JOURNAL OF VIROLOGY, Feb. 1990, p. 613–620
0022-538X/90/020613-08$02.00/0

Vol. 64, No. 2

Mutational Analysis of the Hepatitis B Virus P Gene Product: Domain Structure and RNase H Activity

GERALD RADZIWILL,† WILL TUCKER, AND HEINZ SCHALLER*

Zentrum für Molekulare Biologie, University of Heidelberg, Im Neuenheimer Feld 282, 6900 Heidelberg, Federal Republic of Germany

Received 7 August 1989/Accepted 23 October 1989

To correlate the hepatitis B virus P gene with the enzymatic activities predicted to participate in hepadnavirus reverse transcription, a series of P gene mutants containing missense mutations, in-phase insertions, and in-phase deletions was constructed by site-directed mutagenesis. These mutants were tested in the context of otherwise intact hepatitis B virus genomes for the ability to produce core particles containing the virus-associated polymerase activity. The results obtained suggest that the P protein consists of three functional domains and a nonessential spacer arranged in the following order: terminal protein, spacer, reverse transcriptase/DNA polymerase, and RNase H. The first two domains are separated by a spacer region which could be deleted to a large extent without significant loss of endogenous polymerase activity. In cotransfection experiments, all P gene mutants could be complemented in *trans* by constructs expressing the wild-type gene product but not by a second P gene mutant. This indicates that the multifunctional P gene is expressed as a single translational unit and independent of the core gene and furthermore that the gene product is freely diffusible and not processed before core assembly.

Hepatitis B viruses (HBV), or hepadnaviruses, are characterized by a small circular DNA genome of about 3 kilobases which is not covalently closed in that it is gapped in one DNA strand (the plus strand) and contains a terminal protein linked to the 5′ end of the complementary DNA minus strand. This peculiar structure arises intracellularly by reverse transcription of an RNA pregenome through a process which apparently involves the terminal protein during priming of DNA synthesis and which then, by analogy to retroviruses and retroelements, is assumed to be catalyzed by the concerted action of DNA polymerase/reverse transcriptase (RT) and RNase H activities (25).

It has been an open question whether these enzymes are all encoded in the viral genome (12). A virus-associated DNA polymerase has been detected on the basis of its ability to repair the gap in the viral plus strand in vitro (16). Furthermore, sequence analysis indicates that a large open reading frame covering about 80% of the viral genome, and designated the P frame, is the most probable candidate for encoding this enzyme. Finally, short sequences homologous to conserved sequences in retroviral RTs support the notion that at least the hepadnavirus RT is P frame encoded (28).

Initial attempts to prove this hypothesis by biochemical analysis of the viral DNA polymerase turned out to be unsuccessful, since the hepadnavirus enzyme, unlike the retrovirus enzyme, was found to be tightly linked to the viral genome and unable to switch to an exogenously supplied template (20). Therefore, biochemical studies of the DNA polymerase activity have so far been limited to the use of the viral genome as an endogenous template. As for the RNase H activity, there is no experimental evidence of its existence or of the gene that encodes it.

In this report, we use a mutational analysis of the HBV P gene and its transient expression in HepG2 cells (1, 8, 26, 29, 31) to show that the P gene product does indeed contain the hepadnavirus DNA polymerase/RT and the viral RNase H and, furthermore, that these catalytic domains are separated by a nonessential spacer region from a third domain encoding the HBV DNA terminal protein.

* Corresponding author.
† Present address: MPI für Molekulare Genetik, 1000 Berlin, Federal Republic of Germany.

