JOURNAL OF VIROLOGY, July 1995, p. 4283–4291 Vol. 69, No. 7
0022-538X/95/$04.00+0

RNA Sequences Controlling the Initiation and Transfer of Duck Hepatitis B Virus Minus-Strand DNA

JOHN E. TAVIS[1]* AND DON GANEM[1,2,3]

Departments of Microbiology and Immunology[1] and Medicine[2] and Howard Hughes Medical Institute,[3] University of California—San Francisco, San Francisco, California 94143-0502

Received 24 January 1995/Accepted 24 March 1995

Hepadnaviruses replicate by reverse transcription of an RNA pregenome. Reverse transcription initiates within the stem-loop (SL) of the ε RNA packaging signal and is discontinuous: the nascent minus-polarity DNA is transferred to direct repeat 1 (DR1) at the 3′ end of the pregenomic RNA prior to extensive elongation. In this study we analyzed the initiation and transfer of duck hepatitis B virus minus-strand DNA by using functional viral polymerase expressed in yeast cells. We extensively mutagenized both DR1 and the SL and observed the effects on reverse transcription initiation and on the transfer and subsequent extension of minus-strand DNA. Our results indicate that sequences throughout the SL affect initiation and that minus-strand DNAs initiated at three locations within the SL are competent for transfer to DR1. A short region of homology between the 5′ end of minus-strand DNA and DR1 was necessary but not sufficient to direct the transfer and subsequent extension reactions. This homology was tolerant of minor substitutions, and 2 nucleotides of homology mediated transfer accurately. Mutations had greater detrimental effects on transfer and subsequent extension of minus-strand DNA when they were placed in DR1 than when they were placed in the SL. Efficient transfer of minus-strand DNA from a mutant SL to DR2 was observed in the yeast system. The hexanucleotide AAUUAC was identified as the primary *cis* element of the transfer acceptor, but this element was also insufficient to independently specify the acceptor location. Therefore, additional information, possibly positional context or unrecognized RNA secondary structure, is required.

The hepatitis B viruses (hepadnaviruses) are unique among animal DNA viruses in that they replicate their genomes by reverse transcription. The template for reverse transcription is a terminally redundant mRNA referred to as the pregenomic RNA (pgRNA) (Fig. 1). Within the redundancy of the pgRNA are signals important for reverse transcription, especially direct repeat 1 (DR1) and the stem-loop (SL) of the packaging signal, ε. Encapsidation of the pgRNA requires that the viral reverse transcriptase (P) recognize the pgRNA by binding to the 5′ SL; without this binding, neither P nor the pgRNA is incorporated into viral cores (1, 3). In addition to its role in RNA packaging, the P-SL binding reaction also plays a central role in genomic replication. The origin of reverse transcription is within the SL, and hence the P-SL binding reaction provides the sequence specificity for the initiation of DNA synthesis as well as for packaging of the appropriate message (2, 4, 8, 10).

The SL was not identified as the origin of reverse transcription until recently because the 5′ end of minus-strand DNA is not located within the SL-coding sequences in the mature viral genome but rather is found within DR1. This implies that the minus-strand DNA chain must have been transferred to DR1 following initiation within the SL. The minus-strand DNA transfer reaction was directly confirmed for duck hepatitis B virus (DHBV) by introducing mutations into the SL template and observing transfer of these mutations to DR1 (8, 10). These and other experiments also established that transfer was from the 5′ copy of the SL (the same copy of the SL to which P binds prior to encapsidation) to the 3′ copy of DR1 within the pgRNA (Fig. 1). The template transfer must occur very early in the synthesis of minus-strand DNA, as only 4 nucleotides (nt) (3 nt in the mammalian hepadnaviruses) is identical between the initiation site within the SL and the DR1 acceptor sequences. It is this short homology between the nascent minus-strand DNA and the pgRNA at DR1 that presumably allows hybridization and mediates strand transfer.

We have been studying the initiation of DHBV minus-strand DNA synthesis and the transfer of these DNAs to DR1 by employing DHBV P protein expressed in yeast cells. Our system uses the retrotransposon TY as a vector and produces functional DHBV P as a fusion protein with the TY capsid-like protein, TYA (the fusion protein is analogous to a retroviral *gag-pol* polyprotein). The TYA-DHBV P fusion (TYDP) is packaged into virus-like particles (VLPs) along with its chimeric TY-DHBV mRNA by TYA. The 3′ noncoding region of the encapsidated chimeric mRNA contains both DR1 and the SL (Fig. 1). TYDP recognizes the SL at the 3′ end of the chimeric RNA, initiates reverse transcription at the authentic position within the SL (DHBV nt 2576), and accurately transfers the nascent minus-strand DNA to DR1 (5′ end at nt 2537) (Fig. 1) (7, 8).

Four questions concerning the DHBV minus-strand DNA synthesis mechanism were addressed in this study. (i) What are the features of the SL that specify the initiation site of minus-strand DNA synthesis? (ii) What is the role of the 4-nt homology between the SL and DR1 in mediating transfer of the minus-strand DNA to DR1 and its subsequent elongation? (iii) Is minus-strand DNA transferred exclusively to DR1, or can other sites in the viral genome (particularly DR2) act as acceptors? (iv) What is the nature of the RNA sequences constituting the acceptor site for minus-strand DNA transfer to DR1?

* Corresponding author. Present address: Department of Molecular Microbiology and Immunology, St. Louis University School of Medicine, 1402 S. Grand Blvd., St. Louis, MO 63104. Phone: (314) 577-8441. Fax: (314) 773-3403. Electronic mail address: TAVISJE@SLUVCA.SLU.EDU.

