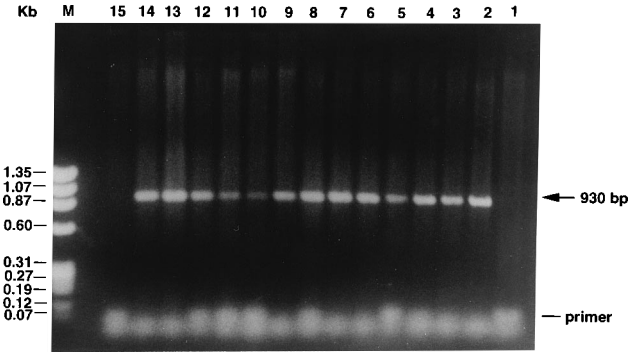

FIG. 5. PCR analysis of LMH-D2 subclone-specific DHBV-cell DNA junctions. Nuclear DNAs of the LMH-D2 second-generation subclone P1(21)-2 (lane 2), its parental population P1(21) (lane 1), 12 single-cell subclones derived from P1(21)-2, (lanes 3 to 14), and another second-generation subclone, P1(21)-7 (lane 15), derived from P1(21), were assayed by PCR for the specific DHBV-cell junction DNA shown in Fig. 4B (junction 1) by using a cellular DNA primer and a DHBV primer. The sequences of the primers used for PCR were 5′-CAGCT GTCTGCTTGGAGATTT-3′ and DHBV nt 1634 to 1653. A 930-bp fragment amplified by the primer pair was electrophoresed on an agarose gel along with a size marker (M) of the *Hae*III digest of ϕX174 DNA, whose sizes are indicated.

P1(21)-7 (Fig. 5, lane 15), derived from P1(21), neither of which contained the new integration in Southern analysis. The predicted DNA fragments containing viral junctions 2, 3, and 4 shown in Fig. 4B were not amplified from DNA of the parental second-generation subclone P1(21)-9 or the first-generation subclone P1(21) (data not shown). Thus, we detected new DHBV integrations arising in LMH-D2 subclones which contained unique virus-cell junctions. Therefore we conclude that the region of DHBV DNA between DR1 and DR2 serves as a substrate for integration and that the terminal redundancy r is a preferred integration sequence within this region.

Proposed mechanism for the b integration. On the basis of our sequence analysis, we propose a three-step process by which the b integration may have occurred (Fig. 6). The basic steps of the integration mechanism include (i) enzymatic modification of the DHBV integration substrate and cleavage of cellular DNA at staggered positions (Fig. 6A and B), (ii) integration of a linearized DHBV molecule (Fig. 6C), and (iii) resolution of the ends of the newly integrated DHBV DNA (Fig. 6D). The enzymatic modifications of DHBV DNA required in step 1 include removal of the terminal protein, cleavage of the minus strand at nt 2532, and removal of the plus-strand RNA primer. This would produce a modified DSL DHBV DNA molecule which would be capable of integrating into cellular DNA (Fig. 6B). Sequence analysis reveals a 3-bp homology with cellular DNA on the right side and 1-bp homology with cellular DNA on the left side of the integration (Fig. 6C, asterisks). The tandem duplication of the 70-bp cellular sequence suggests that the cellular DNA was nicked at staggered positions to provide the integration acceptor (Fig. 6A and B). Interestingly, the potential nick sites in cellular DNA occur within a 9-bp DR which flanks the 70-bp sequence. These sequences may have properties which make them more susceptible to cleavage by cellular enzymes. In our model, the 3′ recessed ends (Fig. 6B) in cellular DNA must be extended by cellular enzymes before the DHBV DNA integrates, because the cellular DNA must be filled in (Fig. 6B and C, dotted arrows) in order to provide the cellular DNA strand with 1- or 3-bp homology to the plus- and minus-strand ends of DHBV DNA, respectively. Finally, resolution of the sites at which viral and cellular DNA ends interact, utilizing nucleic acid enzymes in the nucleus, is required to complete the integration structure.