MATERIALS AND METHODS

Plasmid construction. (i) Basic constructs. The basic constructs used for creating all mutants are shown in Fig. 1. The plasmid used as wild type, pMH −34/2922, contains a terminally redundant HBV genome from nucleotide position (NT) 2922 to NT 84 cloned downstream of the metallothionein promoter upstream element (NT −280 to NT −34) (modified from reference 15; M. Niepmann, Ph.D. thesis, University of Heidelberg, Heidelberg, Federal Republic of Germany, 1990). The single-stranded bacteriophages M13 mp18-HBV(+) and M13 mp18-HBV(−) contain the full-length HBV sequence (*Eco*RI-*Eco*RI) plus and minus strand, respectively, in the M13 vector mp18.

(ii) Point mutations. Point mutations were constructed by oligonucleotide-directed mutagenesis according to a standard protocol (32) by using M13 mp18-HBV(+) or M13 mp18-HBV(−). The mismatched oligonucleotides used, the resulting nucleotide exchanges, and the newly created restriction enzyme cleavage sites are listed in Table 1.

Phage clones containing the mutated sequences were detected by plaque hybridization with the ^{32}P-labeled mutagenizing oligonucleotides. Phage clones were plaque purified, and the nucleotide exchanges were confirmed by restriction enzyme analysis or DNA sequencing. For functional analysis, the mutated sequences were transferred into the HBV wild-type plasmid pMH −34/2922 by using DNA fragments from replicative-form DNA of the mutated M13 mp18-HBV phages. The following fragments were used: mutant P1, *Bst*II-*Bst*II (NT 760 to NT 915); mutant P2, *Spe*I-*Acc*I (NT 1961 to NT 2106); mutants P3 to P5, *Nsi*I-*Nco*I (NT 2346 to NT 2654). The transferred cassettes were sequenced to exclude the possibility that other mutations had occurred.

(iii) Deletion mutants. Mutant HBVΔ12 carries a 401-

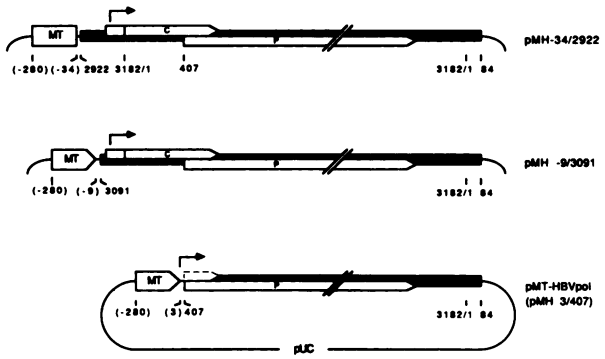

FIG. 1. Schematic representation of the basic plasmid constructs. The HBV DNA is shown as a black bar; the C and P genes are shown as open arrows. The human metallothionein (MT) promoter sequence is shown as an open arrow in the case of a functionally intact promoter and as an open box in the case of a 3′ truncated promoter element. Junction sequences between the MT sequence and the HBV sequence are 13 nucleotides for pMH −34/2922, 5 nucleotides for pMH −9/3091, and 20 nucleotides for pMH 3/407. The transcription initiation sites (M. Niepmann, unpublished data) are indicated by black arrows.

base-pair (bp) deletion in the HBV genome, from *Bam*HI (NT 1009) to *Xho*I (NT 1409), replaced with 17 bp of a synthetic DNA duplex carrying an *Eco*RI site and formed by the two 21-mer oligonucleotides 5′ GATCCGAATTCACCT CTGCAC 3′ and 5′ TCGACTGCAGAGGTGAATTCG 3′ (complementary bases are underlined). Mutant HBVΔ1 carries a 270-bp deletion (NT 1009 to NT 1279), which was obtained by replacing most of the linker sequence inserted in mutant HBVΔ12 with the *Eco*RI-*Xho*I (NT 1280 to NT 1409) HBV wild-type fragment. In mutant HBVΔ2, part of the inserted linker sequence was replaced by the *Bam*HI-*Eco*RI (NT 1004 to NT 1280) HBV wild-type fragment, resulting in a construct in which 123 bp of HBV sequence (NT 1285 to NT 1408) are deleted and 9 bp of foreign sequence remain.

