JOURNAL OF VIROLOGY, May 1991, p. 2155–2163
0022-538X/91/052155-09$02.00/0

Vol. 65, No. 5

Evidence that Less-than-Full-Length *pol* Gene Products Are Functional in Hepadnavirus DNA Synthesis

TSUNG-TEH WU, LYNN D. CONDREAY,† LAURA COATES, CAROL ALDRICH, AND WILLIAM MASON*
Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111

Received 11 September 1990/Accepted 8 January 1991

Duck hepatitis B virus mutants containing frameshift or stop codon mutations in a portion of the viral *pol* gene separating the terminal protein and reverse transcriptase domains had a leaky phenotype and, depending on the location and type of mutation, synthesized up to 10% as much viral DNA as did the wild type. This region of the *pol* gene had previously been reported to be refractory to missense mutations; in fact, the leakiness of most of our mutants appeared attributable to translational suppression, which would also be expected to introduce amino acid changes. However, at least one mutant (pH1093+2), which was ca. 10% as active as the wild type, appeared to use a novel pathway to express the viral *pol* gene. Our analyses indicated that pH1093+2 synthesized the viral reverse transcriptase as a fusion protein with the amino-terminal portion of the pre-S envelope protein. Thus, in this case, the products of the terminal-protein and reverse transcriptase domains of the *pol* gene would function as separate protein species, though perhaps noncovalently joined in a dimeric structure during assembly of DNA replication complexes. Evidence was also obtained that was consistent with the idea that the wild-type *pol* gene may, at least in certain instances, be expressed as functional, subgenic polypeptides.

The *pol* gene of the hepadnaviruses appears to encode at least three functional units with a combined molecular size on the order of 90 kDa. They are NH_2-terminal protein-reverse transcriptase-RNase H-COOH (Fig. 1) (1, 31). The terminal-protein region is presumed to act as the primer of reverse transcription (12, 26), and there is suggestive evidence that the terminal protein domain is either produced as a separate translational product or proteolytically cleaved from downstream domains prior to or during the priming reaction (4). Some DNA polymerase activity gel studies on disrupted viral particles have identified both ca. 90- to 110-kDa and ca. 60- to 70-kDa activities, with most of the activity at 60 to 70 kDa (3, 27a). The latter group has also detected reverse transcriptase activity at ca. 40 to 50 kDa. Evidence that the 60- to 70-kDa species is a viral gene product spanning the reverse transcriptase and RNase H domains has been obtained by Western immunoblot analyses (2, 23). It is not known, however, which if any of the proteins identified on activity gels is actually functional in intracellular viral DNA synthesis.

Little is known about the mechanism by which *pol* open reading frame (ORF) products are produced. It is believed that one or more of the pregenomic RNAs of hepadnaviruses, which begin ca. 500 nucleotides (nt) upstream of the *pol* ORF, serve as messages for translation of at least the ca. 90-kDa *pol* product, and there is good genetic evidence that translation of an essential *pol* gene product initiates at the first AUG of the *pol* ORF (8, 32, 33). Whether splicing, proteolytic cleavage, or additional, downstream translational initiation events are involved in producing *pol* gene products that are less than 90 kDa in length is unclear.

Bartenschlager and Schaller (1) and Li et al. (21) have reported the existence of a spacer region between the terminal protein and reverse transcriptase domains of at least 153 bp (nt 1093 to 1245) that is refractory to missense mutations and in-frame deletions. One possible explanation for this result is that the spacer lies downstream of a domain essential to terminal-protein activity and either within an intron which is removed to produce the mRNA for reverse transcriptase or upstream of a second translation initiation site in *pol*. Splicing events that may be compatible with the former possibility have been reported, though their functional significance is unclear (9, 36, 39). To explore these issues further, we examined the effect of creating single frameshift or stop codon mutations at several different sites within or just downstream of this spacer region. All of these mutations reduced the capacity for viral DNA synthesis 10- to 1,000-fold or more, with the amount of residual synthesis being a characteristic of each mutant. Analysis of viruses containing multiple mutations in the spacer region suggested that the residual activity of individual mutants was due neither to excision of the mutations by splicing out of the spacer region nor to downstream initiation; therefore, leakiness of mutant viruses with stop codon or frameshift mutations in the spacer region was probably due to translational stop codon and frameshift suppression. Surprisingly, one frameshift mutant synthesized viral DNA at ca. 10% of wild-type levels. This mutant appeared to synthesize a functional reverse transcriptase as part of a pre-S/*pol* fusion protein. This result is consistent with the idea that a less-than-full-length *pol* gene product may be functional in viral DNA synthesis.

MATERIALS AND METHODS

Plasmids. Infectious duck hepatitis B virus (DHBV) is produced following transfection of human liver tumor cells with pSP.DHBV5.2Galx2 (30), which contains a tandem *Eco*RI dimer insert of DHBV into the polylinker of pSP65. With this construction, all viral transcription in the transfected cells is under the control of viral sequences. To simplify the analysis of viral mutations, nontranscribed DHBV sequences characteristic of dimer constructs were

* Corresponding author.
† Present address: Burroughs Wellcome Co., Research Triangle Park, NC 27709.

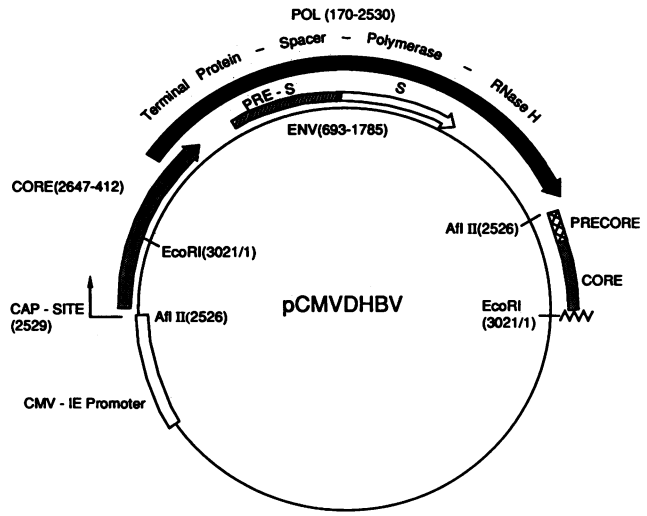

FIG. 1. Schematic representation of the DHBV genome in plasmid pCMVDHBV. In this plasmid, transcription of DHBV pregenomic RNA is under the control of the CMV IE promoter, but the subgenomic mRNAs for envelope protein are still under the control of DHBV promoters. The arrows represent three different ORFs, including core, polymerase (*pol*), and envelope (*env*) protein. The precore-core gene encodes two proteins, but the larger (precore), which is not essential for virus replication (7, 34), cannot be synthesized by this construct. The *env* gene encodes a nested set of two (or perhaps three) proteins, including the larger pre-S protein spanning the entire *env* gene and the S protein, localized to the 3′ terminus of this ORF. The nucleotide numbering system is according to Mandart et al. (24). Two restriction endonuclease sites are also indicated in which *Eco*RI is at a unique site within the DHBV genome and *Afl*II is at a site located 3 bp upstream of the DHBV pregenomic RNA cap site.