Additional characteristics of hepadnavirus integrations. Other DHBV DNAs such as OC or CCC molecules could also be modified to produce integration substrates similar to that in the b integration. Thus, a nick adjacent to nt 2532 on the plus strand of an OC DHBV DNA or staggered nicks near nt 2532 on a CCC DNA would likely linearize the DNA molecule, which could also be a potential integration substrate. However, the ends of such linearized OC or CCC molecules would be expected to have structures different from those of a DSL DNA and therefore would presumably require modifications different from those proposed in Fig. 6 in order to produce the substrate for b integration. The additional junction fragments we cloned illustrate that the specific DHBV nucleotides which become linked to cellular DNA vary within the cohesive overlap region and extend at least to DR2. These integrations may result from a scenario that an enzymatically modified DSL DHBV DNA molecule described in Fig. 6 is further subjected to DNase attacks prior to being engaged in integration reactions. It remains to be seen whether the simple, unrearranged hepadnavirus integrations containing less than a full genome originate from DSL viral DNAs.

The inverse PCR strategy used in this study could potentially detect virus-cell junctions located in a 1-kb region of DHBV DNA surrounding the DR1 and DR2 sequences (Fig. 1). However, all the virus-cell junctions we cloned were located within the cohesive overlap between or within DR1 and DR2. A previous report of three DHBV integrations isolated from a duck hepatoma also revealed a strong preference for viral integration sites in the cohesive overlap region (6). Thus, our

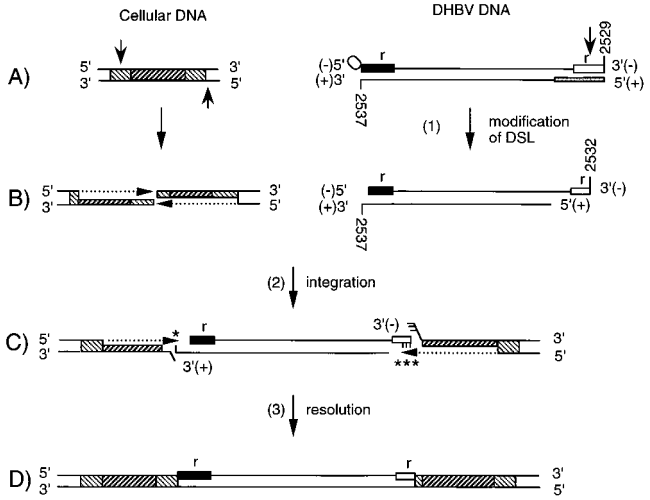

FIG. 6. Proposed pathway for the b integration. (A) DSL DHBV DNA containing the terminal protein (oval) at the 5′ end of the minus strand is shown on the right. Open and closed boxes, 9-bp terminal redundancy (r); thick line, at the 5′ end of the plus strand, RNA primer. At the integration site of the cellular DNA, two staggered nicks (vertical arrows) occurred in the 9-bp DR (thin hatching) located near or at the ends of a 70-bp region (thick hatching). (B) To create the substrate for the b integration, the DSL molecule underwent modifications including removal of the terminal protein, removal of the RNA primer, and cleavage near the 3′ end of the minus strand at nt 2532 (vertical arrow). At the cellular integration site, DNA synthesis occurred to fill in the recessed end (dotted arrows). (C) Initially, during integration of the modified DSL DHBV DNA, 1- (G) or 3-base pairing (ATT) (asterisks) occurred at the left and right junctions, respectively, between the viral and cellular DNA strands. (D) After resolution of the left and right virus cell DNA junctions by cellular nucleases and enzymes, the integrated locus contains a 70-bp cellular DNA DR flanking the DHBV integration.

in vitro system for detecting DHBV DNA integrations correlates well with in vivo data.