(iv) Linker insertion mutants. Mutants P6, P7, P9, and P10 were constructed by J. Salfeld by inserting an 8-bp linker with the sequence 5′ GAAGCTTC 3′ into the HBV genome contained in plasmid pSH2.1 (6). Blunt-end ligation, carried out after filling in at the following restriction sites: mutant P6, *Ava*I (NT 533); mutant P7, *Bgl*II (NT 937); mutant P9, *Nco*I (NT 2654); and mutant P10, *Ava*I (NT 2745), led to 12-bp in-frame insertions.

(v) Further constructs. Plasmid pMT-HBVpol contains the metallothionein promoter sequence from NT −280 to NT +3 (15) linked by a synthetic DNA sequence of 20 nucleotides to the HBV sequence from NT 407 (first nucleotide of the P gene) to NT 84. This synthetic sequence (ACAAGCTTGGT ACCGCCGCC) supplies an optimized 5′ sequence for translational initiation (17). Mutant C1 carries a single nucleotide deletion at position 221 in the HBV genome contained in plasmid pMH −34/2922 (M. Nassal, unpublished data). This results in a frameshift leading to a truncated C protein of the 74 N-terminal amino acids linked to 9 foreign amino acids.

TABLE 1. Missense mutations in the P gene

Mutant	Oligonucleotide[a]	NT exchange	endoR[b]
P1	NT 794–NT 815	NT 804 A→C NT 806 C→G	*Pst*I
P2	NT 2016–NT 2032	NT 2024 G→C	*Nsi*I
P3	NT 2549–NT 2570	NT 2559 G→C	
P4	NT 2570–NT 2589	NT 2578 T→C NT 2580 C→A	*Pvu*I
P5	NT 2609–NT 2624	NT 2616 A→T	*Hpa*I

[a] Position of the mutagenic oligonucleotides.
[b] New restriction enzyme cleavage sites created by nucleotide exchanges.

Detection of core particle-associated HBV polymerase activity after transient expression of HBV DNA in tissue culture. Transfection of HepG2 cells and isolation of cytoplasmic core particles were carried out by methods described previously (22), with minor modifications. Briefly, HepG2 cells grown in 10-cm dishes were transfected by the $Ca_3(PO_4)_2$ precipitation technique with 3.5 pmol of plasmid (about 15 μg of DNA) or, in the case of linker insertion mutants P6 to P10, circularized HBV DNA (about 7 μg). For complementation assays, 3.5 pmol of each of two plasmids were pooled and cotransfected. Transfection efficiencies were determined by testing a portion of the culture medium 3 days after transfection for the presence of hepatitis B e antigen and hepatitis B surface antigen by radioimmunoassay (Abbott Laboratories, North Chicago, Ill.). The values did not vary by more than ±20%, except for mutants HBVΔ1 and HBVΔ12, for which the values were reduced to one-third that of values for the wild type.

The medium was collected, and the cells were lysed in 500 μl of TNE (10 mM Tris hydrochloride [Tris-HCl] [pH 8.0], 100 mM NaCl, 0.1 mM EDTA)–1% Nonidet P-40. The lysate was clarified by centrifugation, and the supernatant was subjected to immunoprecipitation. Antisera used for immunoprecipitation of cytoplasmic core particles were directed against a recombinant hepatitis B core antigen (30). A portion of the immunoprecipitated material was suspended in 30 μl of polymerase buffer (50 mM Tris-HCl [pH 8.0], 40 mM $MgCl_2$, 50 mM NH_4Cl, 1% Nonidet P-40, 0.3% β-2-mercaptoethanol) supplemented with 12.5 μM each of dCTP, dGTP, and dTTP and 10 μCi of [α-^{32}P]dATP (3,000 Ci/mmol; Amersham Buchler GmbH & Co. KG, Braunschweig, Federal Republic of Germany) and incubated for 90 min at 37°C. After addition of unlabeled dATP to 12.5 μM, the incubation was continued for an additional 60 min. Subsequently, sodium dodecyl sulfate (SDS) and proteinase K were added to final concentrations of 1% and 500 μg/ml, respectively, and after 60 min at 37°C the sample was passed through a Sephadex G100 column to separate the unincorporated [α-^{32}P]dATP from the labeled nucleic acids. The ^{32}P-labeled nucleic acids were ethanol precipitated, dissolved in TNE–1% SDS, and analyzed on a 1.5% agarose gel containing 0.1% SDS (7).