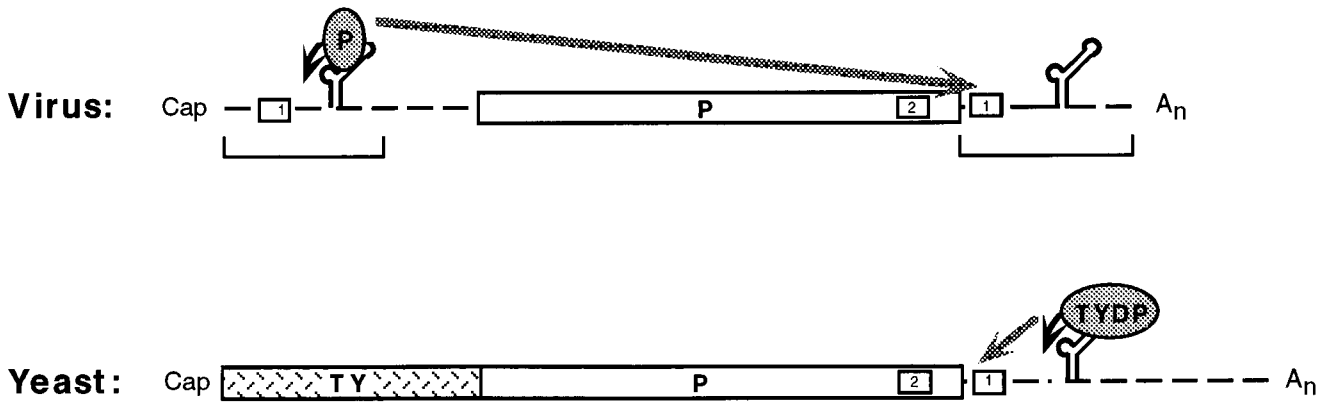


FIG. 1. Template RNAs and minus-strand DNA transfer in DHBV and in the yeast TYDP system. (Top) Structure of the viral pgRNA. The pgRNA is capped and polyadenylated and is terminally redundant (brackets). DR1 and DR2 are indicated by boxes with 1 and 2, respectively, and the SL of ε is drawn. The P ORF is shown by an open box labeled P. The viral polymerase is indicated by a shaded oval labeled P, and the protein-linked nascent minus-strand DNA is shown by a solid arrow. Transfer of the minus-strand DNA from the 5′ SL to the 3′ DR1 is indicated by a shaded arrow. (Bottom) Chimeric TY-DHBV mRNA found in the VLPs. This RNA is capped and polyadenylated and contains Ty sequences at its 5′ end (hatched box); TYA is fused to DHBV P by ribosomal frameshifting. DHBV noncoding sequences are shown by a dashed line. The TYDP fusion polymerase is shown by the shaded oval labeled TYDP; DR1, DR2, the SL, and the nascent minus-strand DNA are as in the top panel. The transfer of the minus-strand DNA from the SL to DR1 is shown by a shaded arrow.

MATERIALS AND METHODS

DNA constructs and nomenclature. All constructs are derivatives of the yeast shuttle plasmid pTYBDP-DR1/SL (8). pTYBDP-DR1/SL contains the DHBV type 3 P open reading frame (ORF) and 3′ noncoding sequences within the yeast retrotransposon TY1-H3; expression of these sequences is under the control of the inducible *GAL1* promoter. Mutations introduced into pTYBDP-DR1/SL are indicated by either DRM or SLM, with numbers, for DR mutant or SL mutant, respectively (e.g., pTYBDP-DRM2/SLM11). The prefix pTYBDP- has been omitted below for simplicity. Mutations at DR1 or the SL can be reassorted by virtue of a *Hin*dIII site created between the elements (DHBV nt 2555 and 2556, CT to AA); this alteration had no effect on minus-strand DNA synthesis or transfer (8). All mutations either were downstream of the P termination codon at nt 2530 or were silent in the P ORF.

Mutagenesis. Mutations were introduced by PCR with mutant primers. PCR products were cloned directly into pTYBDP-DR1/SL or its derivatives, and the entire PCR-derived sequence was confirmed to ensure the absence of secondary mutations.

Isolation of VLPs and VLP-associated DNA. VLPs were isolated as described previously (7). Briefly, yeast cultures containing TYDP expression plasmids were induced by addition of galactose, and 20 to 22 h later cells were collected and lysed. The clarified lysate was layered onto a three-layer sucrose step gradient (60, 30, and 20%) and centrifuged at 100,000 × *g* for 3 h. Particulate matter at the interface between the 60 and 30% sucrose steps was collected, concentrated by centrifugation, and suspended at a concentration of 1 μg/μl; such extracts are highly enriched for VLPs. DNA was isolated from these extracts by proteinase K digestion, phenol and chloroform extraction, and ethanol precipitation as described previously (7).

Primer extension. The 5′ ends of the DHBV minus-strand DNAs produced in the chimeric VLPs were detected by reiterative primer extension with *Taq* DNA polymerase and a thermocycler as described previously (8). The primer to detect 5′ ends at DR1 and the SL was DHBV nt 2453 to 2475 (plus polarity), and the primer to detect 5′ ends at DR2, DR1, and the SL was nt 2294 to 2330 (plus polarity). The primer extension products produced by *Taq* DNA polymerase migrate roughly 1.5 bp more slowly than the corresponding position indicated by the homologous sequencing ladders included in the figures. This discrepancy is because (i) the primer for the sequencing ladder is unphosphorylated, while the primer for the primer extension is phosphorylated, and (ii) under the conditions employed, *Taq* DNA polymerase produces products longer than the template, and since the polymerase lacks a 3′-to-5′ exonuclease activity, the untemplated nucleotide(s) is not removed. This effect was characterized in preliminary primer extensions with plasmid DNA linearized within the coding sequence for the bulge of the SL (pTYBDP-SLM2 cut with *Pml*I) and compared with a homologous sequencing ladder (data not shown).

Primer extensions performed in this manner were highly reproducible and upon repeated analysis revealed the same ratio of ends at DR1 and within the SL for DNA from a given construct. All 21 samples in Fig. 8A have the same SL sequence and, with the exception of DRM18, give the same ratio of ends within the bulge and upper stem. It is not known why the upper stem position of DRM18 is activated, but this too was internally consistent upon repeat analyses of this DNA. The major difference between these primer extensions and those previously described (8) is the appearance of an additional set of bands between DR1 and the SL. These bands were determined to be artifacts for the following reasons: (i) they were also detected in control primer extension samples with linearized plasmid templates, (ii) they became apparent when the lot of *Taq* DNA polymerase employed in the reactions was changed, and (iii) they appeared with a stronger intensity relative to that of the authentic bands in primer extension reactions run through a greater number of cycles. The intensity of these bands also varies with the sequence at the SL, presumably because the sequence alterations affect the local DNA structure and therefore affect the premature termination by *Taq* DNA polymerase.

RESULTS

Reverse transcription initiates at multiple sites. Primer extension analysis of DNA isolated from the TYDP mutants shown in Fig. 2 revealed that reverse transcription initiated (and extended at least as far as nt 2453, the site of the primer employed for primer extension) at two sites within the wild-type SL. The dominant position is nt 2576, within the bulge of the SL (Fig. 2, lane 1 and diagramed with a heavy arrow in the top line at right). Initiation and extension at a minor position accounting for roughly 10% of the minus-strand DNAs are at nt 2582, within the upper stem of the SL (Fig. 2, lane 1; light arrow in the top line of the diagram) (although we interpret the DNAs detected at this position as initiation events, other interpretations are possible [see Discussion]). Mutating 5 of the 6 nt in the bulge of the SL (including the template nucleotide for the dominant initiation site) alters the detection of 5′ ends in two ways (SLM2; Fig. 2, lane 2). First, the start site in the bulge is shifted 5′ by 2 nt to the nearest template C, and second, minus-strand DNA within the upper stem at nt 2582 is detected with an efficiency equal to that for the bulge initiation site. Not all bulge mutations have these effects; a smaller bulge mutation (SLM3; 2 of 6 nt altered without affecting nt 2576) reduces initiation and extension at nt 2576 and eliminates detectable minus-strand DNA at nt 2582 (Fig. 2, lane 3). Altering the template nucleotide for the minor initiation site (nt 2582 from C to U [and nt 2599 from G to A to maintain base pairing in the upper stem]) has no effect on initiation and extension within the bulge, eliminates minus-strand DNA at nt 2582, and generates a minor initiation-extension product opposite an A at nt 2581 (SLM10; Fig. 2, lane 4 [compare the position of the upper stem band with that of the initiation at nt 2582 in lane 5]).