eliminated and synthesis of pregenomic RNA was placed under control of a heterologous promoter, in particular the cytomegalovirus (CMV) immediate-early (IE) promoter as used by Seeger and Maragos (35) with woodchuck hepatitis virus. The DHBV sequence was removed from pSP.DHBV5.1Galx2 (which is identical to pSP.DHBV5.2Galx2 except that the insert is in the opposite orientation) by cleavage at a *Sal*I site in the polylinker followed by partial digestion with *Afl*II. This fragment was then cloned into a derivative of plasmid pBC12/CMV/IL-2, which was described by Cullen (11). To create this derivative, a *Sac*I (position 663)-to-*Sma*I (position 2142) fragment was deleted from pBC12/CMV/IL-2 and a polylinker was inserted in its place to create plasmid pTTW-1 (10). Since the *Sac*I site of pBC12/CMV/IL-2 is 18 bases upstream of the CMV IE promoter cap site, the polylinker was designed to contain an *Afl*II site placed so that insertion of the DHBV fragment between this site and a downstream *Sal*I site in the pTTW-1 polylinker would put the DHBV pregenome cap site (DHBV position 2529 [22] according to the numbering system of Mandart et al. [24]) in the position originally occupied by the CMV IE cap site (Fig. 1). It should be noted that this construct, pCMVDHBV, still contains two DHBV promoters adjacent to intact ORFs, one just upstream of and the other within the pre-S region of the viral envelope gene. However, this construct lacks the capacity to synthesize precore protein because translation of this protein initiates upstream of the pregenome cap site. Previous reports have demonstrated that precore protein is nonessential to virus replication (7, 34). We have not found any significant differences in virus production from selected cells transfected with this construct or with the dimer construct (10).

In DHBV, the core gene is, by convention, in the first reading frame (RF), the *pol* gene is in the second RF, and the *env* gene is in the third RF. Frameshift mutations (+2) were introduced in pCMVDHBV at a number of sites within or flanking the putative spacer regions of the DHBV *pol* ORF by restriction endonuclease cleavage and repair of the four-base overhanging ends with the Klenow fragment of DNA polymerase I: at the *Hin*dIII sites, DHBV positions 1093 (pH1093+2) and 1334 (pH1334+2); at the *Xho*I site, 1212 (pX1212+2); at the *Xba*I site, 1358 (pXb1358+2); and at the *Bam*HI site, 1658 (pB1658+2). Frameshift mutations (+1) were also produced at all but the *Bam*HI site by adding a 10-base *Sal*I linker [d(CGGTCGACCG); New England BioLabs, Inc.] at the respective endonuclease cleavage sites following repair with the Klenow fragment of polymerase I; plasmid designations were given as described above except that the +2 was changed to a +1. In addition to these frameshift mutants, we constructed one plasmid, pH1093SP, in which a linker [d(CTAGTCTAGACTAG); New England BioLabs] containing a stop codon in all three ORFs was inserted, as described above, at the *Hin*dIII site at position 1093 and another plasmid, pH1093SP3, in which the linker [d(CTAGTCTAGACTAGTGATAAGCT)] placed three sequential stop codons in the *pol* ORF and one stop codon in each of the other two ORFs. Plasmid pHSPX is doubly mutated, containing the stop codons of pH1093SP and the frameshift of pX1212+2. The anticipated effects of the different mutations on the *pol* and *env* ORFs are summarized in Fig. 2.

Three derivatives of pH1093+2 were designed to express pre-S/*pol* fusion proteins in the absence of viral core antigen or terminal-protein expression. The cloning site in each case was a *Sma*I sequence at the 3′ end of the polylinker of pTTW-1, 26 bp downstream of the CMV IE promoter cap site. Plasmid pFUS-1 contains an *Eco*RV fragment (positions 721 to 2652), of the DHBV sequence in pH1093+2, inserted into the *Sma*I site of pTTW-1 in the 5′-to-3′ orientation. Plasmid pFUS-2 contains an *Ssp*I-*Nsi*I fragment of DHBV from pH1093+2 (positions 678 to 2845), blunt ended by exonuclease digestion with the Klenow fragment of DNA polymerase I, also inserted into the *Sma*I site of pTTW-1. It should be noted that pFUS-1 lacks the first AUG of the *env* ORF, which is apparently not required for pre-S translation (5), and also lacks the DHBV polyadenylation signal. However, a polyadenylation site exists downstream of the insert within a eukaryotic sequence retained from pBC12/CMV/IL-2. By contrast, pFUS-2 includes the entire *env* ORF and contains the viral polyadenylation signal. Plasmid pFUS-3 contains an *Eco*RV-to-*Nsi*I fragment from pH1093+2 (positions 721 to 2845). Thus, this plasmid lacks the first AUG of the *env* ORF but retains the DHBV polyadenylation site. Plasmids similar to pFUS-2 were also constructed by using *Ssp*I-*Nsi*I inserts from pCMVDHBV, pH1093+1, and 1S (env-) (38) to create pPreS, pPreS+1, and pSM, respectively.

Plasmid DNAs for transfection of eukaryotic cells were isolated by an alkaline lysis procedure (25) and then subjected to isopycnic centrifugation in a CsCl gradient. The presence of the desired mutations in the plasmid preparations was confirmed by DNA sequencing.

Cell lines. The human hepatoma cell line Huh-7 (27, 30) and the chicken hepatoma cell line LMH (19) were maintained in nutrient medium at 37°C, changed daily, as de-