Yang and Summers have recently reported that DSL DHBV molecules are efficiently converted into CCC DNA molecules in the nucleus in primary duck hepatocytes (29). Sequence analysis of the CCC DNA molecules across the joint between the 5′ ends of the plus and minus strands revealed a high rate of illegitimate recombination which created mutant CCC DNAs. One of the most highly preferred mutant joints occurred between nt 2537, which is the 5′ end of the minus strand, and the 9-nt r region at the 3′ end of the minus strand. These positions are the same as the junctions of the b integration in this report. The data of Yang and Summers demonstrate that the ends of DSL DHBV molecules participate in illegitimate recombination in both inter- and intramolecular reactions. Our data suggest that appropriately modified DSL molecules also participate in illegitimate recombination with chromosomal DNA.

One remaining question is why the cohesive overlap region is apparently a preferred integration site. Assuming that integration must be mediated by cellular enzymes, structural features of this region of virion DNA which would confer selective sensitivity to a specific nucleic acid enzyme might be expected to enhance integration mediated by that enzyme. The nucleotide sequences across the r region of DHBV DNA have a strong resemblance to a hexadecameric sequence which is a high-efficiency topoisomerase I cleavage sequence (2). In separate work, we have mapped topoisomerase I cleavage sites at either end of the DHBV r sequence (6a). Since the ends of the r region of the DHBV virion DNA have the structure of topoisomerase I suicide substrates, cleavage at these sites might be expected to lead to illegitimate recombination with cellular DNA (1). Previous work demonstrating that topoisomerase I can mediate integration of woodchuck hepatitis virus DNA in its corresponding r region (25) suggests that the mechanism may be common for hepadnaviruses.

We thank T. T. Wu and W. S. Mason for the generous gift of the LMH-D2 cell line; J. Summers for communicating results prior to publication; D. A. Shafritz, V. Prassad, J. Brown, and T. Harris for carefully reading the manuscript; and T. Harris for help with computer work.

This work was supported by Public Health Service grant CA37232, Cancer Center grant P30CA13330, and an LRC grant. S.S.G. was supported by an NIH training grant (CA09060).

REFERENCES

1. **Champoux, J. J., and P. A. Bullock.** 1988. A possible role for the eucaryotic type I topoisomerase in illegitimate recombination, p. 653–666. *In* R. Kucherlapati and G. Smith (ed.), Genetic recombination. American Society for Microbiology, Washington, D.C.
2. **Christiansen, K., B. J. Bonven, and O. Westergaard.** 1987. Mapping of sequence-specific chromatin proteins by a novel method: topoisomerase I on Tetrahymena ribosomal chromatin. J. Mol. Biol. **193:**517–525.
3. **Condreay, L., C. Aldrich, L. Coates, W. Mason, and T.-T. Wu.** 1990. Efficient duck hepatitis B virus production by an avian tumor cell line. J. Virol. **64:**3249–3258.

3a. **Gong, S. S., and C. E. Rogler.** Unpublished data.

4. **Hino, O., K. Ohtake, and C. E. Rogler.** 1989. Features of two hepatitis B virus (HBV) DNA integrations suggest mechanisms of HBV integration. J. Virol. **63:**2638–2643.
5. **Hsu, T., T. Moroy, J. Etiemble, A. Louise, C. Trepo, P. Tiollais, and M.-A. Buendia.** 1988. Activation of c-myc by woodchuck hepatitis B virus insertion in hepatocellular carcinoma. Cell **55:**627–635.
6. **Imazeki, F., K. Yaginuma, M. Omata, K. Okuda, M. Kobayashi, and K. Koike.** 1988. Integrated structures of duck hepatitis B virus DNA in hepatocellular carcinoma. J. Virol. **62:**861–865.

6a. **Jensen, A. D., et al.** Unpublished data.