Altering the apical loop of the SL also affects the minus-strand DNA 5′ ends. Mutating nt 5 and 6 of the loop eliminates

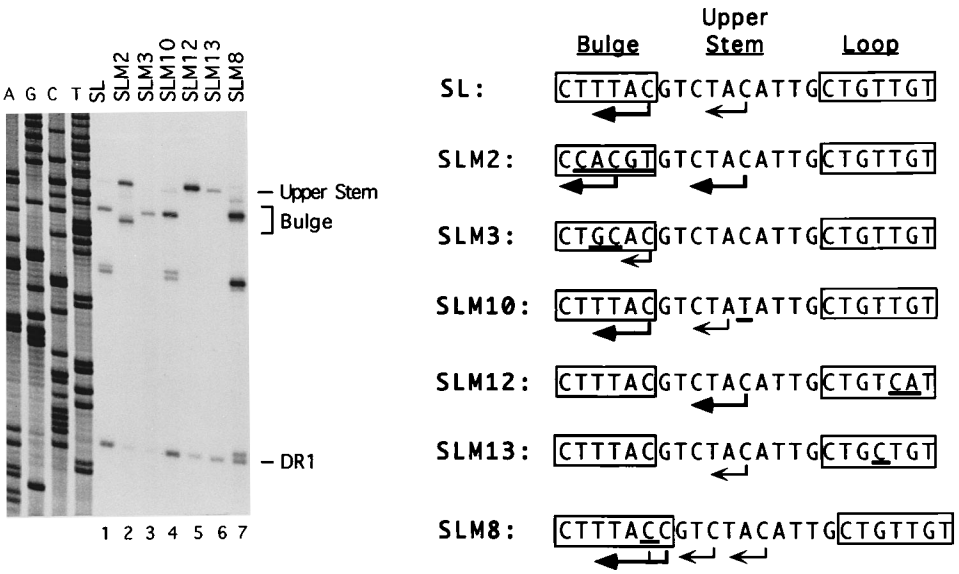


FIG. 2. Mutations throughout the SL affect initiation of reverse transcription. (Left) Primer extension analysis of chimeric VLP-derived DNAs that detects minus-strand DNA 5′ ends mapping to DR1, to the bulge of the SL, and to the upper stem of the SL. A sequencing ladder from the same primer is shown in the left four lanes. The wild-type SL is shown in lane 1, and mutant SLs are shown in lanes 2 to 7. The identities of the mutants are indicated above the lanes, and all SL constructs are paired with wild-type DR1. (The doublet bands between the DR1 and bulge positions are artifacts [see Materials and Methods].) (Right) DNA coding sequences for the bulge, upper stem, and loop of the SL for the constructs employed. The sequence for the wild-type SL is shown in the top line, and sequences for the mutant SLs are shown below the wild-type sequence. Nucleotides in the bulge and loop are boxed, and mutant nucleotides are underlined. Minus-strand DNA 5′ ends are indicated with arrows below the sequence, and the widths of the arrows semiquantitatively represent the relative strengths of the observed signals.

initiation and extension from nt 2576 and increases the detection of minus-strand DNAs within the upper stem at nt 2582 (SLM12; Fig. 2, lane 5). (This represents a rare discordance between the TYDP system and the in vitro DNA priming reaction described by Wang and Seeger [9]. In that assay, this double mutation appears to reduce overall P-linked primer synthesis 10-fold [4]. The reason for this different behavior is unknown.) A substitution at loop nt 4 also eliminates initiation and extension from nt 2576, but this alteration does not increase the detection of minus-strand DNAs at nt 2582 (SLM13; Fig. 2, lane 6).

The polymerase can also initiate reverse transcription from the SL when the bulge has been enlarged by 1 nt. In this case (SLM8; Fig. 2, lane 7) initiation occurs at two locations within the bulge, both at the normal location at the 3′ end of the bulge and also 1 nt 5′ to this position. Minor initiation-extension products are also detected within the upper stem.

These results indicate that P can initiate reverse transcription from multiple sites within the bulge and upper stem of the SL (see Discussion). The polymerase prefers to initiate with a dGTP residue opposite a template C, but under certain situations it can initiate opposite a template A (presumably by using TTP).

Minus-strand DNAs from both the bulge and upper stem can be transferred to DR1. Three reverse transcription start sites in the SL have short sequences of homology to the acceptor site at DR1 (Fig. 3, shaded sequences in diagram). The homology between DR1 and the dominant start site at nt 2576 in the bulge is 4 nt, that between DR1 and the upper stem site at nt 2582 is 3 nt, and that between DR1 and the ectopic bulge site in SLM2 is 2 nt. We created a series of mutations to determine which of these nascent minus-strand DNAs were competent for transfer to DR1 and subsequent elongation. In mutant DR1/SLM10, minus-strand DNAs at the upper stem position 2582 have been ablated, and the starts at nt 2581 are not detectably transferred to the DR1 acceptor site, or if they are transferred, they are not subsequently extended (such transfers would be expected to result in minus-strand DNAs 1 nt shorter than those resulting from wild-type transfers to DR1). Minus-strand DNAs transfer and extend efficiently in this mutant, demonstrating transfer of DNAs from nt 2576 in the bulge to DR1 (Fig. 3, lane 2; transfer from the bulge is also reported in reference 8). Mutant DR1/SLM2 initiates strongly at an ectopic position within the bulge (nt 2574) and also within the upper stem; minus-strand DNA is transferred and extended with moderate efficiency in this case (Fig. 3, lane 3). To determine if minus-strand DNA within the mutant bulge could be transferred, the upper stem initiation site was ablated in mutant DR1/SLM11. In this case transfer and extension occurred, albeit weakly (Fig. 3, lane 4 [see Fig. 4, lane 3, for a darker exposure of same mutant]). Minus-strand DNA initiated within the upper stem could also be transferred to DR1, as is shown by mutant DR1/SLM12; there are no detectable starts within the bulge, but transfer to DR1 still occurs with reasonable efficiency (Fig. 3, lane 5). These results indicate that minus-strand DNA that has initiated at the wild-type location within the bulge, at an ectopic site within the bulge, or within the upper stem can be transferred to DR1 and subsequently be elongated. They also indicate that 2 nt of homology between the nascent minus-strand DNA and DR1 is sufficient to mediate transfer and extension (albeit inefficiently).