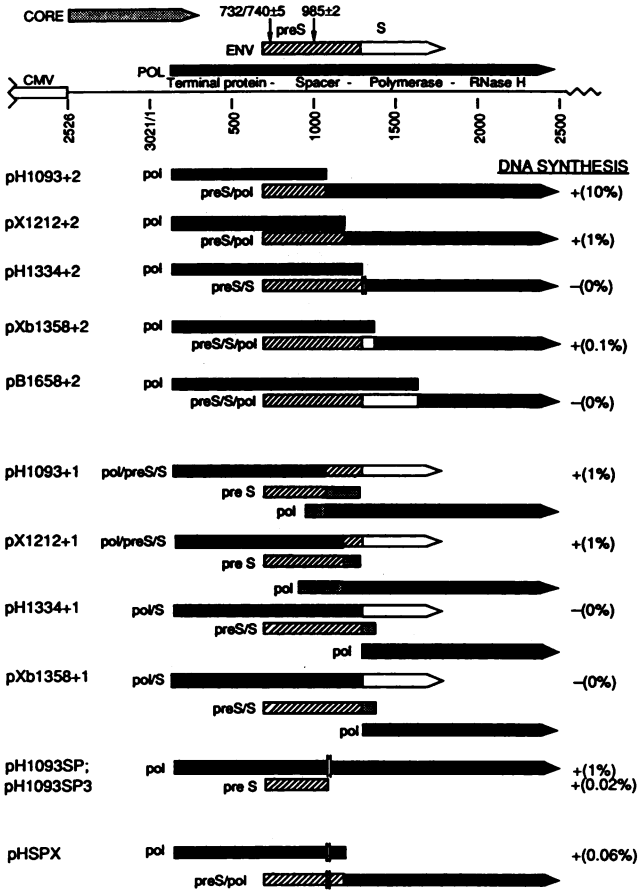

FIG. 2. Predicted effects of different frameshift and stop codon insertion mutations on the *pol* and *env* ORFs of DHBV. The strategies used to construct the mutant plasmids and the plasmid designations are described in Materials and Methods. Relative to the core ORF, the *pol* ORF is +1 and the *env* ORF is +2. In the case of the +2 frameshift mutants, the *pol* ORF was shifted into the noncoding frame (core ORF) and terminated, while the *env* ORF was shifted into the *pol* ORF, creating a *env-pol* fusion gene. In the case of the +1 frameshift mutants, the *pol* ORF was shifted into *env* ORF, creating a *pol-env* fusion gene, the *env* ORF was shifted into the noncoding frame, and the noncoding frame was shifted into the *pol* ORF. In no instance did the latter shift introduce an AUG from the noncoding frame onto the 5′ end of the carboxy-terminal portion of the *pol* ORF. Plasmid pHSPX was a double mutant containing the stop codon insertion of pH1093SP and the +2 frameshift mutation of pX1212+2. An estimate of the amount of viral DNA synthesis for each mutant was made by transient transfection of the plasmid into LMH cells and comparison with the amount of viral DNA produced following transfection with the wild-type plasmid (pCMVDHBV).

scribed by Pugh et al. (30) except that for routine maintenance of the cell line, LMH growth medium contained 10% calf serum instead of fetal calf serum. Cells were trypsinized and passaged at a 1:3 dilution every 4 to 5 days. Regular transfer was especially important with the LMH cells, which tended to clump together and became difficult to trypsinize as the cultures became heavy. The LMH cells produce 10 to 20 times more virus than do Huh-7 or HepG2 liver tumor cells after transfection, either with a plasmid containing a tandem dimer insert of DHBV or with pCMVDHBV (10).

For transfection, 1.65×10^6 Huh-7 or 3.2×10^6 LMH cells per 60-mm-diameter tissue culture dish were grown for ca. 24 h in growth medium containing 10% fetal calf serum (10, 30). Fresh medium was placed on the cells, and transfection with 10 μg of a single plasmid, or 5 μg of each in cotransfections, was done by the calcium phosphate precipitation method of Graham and van der Eb (13). The precipitate was left on the cells for 14 to 16 h. The medium was then changed, and unless indicated otherwise the cells were harvested after another 4 days, with medium changes at daily intervals. The procedure for isolation of intracellular viral cores (which are the viral DNA replication complexes [37]) from transfected cultures was as described by Pugh et al. (30). Where indicated, culture medium from the transfected cells was removed, clarified by centrifugation to remove floating cells, and stored at −80°C for subsequent infectivity assays.

Hepatocyte cultures. Primary hepatocyte cultures were prepared and maintained in serum-free medium as described previously (28, 40). The cultures were infected the next day with DHBV and, after 1 day for virus adsorption, were maintained in medium containing 100 μg of suramin per ml to prevent further virus uptake (28). At 8 days postinfection, the cultures were rinsed with phosphate-buffered saline and stored at −80°C for subsequent DNA extraction.

Analysis of DHBV DNA. DHBV DNA was extracted from viral cores and from primary hepatocyte cultures (28, 30) and subjected to electrophoresis in 1.5% agarose slab gels containing 0.04 M Tris-HCl, (pH 7.2), 0.02 M sodium acetate, and 1 mM EDTA. Nucleic acids were transferred to nitrocellulose sheets (41), and DHBV DNA was detected by hybridization with ^{32}P-labeled, cloned DHBV DNA representing the entire viral genome. The nitrocellulose sheets were then subjected to autoradiography at −70°C, using DuPont Cronex Lightning-Plus intensifying screens and Kodak XAR-5 X-ray film. Signal intensities were quantitated by densitometric scanning of X-ray films or by direct analysis of the nitrocellulose sheets with an AMBIS Radioanalytic imaging system.

Protein blotting. Cytoplasmic extracts of transfected cell cultures were assayed immunologically for DHBV *env* gene products. Monolayers were lysed in 1.0 ml of 0.15 M NaCl–0.01 M Tris-HCl (pH 7.5)–0.5% (vol/vol) Nonidet P-40–0.2 trypsin inhibition units of aprotinin (Sigma) per ml. After 15 min at room temperature, the lysed cells were harvested with a rubber policeman and nuclei were removed by centrifugation for 5 min at 12,000 × *g* and 4°C. The protein in the supernatants was then precipitated by addition of 1/10 volume of 1 M Tris-HCl (pH 8) and 5 volumes of ethanol and storage for at least 1 day at −20°C. Precipitates were collected by centrifugation for 15 min at 12,000 × *g* and 4°C, dried, and dissolved in electrophoresis sample buffer (20). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of samples (1/16) of each extract, electrophoretic transfer to nitrocellulose, and detection of viral antigens with a reactive rabbit serum were done according to a published procedure (15). To detect the 36-kDa pre-S protein of DHBV, we used rabbit antisera raised to purified virions of DHBV. These regents react primarily with determinants encoded within the pre-S domain of the viral *env* gene (29). Two anti-DHBV sera were used; one (no. 365) has been previously described (14, 29), and the other was a gift from John Newbold (University of North Carolina, Chapel Hill).

RESULTS

Effects of frameshift mutations within the *pol* gene on viral DNA synthesis. Horwich et al. (17) have found that DHBV

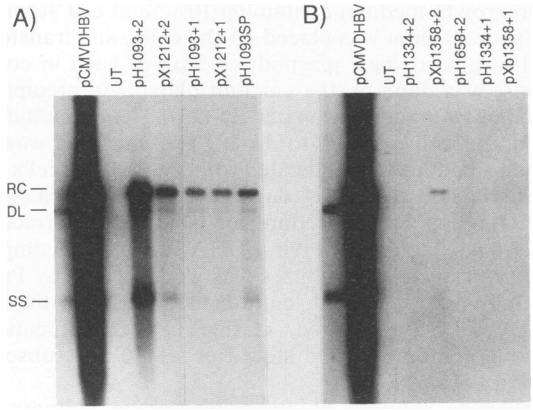

FIG. 3. Evidence that frameshift and stop codon insertion mutations in the spacer region of the *pol* ORF do not completely block DHBV DNA synthesis. Viral DNA present in core particles located in the cytoplasm of LMH cells transiently transfected with wild-type and mutant plasmids was extracted and subjected to Southern blot analysis. Each lane received one-quarter of the DNA extracted from a single 60-mm petri dish. (A) Southern blot analysis of the replicative forms of DHBV DNA produced following transfection with mutant plasmids with a leaky phenotype despite the presence of frameshift and stop codon insertion mutations in the spacer region of *pol*. The replicative intermediate patterns of the mutants, including relaxed circular (RC), double-stranded linear (DL), and single-stranded linear (SS) DNA, was similar to the wild-type pattern. (B) Similar Southern blot analysis of frameshift mutants with little (pXb1358+2) or no viral DNA synthesis. UT, Untransfected cells.