Annealing of reaction products to cloned HBV plus- and minus-strand DNA. Strandedness of DNA products labeled in the endogenous polymerase assay was determined by annealing to M13 single-stranded phage DNA containing a full-length plus or minus DNA strand of HBV. Samples were denatured by heating for 5 min to 100°C in 10 μl of TE (10 mM Tris-HCl [pH 8.0], 0.1 mM EDTA) and then incubated for 90 min at 68°C at a final volume of 20 μl in 0.6 M NaCl–40 mM Tris-HCl [pH 7.5]–2 mM EDTA containing 100 ng of M13 mp18-HBV(+) or M13 mp18-HBV(−) DNA (25). Two micrograms of tRNA was added, and the nucleic acids were ethanol precipitated, resuspended in TE, and analyzed on a 1% agarose gel.

Test for RNA-DNA hybrid molecules. RNA-DNA hybrid molecules were detected by incubating the ^{32}P-labeled products of the endogenous polymerase assay for 30 min at 37°C with 10 U of avian myeloblastosis virus reverse transcriptase (AMV-RT) (Life Sciences, Inc., St. Petersburg, Fla.), 0.6 U of *Escherichia coli* RNase H (Boehringer GmbH, Mann-

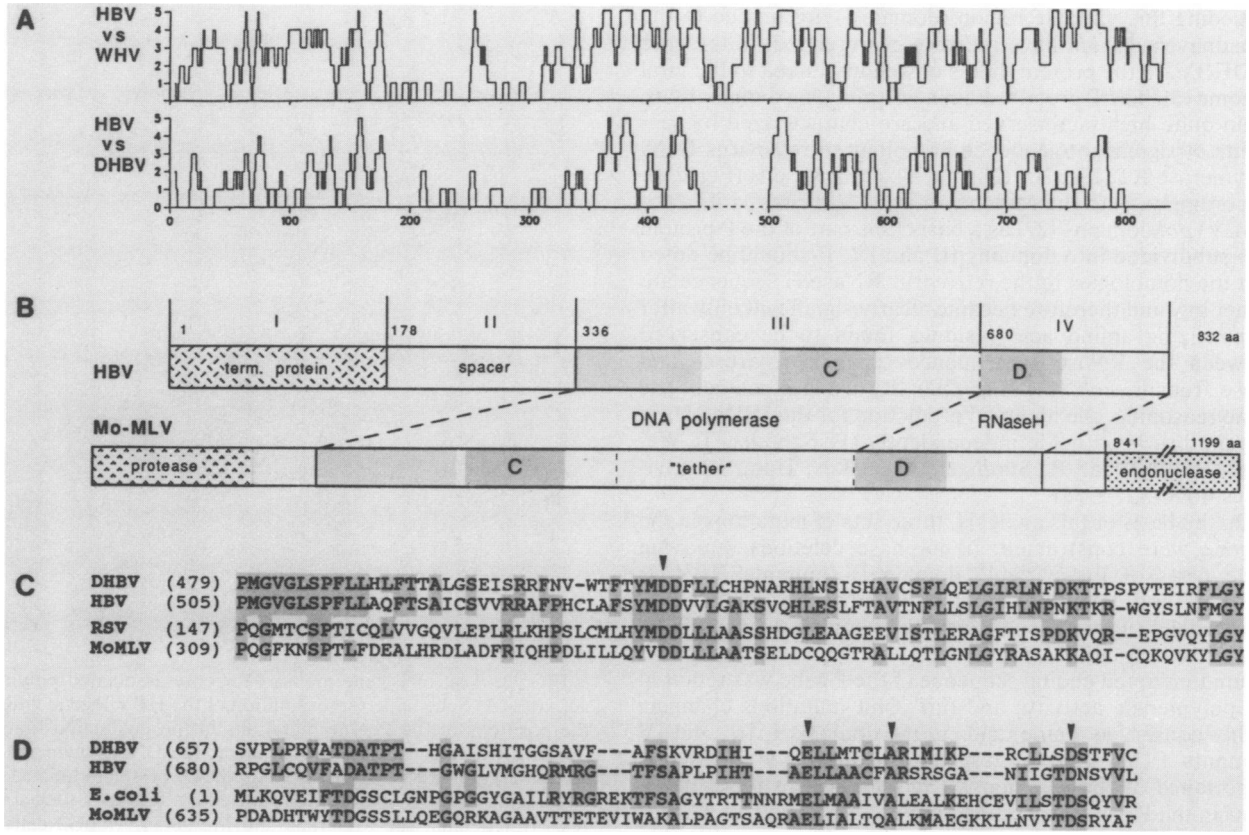

FIG. 2. Structural and functional organization of the HBV P gene as derived from a comparison of its sequence to sequences of other hepadnaviruses and to those of retrovirus *pol* genes. (A) Study of sequence similarity between the P gene products of mammalian hepadnaviruses (HBV, WHV) and those of mammalian (HBV) and avian (DHBV) hepadnaviruses (for details, see Materials and Methods). (B) Segmentation of the P gene into four regions corresponding to regions of high or low amino acid identity between different