Complementarity is necessary but not sufficient for transfer and subsequent extension. We next created a series of mutations in the SL and DR1 to assess the contribution of the homology between the nascent minus-strand DNA and DR1 to the transfer reaction. DR1 was altered in mutant DRM1/SL to remove homology between the minus-strand DNA and the acceptor site; this alteration eliminated transfer of minus-strand DNA or extension of the minus-strand DNA following transfer (Fig. 4, lane 2), indicating that homology is necessary for a successful transfer-extension reaction. In mutant DR1/SLM11 there is a 2-nt homology between the nascent minus-

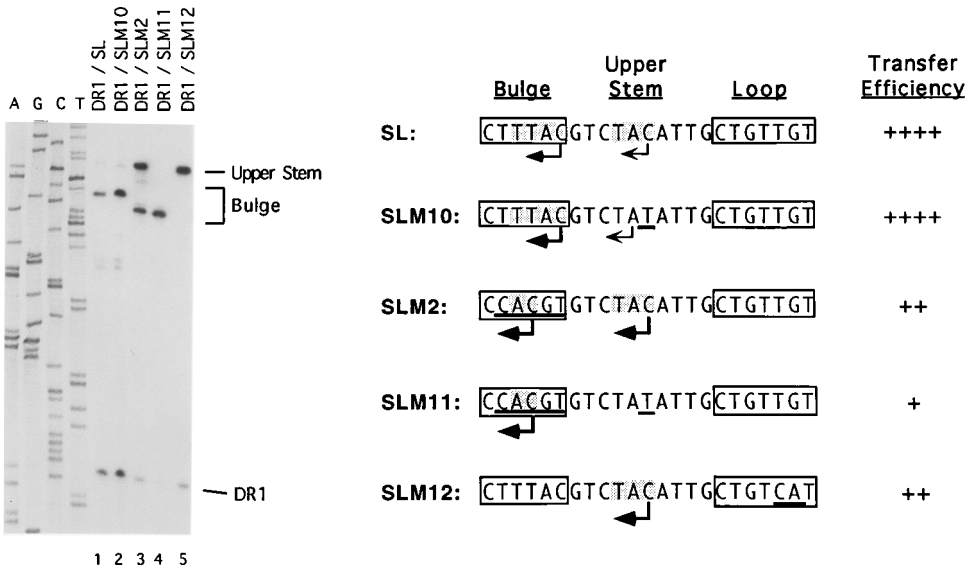


FIG. 3. Minus-strand DNAs from both the bulge and the upper stem can be transferred to DR1. (Left) Primer extension analysis of chimeric VLP-derived DNAs, as in Fig. 2. The wild-type construct (DR1/SL) is shown in lane 1, and constructs bearing wild-type DR1 and mutant SL sequences are shown in lanes 2 to 5. (Right) Coding sequences for the bulge, upper stem, and loop of the SL for the constructs employed. The symbols are as in Fig. 2. In addition, nucleotides homologous to the acceptor at DR1 are shaded, and a visual estimate of the relative transfer efficiency is given (++++, wild-type efficiency; +, barely detectable).

strand DNA and DR1, and in this case transfer and extension occur, but very weakly (Fig. 4, lane 3). Mutations were then inserted into DR1 in DR1/SLM11 to create mutant DRM2/SLM11, in which the homology to the nascent minus-strand DNA of SLM11 was expanded from 2 to 6 nt. Expanding the homology between the nascent minus-strand DNA and DR1 did not improve transfer and extension but rather reduced them to nearly background levels (Fig. 4, lane 4). We also examined two other mutants with 2 nt of homology between the nascent minus-strand DNA and DR1; both completely failed to transfer minus-strand DNA or to extend the DNAs once transferred (data not shown). These data indicate that homology between the nascent DNA and DR1 is necessary but not sufficient for successful transfer and extension and that there is no simple relationship between the number of nucleotides of homology and the efficiency of transfer and/or subsequent extension.

Mutations in DR1 and the SL affect efficiency of transfer and extension differently. The relative effect of mutations in DR1 or the SL on the efficiency of transfer and extension was assessed by inserting identical mutations into either DR1 or the SL and observing their effects. In mutant DR1/SLM3 two substitutions (GC) were created in the bulge so that there was a 2-nt homology (AC) to DR1. In this mutant transfer and extension of minus-strand DNA were moderately efficient (Fig. 5, lane 2). When the same two substitutions (GC) were placed in DR1, generating an identical 2-nt homology (AC) between DR1 and the SL, they had a much more detrimental effect on

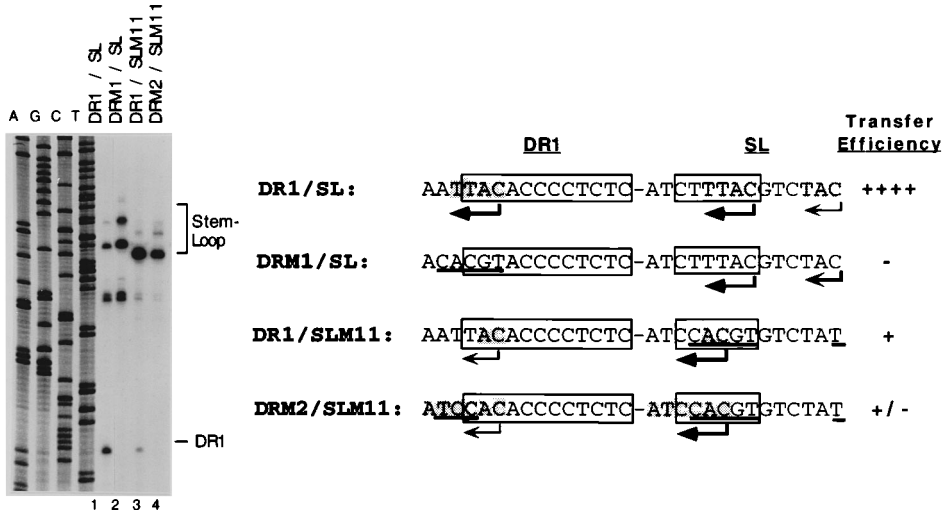


FIG. 4. Homology between the SL and DR1 is necessary but not sufficient for minus-strand DNA transfer to and subsequent extension from DR1. (Left) Primer extension of chimeric VLP-derived DNAs, as in Fig. 2. The wild-type construct (DR1/SL) is shown in lane 1, and constructs bearing mutations at DR1 and/or SL are shown in lanes 2 to 4. (Right) Coding sequences for DR1 and for the bulge and upper stem of the SL for the constructs employed. Sequences of DR1 and the bulge are boxed; all other symbols are as in Fig. 2 and 3. −, no detectable transfer to DR1; +/−, transfer to DR1 at the lower limit of detectability.