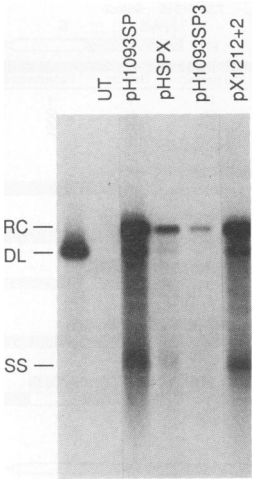

FIG. 4. Effects of multiple mutations in the *pol* spacer region on DHBV DNA synthesis. LMH cells were transfected with two plasmids containing a single mutation (pH1093SP and pX1212+2), a plasmid with double mutations (pHSPX), and a mutant plasmid with three sequential stop codon insertions in *pol* (pH1093SP3). Cells were harvested 5 days posttransfection, and cytoplasmic viral core DNAs were extracted for Southern blot analysis. For definitions of RC, DL, and SS, see the legend to Fig. 3.

DNA synthesis in transfected Huh-7 cells is completely blocked by the introduction of a frameshift mutation at either nt 718 or nt 1290 within the viral *pol* gene. Similar results were obtained by us when LMH cells were transfected with these mutants, which were generously provided by Jesse Summers (data not shown). We found, however, that viral DNA synthesis still occurred following transfection of LMH cells with DHBV DNAs containing a frameshift mutation at position 1093, 1212, or 1358, all of which are sites that are in or just downstream of the *pol* gene spacer region that was defined by Bartenschlager and Schaller (1). The patterns of intermediates in viral DNA synthesis obtained following transfection with DNAs mutated at position 1093, 1212, or 1358 were similar to those observed with the wild-type (Fig. 3), although depressed ca. 10-fold in the case of a +2 frameshift at 1093, ca. 30- to 100-fold with a +2 frameshift at 1212, ca. 100-fold in the cases of a +1 frameshift at 1093 and at 1212, and ca. 1,000-fold in the case of a +2 frameshift at 1358 (summarized in Fig. 2). In contrast, with mutants containing a frameshift at position 1334 or 1658, no (<0.02%) viral DNA synthesis was detected. Finally, insertion of a single stop codon into the *pol* ORF at position 1093, like the frameshift mutations, reduced viral DNA synthesis only ca. 100-fold. A preliminary analysis in Huh-7 cells yielded similar results: pH1093+2 and pX1212+2, but not pH1334+2, produced detectable yields of replicative forms of DHBV DNA (data not shown).

One possible explanation for these results is that the spacer region containing the leaky mutations may have been removed from the *pol* mRNA(s) by splicing. If this were the case, then two spacer region mutations placed in tandem should not reduce viral DNA synthesis to any greater extent than either mutation alone. The double mutant pHSPX, combining the stop codon insertion of pH1093SP and the frameshift of pX1212+2, was, in fact, ca. 20-fold less active than either parent (Fig. 4). Likewise, pH1093SP3, in which a linker inserted at position 1093 introduced three sequential stop codons into the *pol* ORF (as well as a single stop codon in each of the other two ORFs), was ca. 50-fold less active than pH1093SP, which has only a single stop codon at this position in the *pol* ORF (Fig. 4). Taken together, these results suggested that *pol* gene expression by spacer region mutants was not due to a bypass of the mutations by RNA splicing or internal translation initiation downstream of the spacer region.

Although our data were not consistent with the possibility that sequences in the spacer region were spliced out of *pol* mRNA(s), several other explanations of the results remained. Alternative explanations included exact deletion of mutant sequences to restore the parental (wild-type) genotype, larger in-frame deletions which maintained the *pol* ORF, phenotypic suppression during *pol* mRNA transcription or translation, or even low-level contamination of mutant plasmid DNAs with wild-type DHBV. Also, for the frameshift mutations, second-site frameshift mutations could act to compensate for the first mutation and maintain upstream and downstream portions of *pol* in frame, while for the stop codon mutation in pH1093SP, a single base change could remove the stop codon from the *pol* reading frame.

To determine whether our results could be due to contamination of mutant plasmids with wild-type sequences, we examined the consequences of cotransfection of LMH cells with 5 μg of pH1334+2 DNA (Fig. 2) and various amounts (5, 0.5, and 0.05 μg) of the wild type, pCMVDHBV (Fig. 1). Since pH1334+2 is totally defective in viral DNA synthesis, mutant pregenomes produced by this vector will be reverse transcribed only in cells that are also transfected by pCMVDHBV. If transcomplementation of *pol* gene activity was weak in LMH cells, with preferential reverse transcription of the pregenome expressing the *pol* gene product(s), as previously reported in Huh-7 cells (16), then the wild-type se-

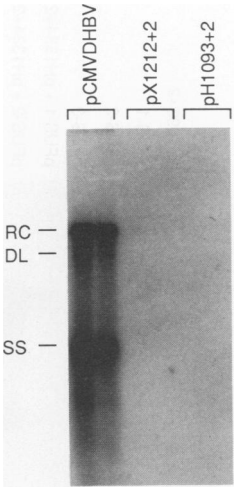

FIG. 5. Evidence that frameshift mutants that synthesized viral DNA did not produce infectious virus. LMH cells were transfected with either the wild-type plasmid, pCMVDHBV, or the mutant plasmids pH1093+2 and pX1212+2 (see Fig. 2). Tissue culture fluids from these transfected cells were collected between days 4 and 5 posttransfection and clarified by centrifugation, and 1-ml aliquots were used to infect primary hepatocyte cultures. After virus adsorption for 24 h, culture medium was removed and fresh medium with suramin (100 μg/ml) was used to prevent further virus infections. Hepatocytes were harvested at 8 days after infection. Total nucleic acids were extracted, and one-quarter of each sample from duplicate plates was assayed by Southern blot analysis. For definitions of RC, DL, and SS, see the legend to Fig. 3.

quence should always be prominent among the reverse transcription products, irrespective of the ratio of mutant to wild-type plasmid in the cotransfections. On the other hand, if transcomplementation of *pol* defects was highly efficient in LMH cells, then a large excess of mutant to wild-type plasmid in a cotransfection might result in replicative forms of viral DNA that were almost exclusively of the mutant genotype. In fact, irrespective of the ratio and even with a 1,000-fold excess of mutant to wild-type plasmid DNA, ca. 25 to 50% of the replicative forms had the mutant genotype and the remainder were wild type (see Fig. 8A; unpublished observations). These results are again consistent with a *cis* preference for reverse transcription, as found in Huh-7 cells (16). The contribution of wild-type DHBV sequences (e.g., as contaminants or revertants to wild type) to *pol* mutant DNA synthesis should therefore be readily discernible.