hepadnaviruses and amino acid sequence similarities to the domain of RT/RNase H retroviruses. (C) Alignment of the amino acid sequence of the HBV P gene segment shaded and labeled C in panel B (amino acids [aa] 505 to 587) with the corresponding DHBV sequence (aa 479 to 561) and a region of the retrovirus P gene product of Rous sarcoma virus (RSV) (aa 147 to 228) and Mo-MuLV (MoMLV) (aa 309 to 391). Amino acids identical between HBV and DHBV or between one of the hepadnaviruses and RSV and/or Mo-MuLV are shaded. (D) Alignment of the amino acid sequence of the HBV P gene segment in region IV shaded and labeled D in panel B (aa 680 to 741) with the corresponding DHBV sequence (aa 657 to 719), the N-terminal part of *E. coli*. RNase H (aa 1 to 74), and a region of the RNase H of Mo-MuLV (aa 635 to 705). Amino acids identical between HBV and DHBV or between one of the hepadnaviruses and *E. coli* and/or Mo-MuLV are shaded. Highly conserved amino acids that were changed in the HBV P gene by point mutations (amino acid positions 540, 718, 725, and 737; Table 2) are indicated by arrowheads.

heim, Federal Republic of Germany), or RNase A (10 μg/ml) in an appropriate buffer and by analyzing them on a 1% agarose gel. The buffer for the RNase H reactions with *E. coli* RNase H or AMV-RT consisted of 20 mM Tris-HCl (pH 8.3), 10 mM $MgCl_2$, 1 mM dithiothreitol, and 50 mM NaCl; the buffer used for completing the viral DNA synthesis with AMV-RT consisted of 50 mM Tris-HCl (pH 8.3), 6 mM $MgCl_2$, 10 mM dithiothreitol, 100 mM NaCl, and dATP, dCTP, dGTP, and dTTP, each at 200 μM.