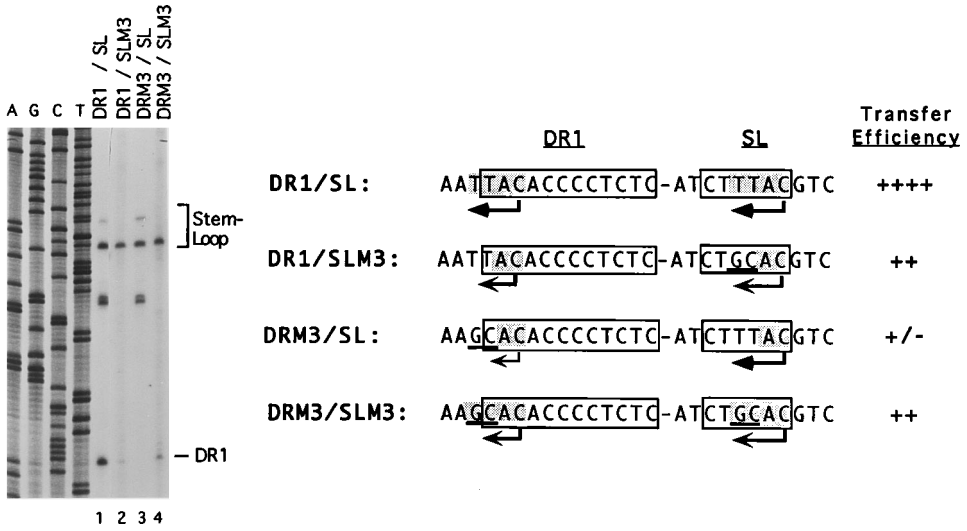


FIG. 5. Mutations affect minus-strand DNA transfer and extension efficiency to different extents when placed in DR1 or SL. (Left) Primer extension of chimeric VLP-derived DNAs, as in Fig. 2. The wild-type construct (DR1/SL) is shown in lane 1, and constructs bearing mutations at DR1 and/or SL are shown in lanes 2 to 4. (Right) Coding sequences for DR1 and for the bulge of the SL for the constructs employed. All symbols are as in Fig. 2 to 4.

the efficiency of transfer and extension (DRM3/SL; Fig. 5, lane 3). The greater detrimental effect of mutations to DR1 than to the SL is a general feature of the transfer reaction. We analyzed over 30 mutants with mutations in either DR1 or the SL and observed that the transfer-extension machinery tolerates mutations to the SL donor element better than it does alterations to the DR1 acceptor motif.

Transfer of minus-strand DNAs to DR2. Minus-strand DNA is normally transferred to DR1 during reverse transcription. We asked if minus-strand DNA could be transferred to and subsequently extended from the second copy of the same sequence, termed DR2, found 42 nt 5′ of DR1 in the template RNA. Primer extension analysis of DHBV viral DNA isolated from intracellular cores produced in LMH cells detected very low levels of DHBV minus-strand DNA 5′ ends at DR2 (nt 2479) (data not shown). These ends were estimated to account for less than 1% of the DNA chains.

The presence of minus-strand DNA 5′ ends at DR2 was also assessed in DNA extracted from the chimeric TYDP constructs shown in Fig. 6 by primer extension with a primer that annealed to the minus-strand DNA downstream of DR2. This primer allowed the simultaneous detection of ends mapping to DR2, DR1, and the SL. Analysis of DNA from DR1/SL revealed no minus-strand DNA 5′ ends at DR2 (DR1/SL; Fig. 6, lane 1), and disruption of the normal acceptor sequence at DR1 resulted in the detection of only trace levels of 5′ ends at DR2 (DRM1/SL; Fig. 6, lane 2). However, when the mutant DR1/SLM2 was analyzed, minus-strand DNA 5′ ends were detected with equal efficiency at DR1 and DR2 (Fig. 6, lane 3). SLM2 contains a 5-of-6-nt substitution in the bulge of the SL

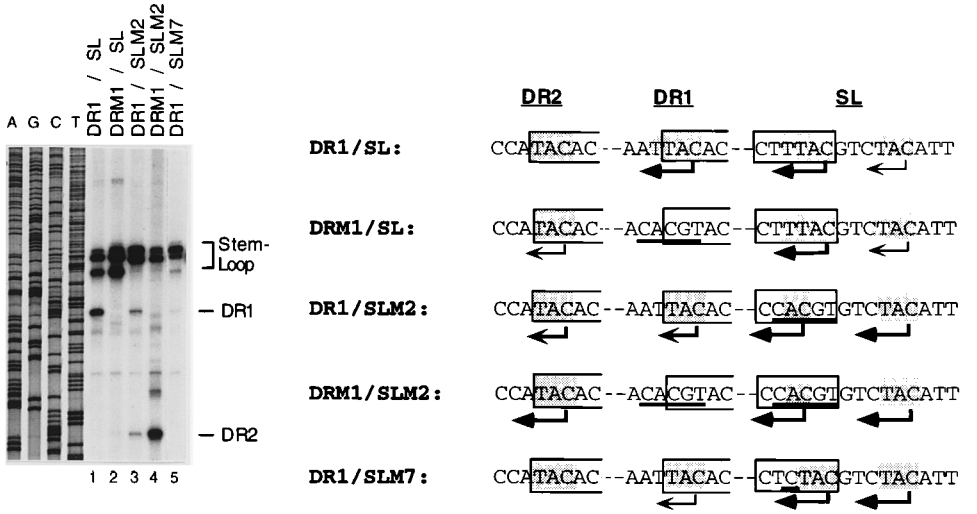


FIG. 6. Minus-strand DNA can be transferred to DR2 in yeast cells. (Left) Primer extension analysis of chimeric VLP-derived DNAs with a primer that detects 5′ ends at DR2, DR1, and the SL. The wild-type construct (DR1/SL) is shown in lane 1, and constructs bearing mutations at DR1 and/or SL sequences are shown in lanes 2 to 5. A sequencing ladder from pTYBDP-DR1/SL with the oligonucleotide employed for primer extension is shown. (Right) Coding sequences for DR2, DR1, and the bulge and upper stem of the SL for the constructs employed; sequences of DR2, DR1, and the bulge are boxed. The symbols are as in Fig. 2 and 3.