7. **Kawaguchi, T., K. Nomura, Y. Hirayama, and T. Kitagawa.** 1987. Establishment and characterization of a chicken hepatocellular carcinoma cell line LMH. Cancer Res. **47:**4460–4464.
8. **Kock, J., and H.-J. Schlicht.** 1933. Analysis of the earliest steps of hepadnavirus replication: genome repair after infectious entry into hepatocytes does not depend on viral polymerase activity. J. Virol. **67:**4867–4874.
9. **Lenholf, R., and J. Summers.** 1994. Construction of avian hepadnavirus variants with enhanced replication and cytopathicity in primary hepatocytes. J. Virol. **68:**5706–5713.
10. **Mandart, E., A. Kay, and F. Galibert.** 1984. Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with woodchuck and human hepatitis B virus sequences. J. Virol. **49:**782–792.
11. **Nagaya, T., T. Nakamura, T. Tokino, T. Tsurimoto, M. Imai, T. Mayiomi, K. Kamino, K. Yamamura, and K. Matsubara.** 1987. The mode of hepatitis B virus DNA integration in chromosomes of human hepatocellular carcinoma. Genes Dev. **1:**773–782.
12. **Newbold, J. E., H. Xin, M. Tencza, G. Sherman, J. Dean, S. Bowden, and S. Locarnini.** 1995. The covalently closed duplex form of the hepadnavirus genome exists in situ as a heterogeneous population of viral minichromosomes. J. Virol. **69:**3350–3357.
13. **Loeb, D. D., R. C. Hirsch, and D. Ganem.** 1991. Sequence-specific RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B virus: implications for other reverse transcribing elements. EMBO J. **10:** 3533–3540.
14. **Ogston, C. W., G. J. Jonak, C. E. Rogler, S. M. Astrin, and J. Summers.** 1982. Cloning and structural analysis of integrated woodchuck hepatitis virus sequences from hepatocellular carcinomas of woodchucks. Cell **29:**385–395.
15. **Rogler, C. E., and J. Summers.** 1984. Cloning and structural analysis of integrated woodchuck hepatitis virus sequences from a chronically infected liver. J. Virol. **50:**832–837.
16. **Seeger, C., J. Summers, and W. W. S. Mason.** 1991. Viral DNA synthesis. Curr. Top. Microbiol. Immunol. **168:**41–61.
17. **Shih, C., K. Burke, M. J. Chou, J. B. Zeldis, C. S. Yang, C. S. Lee, K. J. Isselbacher, J. R. Wands, and H. Goodman.** 1987. Tight clustering of human hepatitis B virus integration sites in hepatomas near a triple-stranded region. J. Virol. **61:**3491–3498.
18. **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.
19. **Summers, J., and W. Mason.** 1982. Replication of the genome of hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell **29:**403–415.
20. **Summers, J., P. Smith, M. Huang, and M.-S. Yu.** 1991. Morphogenetic and regulatory effects of mutations in the envelope proteins of an avian hepadnavirus. J. Virol. **65:**1310–1317.
21. **Takada, S., Y. Gotoh, S. Hayashi, M. Yoshida, K. Koike.** 1990. Structural rearrangement of integrated hepatitis B virus DNA as well as cellular flanking DNA is present in chronically infected hepatic tissues. J. Virol. **64:**822–828.
22. **Transy, C., C.-A. Renard, and M.-A. Buendia.** 1994. Analysis of integrated ground squirrel hepatitis virus and flanking host DNA in two hepatocellular carcinomas. J. Virol. **68:**5291–5295.
23. **Triglia, T., M. G. Peterson, and D. J. Kemp.** 1988. A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res. **16:**8186.
24. **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.
25. **Wang, H. P., and C. E. Rogler.** 1991. Topoisomerase I-mediated integration of hepadnavirus DNA in vitro. J. Virol. **65:**2381–2392.
26. **Wang, G., and C. Seeger.** 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J. Virol. **67:**6507–6512.
27. **Yaginuma, K., H. Kobayashi, M. Kobayashi, T. Morishima, K. Matsuyama, and K. Koike.** 1987. Multiple integration site of hepatitis B virus DNA in hapatocellular carcinoma and chronic active hepatitis tissues from children. J. Virol. **61:**1808–1813.
28. **Yaginuma, K., M. Kobayashi, E. Yoshida, and K. Koike.** 1985. Hepatitis B virus integration in hepatocellular carcima DNA: duplication of cellular flanking sequences at the integration site. Proc. Natl. Acad. Sci. USA **82:** 4458–4462.
29. **Yang, W., and J. Summers.** 1995. Illegitimate replication of linear hepadnavirus DNA through nonhomologous recombination. J. Virol. **69:**4029–4036.