Computer sequence analysis. Computer sequence analysis was carried out by using a program for sequence similarity comparison (O. Steinau, unpublished data). A "window" of five amino acids in the HBV sequence (11) was compared, in one-amino-acid steps, to the sequences from woodchuck hepatitis B virus (WHV) or duck hepatitis B virus (DHBV) (10, 23), resulting in an identity value between 0 and 5. For alignment with the HBV P amino acid sequence, the following changes were introduced into the WHV and the DHBV sequences: WHV deletions, amino acids 18 to 22, 184 to 223, and 356 to 357; empty space, two amino acids between amino acids 501 and 502; DHBV deletions, amino acids 19 to 51, 360 to 364, 549, 647 to 649, and 671; empty space, between amino acids 127 and 128 (10 amino acids), between amino acids 471 and 472 (54 amino acids), and between amino acids 504 and 505 (1 amino acid).

RESULTS

Prediction of functional domains in the P protein and design of mutations. To identify functionally important parts of the HBV P gene product which might be used as preferred targets for a mutational analysis, a sequence comparison of the P gene products encoded in the different hepadnaviruses was carried out. Sequences conserved between hepadnavirus P proteins were also compared to functionally defined amino acid sequences of retrovirus *pol* gene products (Fig. 2C and D). The first part of this analysis (Fig. 2A) shows that the hepadnavirus P protein sequence is well conserved, even between the quite distantly related avian and mammalian members of this virus family, except for a rather variable region covering the second quarter of the gene (amino acids 178 to 336; domain II in Fig. 2B [23]). The sequences

preceding this variable region (domain I) are unique to the hepadnaviruses and have recently been related, in the case of DHBV, to the protein that is covalently linked to the viral genome (5). The P protein sequences following domain II are again quite highly conserved and are characterized by segments of significant sequence homology to retrovirus DNA polymerase/RT (Fig. 2C) (28) and also to RNase H (Fig. 2D). Accordingly, with the Moloney murine leukemia virus (Mo-MuLV) *pol* domains (27) as a basis, this part of the P protein was subdivided into domains III and IV. It should be noted that the homologies to the retrovirus RNase H sequence are rather low and therefore became clearly significant only after emphasizing amino acid residues found to be conserved between the RNase H sequences from retroviruses and those from *E. coli* (14) (Fig. 2D). By using essentially the same reasoning, the identical prediction for the HBV RNase H gene has been made independently (F. Schödel, T. Weimer, H. Will, and R. Sprengel, AIDS Res. Hum. Retroviruses **4**:IX–XI, 1988).

On the basis of this analysis, three sets of mutations in the P gene were constructed: (i) in-phase deletions removing large parts of the variable domain II (mutants HBVΔ1, HBVΔ2, and HBVΔ12), to test whether this region was dispensible; (ii) in-phase insertions in domain I (mutants P6 and P7) or domain IV (mutants P9 and P10), to test whether an uninterrupted coding sequence of the P gene was required for polymerase activity; and (iii) point mutations changing highly conserved amino acids in the domains I, III, and IV (mutants P1 to P5), to test for the predicted functional importance of these regions. The locations of these mutations in the P gene are outlined in Fig. 3A, while the changes in nucleotide sequence and amino acid sequence are listed in Tables 1 and 2. More details of their construction are given in Materials and Methods.

Functional analysis of P gene mutations. The P gene mutations described above were tested for their effect on the synthesis of viral gene products in the context of cloned full-length HBV genomes capable of expressing all viral proteins and of producing viruslike particles indistinguishable from authentic virus. To facilitate the transfer of mutated viral DNA segments, two plasmids were used. Both specify the synthesis of the HBV pregenomic RNA (and its translation products from the core gene and P gene) from terminally redundant HBV genomes without any need for excision and circularization prior to transfection. As outlined in Fig. 1, these constructs differ by using either the authentic HBV core gene promoter (linked to a foreign upstream element; pMH −34/2922) or a foreign promoter (pMH −9/3091) to direct genomic RNA synthesis. Similarly high yields of viral gene products, in particular of viral cores containing viral nucleic acids and DNA polymerase, were obtained upon transfection of HepG2 cells with either plasmid.