On the basis of this observation, we next asked whether our results with the most leaky mutants might be due to reversion to wild type or a precise deletion of the mutated region. The +1 and +2 frameshift and stop codon mutations introduced at positions 1093 and 1212 produced unique restriction endonuclease cleavage sites; moreover, the 3-kbp, relaxed circular viral DNAs in cells transfected with any of these seven mutants were completely digested by the respective endonuclease (data not shown), indicating that the mutated sites were retained in the replicating DNA. This result was incompatible with the possibility that the mutants replicated because the primary mutation had been deleted. The possibility of precise deletion of the mutations was also ruled out by the observation that pH1093+2 and pX1212+2 produced, in transfected LMH cells, 1 to 10% as much replicative viral DNA as did the wild type, but no infectious virus was detectable in the culture fluids (Fig. 5). (This assay was sufficiently sensitive to detect as little as 0.5% of the level of virus produced by the wild-type control.) Thus, the activity of these mutants may reflect either phenotypic suppression of the mutations (e.g., by ribosomal frameshifting or stop codon readthrough) or, for the frameshift mutants, a high frequency of second-site mutation to restore the *pol* reading frame (Fig. 2). We have not yet directly addressed these possibilities; however, we have considered an explanation for the expression of *pol* activity that would uniquely apply to the +2 frameshift mutants.

Evidence that pre-S/*pol* fusion proteins created by +2 frameshift mutations at certain sites in the spacer region of *pol* may be active in DHBV DNA synthesis. Consideration of the data in Fig. 2 revealed that DHBV clones with +2 frameshift mutations at positions 1093, 1212, and 1358 were somewhat more active than clones with +1 frameshifts or stop codons at the same positions. This was especially notable for pH1093+2, which was ca. 3 to 100 times as active as the other clones with mutations at the same position, pH1093+1, pH1093SP, and pH1093SP3. A +2 frameshift mutation at position 1093, as well as at positions 1212 and 1358, would cause termination of translation of the upstream portion of the *pol* gene in the spacer region between the terminal-protein and reverse transcriptase domains but would permit expression of the reverse transcriptase and RNase H domains as a ca. 65-kDa fusion protein with the pre-S/S region of the *env* gene. Two types of experiment suggested that synthesis of these fusion proteins might serve as an alternate mechanism of expression of the reverse transcriptase and RNase H domains to produce functional DNA synthesis complexes, especially in the case of pH1093+2. In the first experiment, we looked for the predicted pre-S/*pol* fusion proteins by immunoblotting with antiviral rabbit sera strongly reactive with the 36-kDa pre-S protein product of the *env* gene. This approach revealed the 36-kDa pre-S protein in extracts of LMH cells transfected with wild-type DHBV (Fig. 6). The 36-kDa protein was not detected in mutant-transfected cultures. Instead, the cells transfected with the +2 frameshift mutants (except pH1334+2) produced new proteins with sizes of ca. 60 to 72 kDa (Fig. 2). This range of electrophoretic mobilities cannot be attributed to differences in predicted lengths of the various fusion proteins (predicted molecular size difference, less than 1 kDa) but may be attributable to different levels of glycosylation, phosphorylation, posttranslational proteolysis, or even in amino acid sequence. In contrast, the cells transfected with the +1 frameshift mutants produced 3′ truncated *env* gene products of ca. 16 to 25 kDa. The presence of two rather than one band in this size range may be indicative of translational initiation from more than one AUG in pre-S; however, the exact sizes of the bands varied unpredictably, suggesting that posttranslational modifications of these aberrant viral gene products might also have occurred.

Inasmuch as the predicted pre-S/*pol* fusion proteins appeared to be synthesized, a second experimental approach was used to determine whether one of these proteins might actually function in viral DNA synthesis. LMH cells were cotransfected with the DNA-negative mutant pH1334+2 (Fig. 2) and either pFUS-1, pFUS-2, or pFUS-3, expression vectors in which restriction fragments encompassing the pre-S/*pol* fusion gene of pH1093+2 were placed downstream of the CMV IE promoter. Transcomplementation apparently occurred in the cotransfection of pH1334+2 with pFUS-2 and pFUS-3 and perhaps with pFUS-1, although in this case with a much lower efficiency (Fig. 7). (We have never

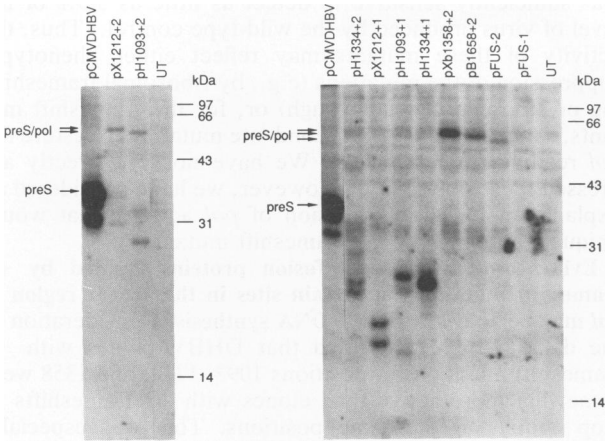

FIG. 6. Detection of the predicted pre-S/*pol* fusion proteins in LMH cells transfected with DHBV mutants possessing +2 frameshift mutations in the viral *pol* gene. LMH cells were transfected with different +2 or +1 frameshift mutation plasmids. Cells were harvested 3 days posttransfection, and proteins were collected and assayed by Western blot analysis. The pre-S protein (ca. 36 kDa) and pre-S/*pol* fusion proteins were visualized by using rabbit antisera raised to purified viruses.

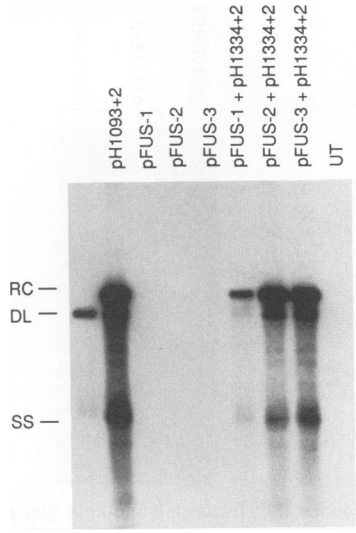

FIG. 7. Transcomplementation by the pre-S/*pol* fusion protein of a DNA-negative mutant. LMH cells were transfected with 10 μg of pH1093+2, pFUS-1, pFUS-2, or pFUS-3 or were cotransfected with 5 μg of pH1334+2 and 5 μg of either pFUS-1, pFUS-2, or pFUS-3. Cells were harvested 5 days posttransfection. Cytoplasmic viral core DNAs were extracted and assayed by Southern blot analysis. For definitions of RC, DL, and SS, see the legend to Fig. 3.

detected any evidence of viral DNA synthesis in cells transfected with pH1334+2 alone.) The ability of pFUS-2 and pFUS-3 to complement pH1334+2 correlated with the observation that these constructs, but not pFUS-1, produced detectable levels of the presumptive pre-S/*pol* fusion protein (Fig. 6) (pFUS-3 data not shown). The failure of pFUS-1 to produce detectable levels of pre-S/*pol* fusion protein suggested that the downstream interleukin-2 polyadenylation signal (11) did not compensate for the absence of the viral polyadenylation signal in this construct. The relative lack of viral DNA synthesis following cotransfection with pFUS-1 also suggested that the activity exhibited by pFUS-2 and pFUS-3 was not due to recombination between the cotransfected plasmids to produce wild-type DHBV genomes. Consistent with this interpretation, we found that a construct (pPreS) similar to pFUS-2 but containing an insert of wild-type DHBV sequences (i.e., an *env* expression vector) was no better than pFUS-1 in promoting viral DNA synthesis following cotransfection with pH1334+2 (data not shown).