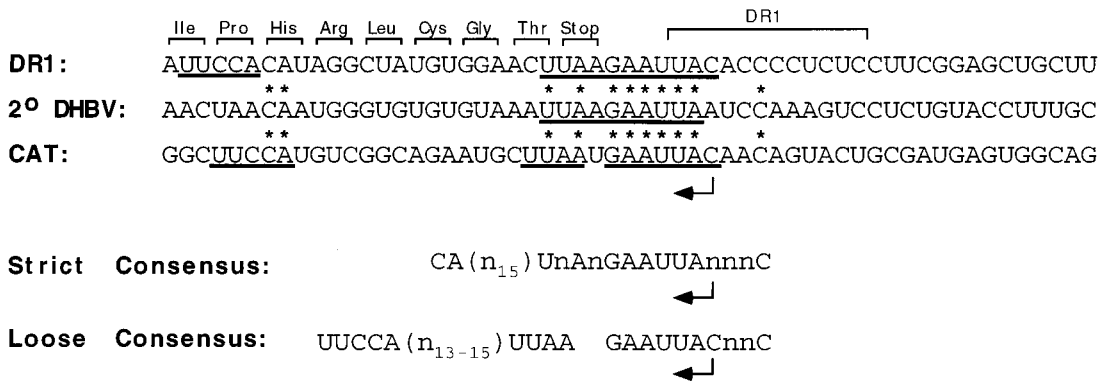


FIG. 7. Minus-strand DNA transfer acceptor consensus sequence. The RNA sequences surrounding DR1 (DHBV nt 2504 to 2560), the secondary (2°) DHBV acceptor site (nt 2094 to 2150), and the CAT acceptor (CAT ORF nt 592 to 648) are shown at top. The translation of the P ORF, DR1 (bracket), and the 5′ end of minus-strand DNA (arrow) are shown. Perfectly conserved positions constituting the strict consensus are indicated by asterisks between the sequences, and prominent motifs forming the loose consensus are underlined. Also shown are representations of the two consensus sequences; n represents any nucleotide, and the discontinuity in the loose consensus represents the additional U in the CAT element.

(diagram in Fig. 6), and so alteration of donor sequences had modified the target site choice. When DR1 was disrupted in DRM1/SLM2, utilization of DR2 in minus-strand DNA transfer and extension was further enhanced (Fig. 6, lane 4), indicating that there may be a competition between DR1 and DR2 for transfer of minus-strand DNA from SLM2.

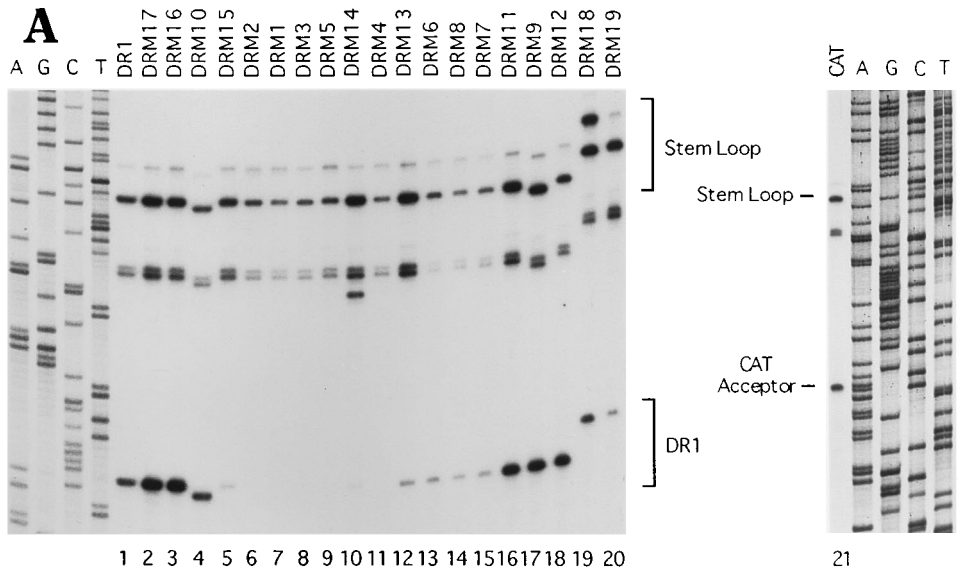
We noted that the upper stem start site at nt 2582 in SLM2 was strongly activated and that this site possessed a 3-nt homology to the acceptor sites at both DR1 and DR2. Mutant SLM7 was employed to test whether production of large amounts of nascent minus-strand DNA with 3 nt of homology to both acceptor sites was responsible for the transfer to DR2. SLM7 contains a single-nucleotide substitution in the bulge and was used because it codes for a nascent primer with the same 3-nt homology as does the upper stem site and because an increase in 5′ ends is detected within the upper stem, similar to those found within SLM2. Primer extension analysis of minus-strand DNA isolated from DR1/SLM7 did not detect 5′ ends at DR2 (Fig. 6, lane 5). Also, ablation of the upper stem start site of SLM2 did not eliminate transfer to DR2 (data not shown). These data indicate that DNA from the ectopic bulge initiation in SLM2 can be transferred to DR2 and be subsequently extended. They also show that a feature of SLM2 other than homology between the nascent minus-strand DNA and the acceptor site is responsible for the efficient transfer to and extension from DR2. The biochemical activity represented by this aberrant transfer and extension either is masked in the context of the native virion or is unique to the TYDP system, because placing the SLM2 mutations in the 5′ SL of DHBV does not result in an elevated level of transfer to DR2 (data not shown).

Identification of the minus-strand DNA transfer acceptor sequence. We next determined the primary sequence element required to form a functional transfer acceptor at DR1. The starting point for this analysis was the observation that in DHBV constructs in which DR2 was disrupted, a DNA species whose 5′ ends mapped to a novel site was generated. Similarily, in DHBV constructs carrying an insertion of the chloramphenicol acetyltransferase (CAT) gene, a new viral minus-strand DNA whose 5′ end mapped within the CAT sequences was identified (3a). Alignment of the sequences at the 5′ ends of the minus-strand DNAs from these two ectopic sites with the wild-type sequence at DR1 allowed the generation of a consensus acceptor sequence. Figure 7 shows the aligned sequences and two potential interpretations. The "strict consensus" allows for no mismatches or insertions among the sequences, whereas the "loose consensus" allows minor insertions and mismatches and also recognizes the potential effect of the sequence UUCCA 13 to 15 nt 5′ of the main consensus sequence in the two strongest acceptors (DR1 and CAT).

The validity of testing the consensus sequence in the yeast TYDP system was checked by determining the ability of CAT sequences to function as a minus-strand DNA transfer acceptor in yeast cells. One hundred nucleotides surrounding the minus-strand DNA 5′ ends observed in the CAT-containing DHBV constructs was inserted into pTYBDP-DR1/SL in place of DR1 (the P ORF and the SL were left intact). Chimeric VLPs were prepared, and protein-linked minus-polarity CAT DNAs were assayed by primer extension. DNAs with 5′ ends mapping to the CAT acceptor site were easily detected; the efficiency of transfer from the SL to the CAT site and subsequent extension was comparable to that of transfer to and extension from DR1 (Fig. 8A; compare lane 21 with lane 1). Utilization of the CAT acceptor site appears to proceed as efficiently in yeast cells as it does in transfected avian cells.