For functional analysis, the mutated viral genomes were introduced and transiently expressed in the human hepatoma cell line HepG2. After 3 days, cytoplasmic core particles were isolated and assayed for the presence of the core-associated DNA polymerase activity incorporating radioactively-labeled deoxynucleotides into the viral genome. Labeled nucleic acids were then characterized by separation on an agarose gel and autoradiography.

Typical results obtained with this assay procedure (designated endogenous polymerase assay) with HBV wild-type DNA and P gene mutants are presented in Fig. 3B and C. When wild-type HBV was used (lane wt), two (or sometimes three) strongly labeled DNA bands, corresponding to linear and circular viral genomes, were observed. The strength and the pattern of these signals were more or less affected when mutant DNA genomes were tested. Thus, mutant P1, which carries two neighboring amino acid exchanges in the N-terminal part of the P gene product, produced essentially the same pattern of labeled DNA as did wild-type HBV, but the

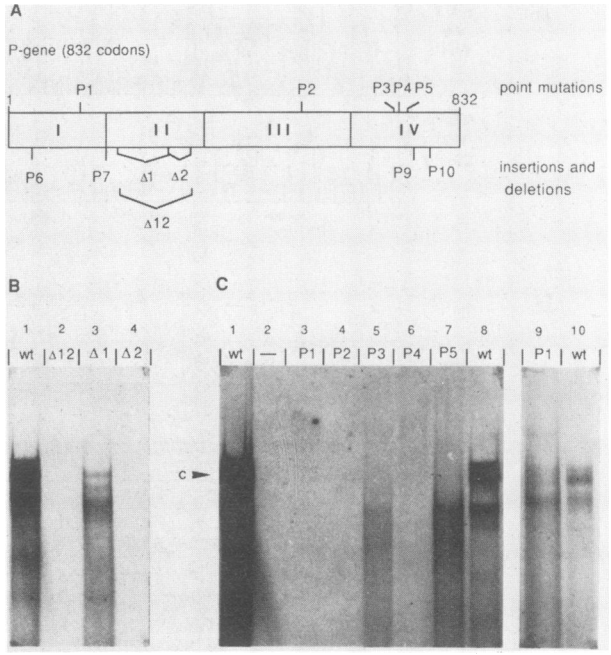

FIG. 3. Test of P gene mutants for core-associated polymerase activity. (A) Schematic representation of the HBV P gene indicating the positions of the P gene mutations in the postulated functional domains (I, terminal protein; II, spacer; III, DNA polymerase; IV, RNase H). (B and C) Analysis of the DNA products labeled in the endogenous polymerase assay, followed by agarose gel electrophoresis (for details, see Materials and Methods). HepG2 cells were transfected with wild-type HBV (wt) or mutant DNA or were mock transfected (—). The same amount of cell lysates was used in each assay, except for lanes 8 (one-third the amount), 9 (one-eighth the amount), and 10 (one-twelfth the amount). The position of the circular HBV DNA is indicated by the arrowhead.

TABLE 2. Changes in the P gene product and their effects on the endogenous polymerase activity

Mutant	Amino acid[a]	Change[b]	% Polymerase activity[c]
P1	133/134	YP→SA	1
P2	540	D→H	<0.3
P3	718	E→H	10
P4	725	A→D	<0.3
P5	737	D→V	20
P6	44< >45	G<KLLG>N	<0.3
P7	178< >179	D<RSFD>L	<0.3
P9	751< >752	W<KLPW>L	<0.3
P10	781< >782	R<EASR>G	<0.3
HBVδ12	201< >335	S<EFTSA>L	<0.3
HBVδ1	201< >292	S< >E	40
HBVδ2	293< >335	F<TSA>L	<0.3

[a] Positions in the P gene product.
[b] Changed amino acids (→) and insertions or deletions (< >).
[c] Incorporation of ^{32}P-labeled nucleotides into viral DNA relative to that in the wild type (about 20,000 cpm) measured in the endogenous polymerase assay (see Materials and Methods).