Though the interpretation of the results in Fig. 7 seemed straightforward, we decided to test a number of additional plasmids to further evaluate the potential of recombination as an explanation for the apparent complementation. Three plasmids, pPreS, pPreS+1, and pSM, were therefore compared with pFUS-2 (Fig. 8C) for the ability to complement pH1334+2. pPreS+1, like pPreS, was homologous to pFUS-2 but contained an insert of DHBV sequences from pH1093+1; pSM was also homologous to pFUS-2 but contained *env* gene mutations (T → A) at nt 1327, 1345, and 1348, which abrogate *env* gene expression without, in the context of the complete viral genome, affecting *pol* activity (38). Complementation experiments (Fig. 8B) revealed that pPreS complemented pH1334+2 ca. 7- to 8-fold less effectively than did pFUS-2 and that pPreS+1 complemented at least 10-fold less efficiently than did pFUS-2. However, pSM appeared to work just as well as pFUS-2 in the complementation assay. Restriction endonuclease digestion of relaxed circular viral DNA from the complementations with *Hin*dIII, which cleaves at positions 14, 1093, and 1334 in the wild-type genome but only at the first two sites in pH1334+2, revealed the presence of a major component of apparently wild-type relaxed circles following cotransfection of pH1334+2 with pPreS+1 or pSM but not with pFUS-2 or pPreS. This result suggested that cotransfection of pSM or pPreS+1 with pH1334+2 resulted in functional *pol* gene products following plasmid recombination to generate wild-type DHBV genomes. In contrast, pFUS-2 by itself must have produced a functional pre-S/*pol* fusion protein which acted in concert with a 5′-terminal *pol* gene product synthesized by pH1334+2 to replicate viral DNA; surprisingly, our results suggested that pPreS may also complement pH1334+2, possibly by production of a truncated *pol* gene product through de novo translation initiation from a methionine codon at nt 905.

A peculiar aspect of the experiment shown in Fig. 8B was the larger amount of viral DNA synthesis after cotransfection with pSM than after cotransfection with pPreS or pPreS+1. A single recombination event between pH1334+2 and pSM to generate a DHBV genome with a wild-type *pol* gene would not reconstitute the *env* gene, whereas a single recombination between pH1334+2 and any of the remaining vectors would restore both *pol* and *env*. In the absence of pre-S protein, viral covalently closed circular DNA, the transcriptional template in natural infections, undergoes 20 to 30 times more amplification in transfected LMH cells (44). This could contribute to a differential amplification of recombinant pregenomes and ultimately replicative forms of DHBV DNA in cells cotransfected with pSM and pH1334+2.

DISCUSSION

Our results suggested that DHBV mutants with frameshift and stop codon mutations in the spacer region of the *pol* gene, which continued to express residual reverse tran-

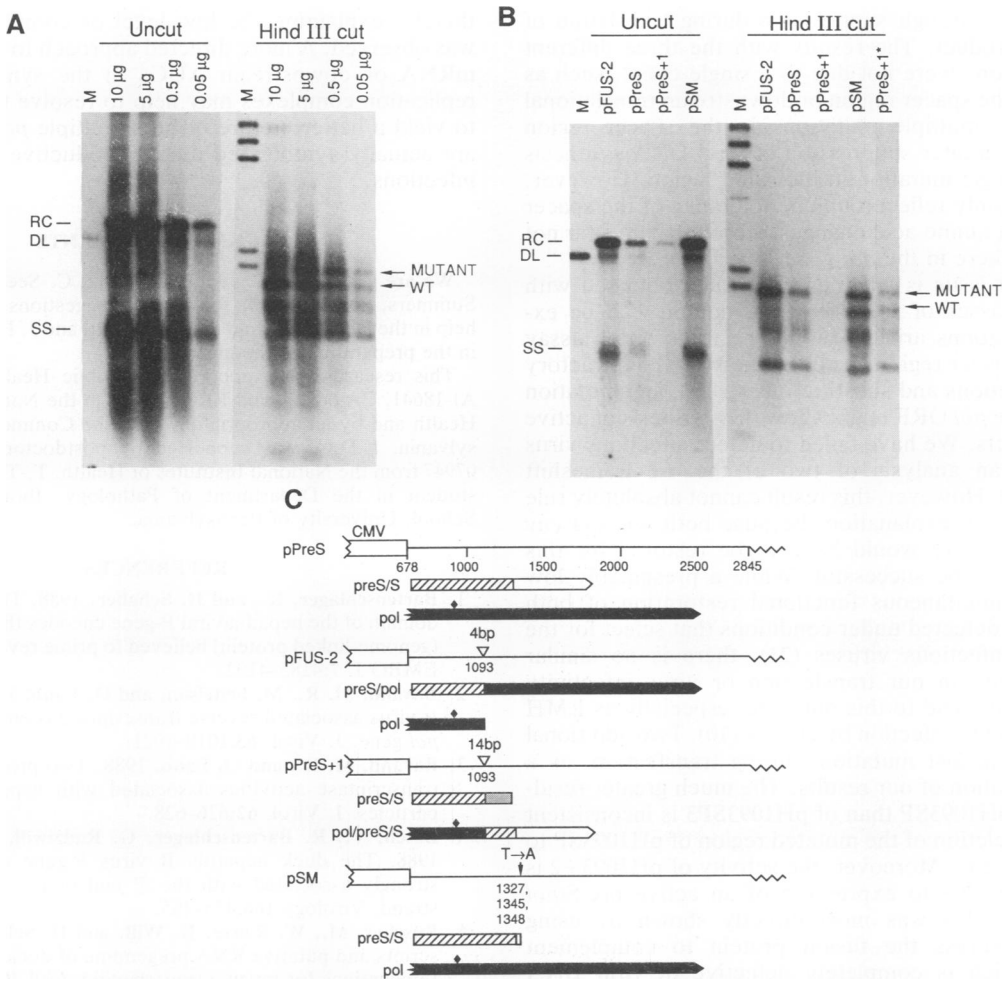

FIG. 8. Evidence that transcomplementation between a pre-S/*pol* fusion protein expression vector and a DNA-negative mutant was not due to recombination between cotransfected plasmids. LMH cells were transfected with 10 μg of pCMVDHBV alone or were cotransfected with 5 μg of pH1334+2 plus 5, 0.5, or 0.05 μg of pCMVDHBV together with pTTW-1 to make a final total of 10 μg of plasmid DNA in each transfection (A) or with 5 μg of pH1334+2 and 5 μg of either pFUS-2, pPreS, pPreS+1, or pSM, respectively (B). Cells were harvested 5 days posttransfection. Cytoplasmic viral core DNAs were extracted and assayed by Southern blot analysis either with uncut samples or with DNAs that were first digested with the restriction endonuclease *Hin*dIII. The unique *Hin*dIII-digested band was labeled as mutant viral DNA (from pH1334+2) or as wild-type (WT) viral DNA (from pCMVDHBV or from recombinant plasmid). The predicted viral ORFs from different plasmids used in transcomplementation assays (B) are summarized in panel C. The *pol* ORF shown here was predicted from the whole *pol* ORF. There is an ATG (◆) in the *pol* ORF at nt 905. Lanes M, 3-kbp double-stranded linear, cloned DHBV DNA (left) or ^{32}P-labeled λ DNA cut with *Hin*dIII (right). For definitions of RC, DL, and SS, see the legend to Fig. 3. *, The exposure time of the autoradiogram in this lane is five times longer than for the other lanes.