The loose consensus was used to guide a mutational analysis of potential *cis* acceptor site sequences. Figure 8A shows a primer extension analysis of the minus-strand DNA transfer-extension reaction from the wild-type SL to 19 DR1 sequences with alterations to the loose consensus sequence. Figure 8B summarizes the data in Fig. 8A; the shaded group of constructs indicates those mutants with greatly reduced transfer-extension efficiency. All of these mutants have alterations in the hexanucleotide AAUUAC (DHBV nt 2532 to 2537); the upstream motif UUCCA does not appear to be required for transfer. Not all alterations to the hexanucleotide render it completely inactive; low levels of transfer to and extension from hexanucleotides containing single-nucleotide changes (DRM14 and DRM13; Fig. 8A, lanes 10 and 12, respectively) or a hexanucleotide carrying a double mutation to the 5′ AA (DRM15; lane 5) are observed. The AAUUAC hexanucleotide contains at its 3′ end the UUAC homology between the nascent minus-strand DNA and DR1.

Two additional mutants (DRM18 and DRM19) were constructed to test the role of sequences adjacent to the AAUUAC hexanucleotide. This mutagenesis is complicated by the fact that the hexanucleotide overlaps the last two codons of the P ORF. Accordingly, silent mutations were created in the last two codons of the P ORF, and 4 nt was inserted 3′ to the P stop codon. In addition, all nucleotides from the 3′ end of the



B

Mutant	Ile Pro Thr * DR1 HinDIII	Transfer Efficiency
WT	AUUCCA(N15)ACUUAAGAAUUACACCCUCUCCUUCGGAGAAGCUU	++++
DRM17	..C..G..	++++
DRM16	..........C.G.A..	++++
DRM10	..........d..	++++
DRM15	..........GG..	+
DRM2	..........UCC..	−
DRM1	..........CACGU..	−
DRM3	..........GC..	+/−
DRM5	..........GCGU..	−
DRM14	..........G..	+
DRM4	..........GU..	−
DRM13	..........U..	+
DRM6	..........G..	+++
DRM8	..........GU..	+++
DRM7	..........U..	+++
DRM11	..........U..	++++
DRM9	..........d..	++++
DRM12	..........AGAU..	++++
DRM18	..........G.G.C.....UAGAUCGAUAGAACUUC..... (.G.C)	+++
DRM19	..........G.G.C.....UUAGAUCGAUAGAACUUC..... (.G.C)	++
CAT	C.....(N13)CU.A.U.....A.AG.ACUGCGAU...UGG.AG	++++

FIG. 8. The minus-strand DNA transfer acceptor element. (A) Primer extension analysis of acceptor consensus sequence mutants. Primer extension of chimeric VLP-derived DNAs is shown as in Fig. 2. The wild-type construct (DR1/SL) is shown in lane 1, and constructs bearing mutations at DR1 are shown in lanes 2 to 20. All constructs contain the wild-type SL. The right panel shows a primer extension analysis of DNA isolated from TYBDP-CAT/SL VLPs (lane 21) and a sequencing ladder of pTYBDP-CAT/SL with the oligonucleotide employed for primer extension. (B) Summary of primer extension data from panel A. The top line shows the RNA sequence of DR1/SL (wild type [WT]; DHBV nt 2504 to 2560). The translation of the P ORF (*, stop codon) and brackets denoting the positions of DR1 and the *Hin*dIII site at the base of the SL are indicated above the wild-type sequence. The loose consensus sequence is indicated by shading in the top line, and the 4-nt homology between the nascent minus-strand DNA and DR1 in the wild-type DR1/SL is shown by an arrow. Mutants are indicated by name at left, sequences identical to those of the wild type are shown by dots, and the 4-nt insertions in DRM18 and DRM19 are shown by a caret below the sequence. A visual estimation of the relative transfer efficiency is shown at right (++++, wild-type transfer efficiency; −, no detectable transfer. Mutants with greatly impaired minus-strand DNA transfer to or extension from DR1 are within the shaded block.

AAUUAC hexanucleotide to the *Hin*dIII site at the base of the SL were also altered (DRM19 carries an additional alteration at nt 2537). These mutations were constructed to disrupt as much of the consensus sequence as possible while leaving a minimal transfer acceptor site in an altered location. The AAUUAC hexanucleotide functioned efficiently in this context and tolerated a substitution at the 3′ C (Fig. 8A, lanes 19 and 20).

DISCUSSION

The recent identification of a novel reverse transcription initiation mechanism employing transfer of a nascent minus-strand DNA from the SL to DR1 in the hepadnaviruses has highlighted the unique nature of this reaction (8, 10). In this study we have addressed several features of the minus-strand DNA initiation and transfer mechanisms in an effort to increase our understanding of these processes.

Sequences throughout the SL affect initiation of reverse transcription. Mutations were introduced into multiple locations within the SL, and their effects on initiation position and efficiency were observed. Data from these experiments indicate that sequences in the bulge, upper stem, and loop of the SL are all involved in specifying the precise initiation sites and the frequency either with which they are used or with which these DNAs are extended sufficiently to be detected by our primer extension assays. The effects of mutations in the SL are not yet predictable, indicating that their effects on reverse transcription may result from subtle alterations to the SL secondary structure and hence its interaction with P.

A minor population of the minus-strand DNA 5′ ends mapped within the upper stem of the SL, and these DNAs were competent for transfer to DR1. Transfer of minus-strand DNA from the upper stem may be physiologically relevant, because transfer of a primer 3 nt or less in length would result in a wild-type progeny molecule. DNA mapping to the upper stem presumably initiates within the upper stem, but formally it could result from initiation within the bulge followed by transfer to the upper stem (in a manner analogous to the transfer to DR1). We consider transfer of minus-strand DNA from the bulge to the upper stem to be unlikely for four reasons: (i) 5′ ends mapping to the bulge can be ablated without eliminating ends in the upper stem (Fig. 2, SLM12 and SLM13); (ii) alterations at DR1 have no consistent effect on the strength of the upper stem signals (as would be expected if there were two potential destinations for the minus-strand DNA during transfer); (iii) the position of the upper stem site in mutant SLM10 (Fig. 2) is inconsistent with a homology-dependent strand transfer; and (iv) alterations to the donor sequence of the bulge do not affect the upper stem site in a manner consistent with a homology-dependent strand transfer.

Role of homology between the nascent minus-strand DNA and DR1 during transfer and subsequent extension. Homology between the nascent minus-strand DNA and the DR1 acceptor is necessary but not sufficient for transfer and subsequent extension. Lesions at DR1 that eliminate homology prevent transfer and extension, and alterations that expand homology beyond the normal 4 nt do not increase the efficiency of transfer and extension. Also, there is no simple relationship between the number of nucleotides of homology and the observed efficiency of transfer and extension.