scriptase activity, did so by different mechanisms. First, the capacity of mutants with +2 frameshifts to synthesize reverse transcriptase was probably due, at least in part, to the expression of a pre-S/*pol* fusion protein with a reverse transcriptase and RNase H that were active either as part of the fusion protein or after proteolytic processing to remove *env* sequences. Evidence for this conclusion was presented for pH1093+2 and was consistent with the immunoblot data (Fig. 6) for pX1212+2 and pXb1358+2. However, if proteolytic release of *env* sequences does occur, it may not reflect a normal cleavage site in *pol*, since the COOH-terminal *pol* fragment that would be released could not be larger than 55 kDa for pH1093+2, 51 kDa for pX1212+2, or 45 kDa for pXb1358+2. Second, the synthesis of reverse transcriptase following the introduction of a +1 frameshift in *pol* could not be due to the synthesis of a pre-S/*pol* fusion protein and, in the absence of an available AUG, is unlikely to be due to translational initiation in the noncoding frame upstream of the mutant genome (Fig. 2). For these mutants, ribosomal frameshifting would seem to be the most plausible explanation for the continued production of reverse transcriptase. Homopolymeric runs of six or more A's or T's upstream of sites of frameshifting have been associated with efficient (1 to 5%) −1 frameshifts in animal and yeast cells (18, 43); however, a similar association with +1 frameshifts was not found in yeast cells (42). Homopolymeric runs were not found upstream of the sites of the DHBV mutations, and it is possible that +1 frameshifts in animal cells employ other sequence motifs. Third, the failure of a stop codon insertion (pH1093SP) to totally block viral DNA synthesis was con-

sistent with readthrough suppression during translation of the *pol* gene product. The results with the three different types of mutations were not due to a single event, such as splicing out of the spacer region or downstream translational initiation, since multiple mutations in the spacer region caused a much greater suppression of viral DNA synthesis than did any single mutation in the same region. However, the results certainly reflect a unique tolerance of the spacer region to accept amino acid changes, a property thus far not observed elsewhere in the *pol* gene.

An important issue is whether the results obtained with the different classes of spacer region mutants can be explained by mutations arising during the transfection assay (6). Since the spacer region has been shown to be refractory to in-frame deletions and substitutions (1, 21), any mutation that restored the *pol* ORF might allow the synthesis of active *pol* gene products. We have failed to detect infectious virus production in an analysis of two of the +2 frameshift mutants (Fig. 4). However, this result cannot absolutely rule out mutation as an explanation, because both *pol* and *env* gene product activity would have to be restored for this infection assay to be successful. While a presumably low frequency of simultaneous functional restoration of both genes has been detected under conditions that select for the outgrowth of infectious viruses (21), there is no similar selective pressure in our transfection or virus infectivity assays that would lead to this outcome, especially as LMH cells are resistant to infection by DHBV (10). Two additional findings argue against mutation, during transfection, as a unifying explanation of our results. The much greater residual activity of pH1093SP than of pH1093SP3 is inconsistent with in-frame deletion of the mutated region of pH1093SP to restore the *pol* gene. Moreover, the activity of pH1093+2 is almost certainly due to expression of an active pre-S/*pol* fusion protein. This was most directly shown by using vectors that express the fusion protein to complement pH1334+2, which is completely defective in viral DNA synthesis. Complementation was associated with expression of the fusion protein (Fig. 6 to 8) and could not be explained simply by recombination between the vectors to produce a wild-type *pol* gene. In this assay, the terminal protein was apparently provided as a COOH-terminal-truncated product of the pH1334+2 *pol* gene (Fig. 2). Thus, at least in this situation, it appears that the terminal protein and reverse transcriptase-RNase H domains can function as separate protein species to direct reverse transcription of the viral pregenome. If so, each might be predicted to possess a recognition sequence for packaging into viral cores or would have to be incorporated as a heterodimer.

A final question is whether the phenomena that we have described provide any insights into the normal pathway of *pol* gene expression. At this point, the answer is unclear. The results in Fig. 8B (pPres complementation of pH1334+2) suggest that a pre-S mRNA could direct the synthesis of a functional 61-kDa *pol* protein, beginning with an AUG at nt 905 and spanning the spacer, reverse transcriptase, and RNase H domains of the *pol* gene. While the size and composition of this protein species are intriguing, the complementation assay that we have used (Fig. 8) to detect the possible synthesis of such a protein may not be optimal. In particular, following cotransfection of pH1334+2 and pPreS, the terminal protein domain of *pol* must be expressed as an aberrant, truncated gene product from plasmid pH1334+2. It is conceivable that this truncated *pol* product has very low functional activity or even interferes with the assembly of DHBV-DNA replication complexes, thereby explaining the low level of complementation that was observed. A more directed approach to the role of pre-S mRNA or downstream AUGs in the synthesis of active replication complexes may help to resolve these issues and to yield a better picture of how multiple *pol* gene products are actually synthesized during productive hepadnaviruses infections.

ACKNOWLEDGMENTS

We are grateful to A. Jilbert, R. Katz, C. Seeger, A. Skalka, J. Summers, and J. Taylor for helpful suggestions, to C. House for help in the laboratory, and to A. Capriotti and J. Ford for assistance in the preparation of the manuscript.

This research was supported by Public Health Service grants AI-18641, CA-06927, and RR-05539 from the National Institutes of Health and by an appropriation from the Commonwealth of Pennsylvania. L.D.C. was supported by postdoctoral fellowship AI-07947 from the National Institutes of Health. T.-T. Wu is a graduate student in the Department of Pathology, Biomedical Graduate School, University of Pennsylvania.

REFERENCES

1. **Bartenschlager, R., and H. Schaller.** 1988. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO J. **7**:4185–4192.
2. **Bavand, M. R., M. Feitelson, and O. Laub.** 1989. The hepatitis B virus associated reverse transcriptase is encoded by the viral *pol* gene. J. Virol. **63**:1019–1021.
3. **Bavand, M. R., and O. Laub.** 1988. Two proteins with reverse transcriptase activities associated with hepatitis B virus-like particles. J. Virol. **62**:626–628.
4. **Bosch, V., R. Bartenschlager, G. Radziwill, and H. Schaller.** 1988. The duck hepatitis B virus P-gene codes for protein strongly associated with the 5′ end of the viral DNA minus strand. Virology **166**:475–485.
5. **Büscher, M., W. Reiser, H. Will, and H. Schaller.** 1985. Transcripts and putative RNA pregenome of duck hepatitis B virus. Implications for reverse transcription. Cell **40**:717–724.
6. **Calos, M. P., J. S. Lebrowski, and M. R. Botchan.** 1983. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. USA **80**:3015–3019.
7. **Chang, C., G. Enders, R. Sprengel, N. Peters, H. E. Varmus, and D. Ganem.** 1987. Expression of the precore region of an avian hepatitis virus is not required for viral replication. J. Virol. **61**:3322–3325.
8. **Chang, L.-J., P. Pryciak, D. Ganem, and H. E. Varmus.** 1989. Biosynthesis of the reverse transcriptase of hepatitis B viruses involves *de novo* translational initiation not ribosomal frameshifting. Nature (London) **337**:364–368.
9. **Chen, P.-J., C.-R. Chen, J.-L. Sun, and D.-S. Chen.** 1989. Identification of a doubly spliced viral transcript joining the separated domains for putative protease and reverse transcriptase of hepatitis B virus. J. Virol. **63**:4165–4171.
10. **Condreay, L. D., C. E. Aldrich, L. Coates, W. S. Mason, and T.-T. Wu.** 1990. Efficient duck hepatitis B virus production by an avian liver tumor cell line. J. Virol. **64**:3249–3258.
11. **Cullen, B. R.** 1986. *trans*-activation of human immunodeficiency virus occurs via a bimodal mechanism. Cell **46**:973–982.
12. **Gerlich, W. H., and W. S. Robinson.** 1980. Hepatitis B virus contains protein attached to the 5′ terminus of its complete strand. Cell **21**:801–809.
13. **Graham, F.-L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
14. **Halpern, M. S., J. M. England, D. T. Deery, D. J. Petcu, W. S. Mason, and K. L. Molnar-Kimber.** 1983. Viral nucleic acid synthesis and antigen accumulation in pancreas and kidney of Pekin ducks infected with duck hepatitis B virus. Proc. Natl. Acad. Sci. USA **80**:4865–4869.
15. **Halpern, M. S., J. M. England, L. Flores, J. Egan, J. Newbold,**

and W. S. Mason. 1984. Individual cells in liver, kidney, and pancreas of DHBV infected ducks express antigens cross-reactive with those on viral surface antigen particles and immature cores. Virology **137**:408–413.

16. **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.
17. **Horwich, A. L., K. Furtac, J. Pugh, and J. Summers.** 1990. Synthesis of hepadnavirus particles containing replication defective duck hepatitis B virus genomes in cultured HuH7 cells. J. Virol. **64**:642–650.
18. **Jacks, T., H. D. Madhani, F. R. Masiarz, and H. E. Varmus.** 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. Cell **55**:447–458.
19. **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.
20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–682.
21. **Li, J.-S., L. Cova, B. Buckland, V. Lambert, G. Deleage, and C. Trepo.** 1989. Duck hepatitis B virus can tolerate insertion, deletion, and partial frameshift mutation in the distal pre-S region. J. Virol. **63**:4965–4968.
22. **Lien, J.-M., C. E. Aldrich, and W. S. Mason.** 1986. Evidence that a capped oligoribonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J. Virol. **57**:229–236.
23. **Mack, D. H., W. Bloch, N. Nath, and J. Sninsky.** 1988. Hepatitis B virus particles contain a polypeptide encoded by the largest open reading frame: a putative reverse transcriptase. J. Virol. **62**:4786–4790.
24. **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.
25. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 88–93. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Molnar-Kimber, K. L., J. Summers, J. M. Taylor, and W. S. Mason.** 1983. Protein covalently bound to minus-strand DNA intermediates on duck hepatitis B virus. J. Virol. **45**:165–172.
27. **Nakabayashi, H., K. Taketa, K. Miyano, T. Yamane, and J. Sato.** 1982. Growth of human hepatoma cell lines with differentiated functions in chemically defined medium. Cancer Res. **42**:3858–3863.

27a. **Oberhaus, S., and J. Newbold.** Personal communication.

28. **Petcu, D. J., C. E. Aldrich, L. Coates, J. M. Taylor, and W. S. Mason.** 1988. Suramin inhibits *in vitro* infection by duck hepatitis B virus, Rous sarcoma virus, and hepatitis delta virus. Virology **167**:385–392.
29. **Pugh, J. C., J. J. Sninsky, J. W. Summers, and E. Schaeffer.** 1987. Characterization of a pre-S polypeptide on the surfaces of infectious avian hepadnavirus particles. J. Virol. **61**:1384–1390.
30. **Pugh, J. C., K. Yaginuma, K. Koike, and J. Summers.** 1988. Duck hepatitis B virus (DHBV) particles produced by transient expression of DHBV DNA in a human hepatoma cell line are infectious in vitro. J. Virol. **62**:3513–3516.
31. **Radziwill, G., W. Tucker, and H. Schaller.** 1990. Mutational analysis of the hepatitis P gene product: domain structure and RNase H activity. J. Virol. **64**:613–620.
32. **Roychoudhury, S., and C. Shih.** 1990. *cis* rescue of a mutated reverse transcriptase of human hepatitis B virus by creation of an internal AUG. J. Virol. **64**:1063–1069.
33. **Schlicht, H.-J., G. Radziwill, and H. Schaller.** 1989. Synthesis and encapsidation of duck hepatitis B virus reverse transcriptase do not require formation of core-polymerase fusion proteins. Cell **56**:85–92.
34. **Schlicht, H.-J., J. Salfeld, and H. Schaller.** 1987. The duck hepatitis B virus pre-C region encodes a signal sequence which is essential for synthesis and secretion of processed core proteins but not for virus formation. J. Virol. **61**:3701–3709.
35. **Seeger, C., and J. Maragos.** 1989. Molecular analysis of the function of direct repeats and a polypurine tract for plus-strand DNA priming in woodchuck hepatitis virus. J. Virol. **63**:1907–1915.
36. **Su, T.-S., C.-J. Lai, J.-L. Huang, L.-H. Lin, Y.-K. Yauk, C. M. Chang, S. J. Lo, and S.-H. Han.** 1989. Hepatitis B virus transcript produced by RNA splicing. J. Virol. **63**:4011–4018.
37. **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.
38. **Summers, J., P. M. Smith, and A. L. Horwich.** 1990. Hepadnavirus envelope proteins regulate covalently closed circular DNA amplification. J. Virol. **64**:2819–2824.
39. **Suzuki, T., N. Masui, K. Kajino, I. Saito, and T. Miyamura.** 1989. Detection and mapping of spliced RNA from a human hepatoma cell line transfected with the hepatitis B virus genome. Proc. Natl. Acad. Sci. USA **86**:8422–8426.
40. **Tuttleman, J. S., J. C. Pugh, and J. W. Summers.** 1986. In vitro experimental infection of primary duck hepatocyte cultures with duck hepatitis B virus. J. Virol. **58**:17–25.
41. **Wahl, G., M. Stern, and G. R. Stark.** 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzybxymethyl paper and rapid hybridization by using dextran sulfate. Proc. Natl. Acad. Sci. USA **76**:3683–3687.
42. **Wilson, W., M. Braddock, S. E. Adams, P. D. Rathjen, S. M. Kingsman, and A. J. Kingsman.** 1988. HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. Cell **55**:1159–1169.
43. **Wilson, W., M. H. Malim, J. Mellor, A. J. Kingsman, and S. M. Kingsman.** 1986. Expression strategies of the yeast retrotransposon Ty: a short sequence directs ribosomal frameshifting. Nucleic Acids Res. **14**:7001–7016.
44. **Yu, M., and J. Summers.** Personal communication.