Determination of the positions within the nascent minus-strand DNA that must be homologous to DR1 for successful transfer and extension is dependent on the site of arrest of minus-strand DNA synthesis within the bulge. If arrest occurs exclusively 1 nt before the 5′ end of the bulge to create a uniform population of transfer substrates, then the transfer observed with DR1/SLM11 (Fig. 3) indicates that the 3′ base of the nascent minus-strand DNA need not be hybridized to the template. However, if arrest occurs at multiple positions within the bulge prior to transfer, then the 5′ ends observed at DR1 in mutant DR1/SLM11 would indicate that a 2-nt product completely homologous to DR1 can be transferred. (In either case, RNA-protein interactions would probably contribute greatly to the transfer reaction, since 2-nt DNA-RNA hybrids are unlikely to be sufficiently stable.) Our data do not distinguish between these possibilities. However, primers with mispairing at positions outside the 3′ nucleotide clearly can be transferred and extended. For example, SLM8 contains a 1-nt insertion in the bulge; mature minus-strand DNA with a 5′ end at nt 2538 is detected in DR1/SLM8 (Fig. 2, lane 7). This lengthened minus-strand DNA could have been produced only if the 3′ base of the lengthened primer was paired to its complement in the template RNA during transfer. In this case, the 5′ base of the nascent minus-strand DNA would not be hybridized to the RNA.

Determination of the minus-strand DNA transfer acceptor element. Mutational analysis of the consensus acceptor sequence identified the hexanucleotide AAUUAC as the primary *cis* element of the minus-strand DNA transfer acceptor. Transfer and subsequent extension of minus-strand DNA were very specific for the UUAC homology within the hexanucleotide; similar motifs elsewhere in the neighboring sequence were not recognized. The hexanucleotide acceptor motif can account for most of this selectivity, including the rejection of DR2 as an adequate acceptor.

However, this hexanucleotide is not the sole determinant of the acceptor, because it occurs twice in the DHBV genome (nt 2537 and 3011) but transfer-extension is to nt 2537 exclusively (both in DHBV and in the yeast TYDP system). Therefore, additional information must be required to specify a fully functional transfer acceptor site. Sequences 3′ to the hexanucleotide are not required for transfer: DRM18 excludes sequences from the hexanucleotide to the base of the SL (Fig. 8A, lane 19), and all sequences 3′ to DR1 can be deleted in the virus without affecting minus-strand DNA synthesis (5). An RNA element further 5′ than the consensus sequences examined in Fig. 8 may provide the additional information to specify the acceptor site, but if so it would fall within the P ORF, and its analysis would be complicated by the overlapping coding sequences. Further experiments are under way to determine if such an element exists.

Alternatively, the additional information may be provided by positional effects within the RNA or by as-yet-unrecognized RNA secondary structures. However, the additional information is unlikely to be the result of a specific conformation imposed on the RNA by the capsid, because the structures of the yeast VLP and the hepadnavirus capsid differ greatly but transfer occurs correctly and efficiently within both particles.

Clearly, much remains to be learned about the factors influencing the selection of the acceptor site. One illustration of the complexity of this process is our observation that certain lesions in the bulge of the SL allow transfer of minus-strand DNA to the CCAUAC sequence at DR2 as well as to the AAUUAC hexanucleotide at DR1, at least in the TYDP system (Fig. 6). This suggests that features of the donor site can influence acceptor site selection. However, there are at present too few examples of this phenomenon to allow speculation about its mechanism.

Generality of the minus-strand DNA transfer signals. Seeger and Maragos have previously identified a short sequence that is essential for the synthesis of minus-strand DNA in

woodchuck hepatitis B virus (WHV) (UUUC, overlapping the 5′ end of DR1 [5, 6]). The UUUC element tolerates limited substitutions and can be shifted short distances around DR1, but it is not functional at other locations in the pgRNA. These characteristics led to the conclusion that UUUC represents the primary *cis* element of the origin of reverse transcription but that additional information was required to obtain full specificity. Seeger and Maragos speculated that the additional information may be provided by structural features of the RNA or by the interaction of the RNA with the viral capsid (6).

Our observations with DHBV are strikingly similar to those of Seeger and Maragos with WHV. Although the UUUC motif of WHV was originally considered an element of the origin of reverse transcription, subsequent data indicate that it represents the core of the minus-strand DNA transfer acceptor (8, 10). The WHV UUUC tetranucleotide is analogous to the AAUUAC hexanucleotide of DHBV in being an essential but insufficient element of the transfer acceptor whose function is position dependent. The UUUC sequence contains at its 3′ end a 3-nt homology (UUC) to the bulge of the SL (presumably representing the nascent minus-strand DNA [8, 10]), as the DHBV hexanucleotide contains the UUAC homology to the nascent minus-strand DNA. The sequences surrounding DR2 in WHV are not usually recognized in minus-strand DNA transfer and extension but, as with DHBV, can be utilized under artificial conditions (6). Thus, the general features of the minus-strand DNA initiation and transfer reactions identified in DHBV appear to be conserved among all hepadnaviruses.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health to D.G. and by institutional funds from St. Louis University granted to J.E.T.

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. **Ganem, D., J. R. Pollack, and J. E. Tavis.** 1994. Hepatitis B virus reverse transcriptase and its many roles in hepadnaviral genomic replication. Infect. Agents Dis. **3:**85–93.
3. **Hirsch, R. C., J. E. Lavine, L. J. Chang, H. E. Varmus, and D. Ganem.** 1990. Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription. Nature (London) **344:** 552–555.

3a. **Loeb, D. D., and D. Ganem.** Unpublished data.

4. **Pollack, J. R., and D. Ganem.** 1994. Site-specific RNA binding by a hepatitis B virus reverse transcriptase initiates two distinct reactions: RNA packaging and DNA synthesis. J. Virol. **68:**5579–5587.
5. **Seeger, C., and J. Maragos.** 1990. Identification and characterization of the woodchuck hepatitis virus origin of DNA replication. J. Virol. **64:**16–23.
6. **Seeger, C., and J. Maragos.** 1991. Identification of a signal necessary for initiation of reverse transcription of the hepadnavirus genome. J. Virol. **65:**5190–5195.
7. **Tavis, J. E., and D. Ganem.** 1993. Expression of functional hepatitis B virus polymerase in yeast reveals it to be the sole viral protein required for correct initiation of reverse transcription. Proc. Natl. Acad. Sci. USA **90:**4107–4111.
8. **Tavis, J. E., S. Perri, and D. Ganem.** 1994. Hepadnavirus reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer. J. Virol. **68:**3536–3543.
9. **Wang, G.-H., and C. Seeger.** 1992. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. Cell **71:**663–670.
10. **Wang, G.-H., and C. Seeger.** 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J. Virol. **67:**6507–6512.