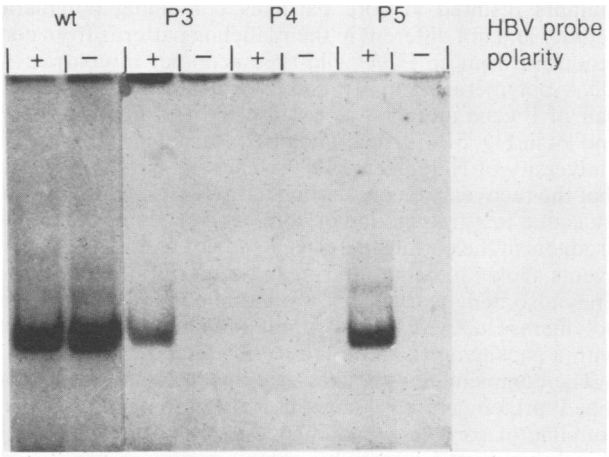

FIG. 4. Test for selective first-strand (minus-strand) synthesis. ^{32}P-labeled products from the endogenous DNA polymerase assay were heat denatured, annealed to phage M13 DNA carrying the HBV genome in plus (lanes +) or minus (lanes −) orientation, and then analyzed for hybrid formation by sizing on a 1% agarose gel (for details, see Materials and Methods). wt, Wild type.

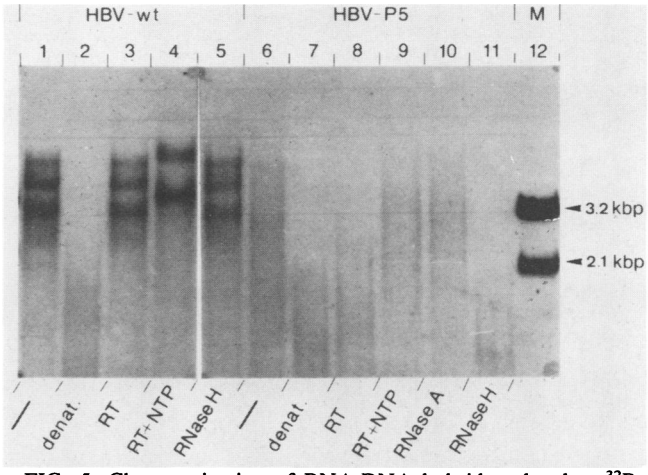

FIG. 5. Characterization of RNA-DNA hybrid molecules. ^{32}P-labeled products of the endogenous DNA polymerase assay using the wild type (HBV-wt) (lanes 1 to 5) or mutant HBV-P5 (lanes 6 to 11) were analyzed on an agarose gel directly after synthesis (lanes 1 and 6), after heat denaturation (lanes 2 and 7), after incubation with AMV-RT/RNase H in the absence (lanes 3 and 8) or presence (lanes 4 and 9) of deoxyribonucleoside triphosphates, after incubation with *E. coli* RNase H (lanes 5 and 11), or after incubation with RNase A (lane 10). The end-labeled *Eco*RI fragments from plasmid pSH2.1 (3.2 kbp = HBV genome; 2.1 kbp = pSV08 [6]) were used as size markers.

signals were reduced to about 1% and were therefore detectable only after the amount of the material analyzed was increased (lanes 3 and 9). Mutants P2 and P4, which contain single-amino-acid exchanges in the regions homologous to retrovirus RT and RNase H, respectively, were completely negative in the endogenous polymerase assay (lanes 4 and 6), as were mutants P6, P7, P9, and P10, which encode insertions of four amino acids in domains I or IV of the P protein (Table 2). Mutants P3 and P5, carrying independent missense mutations in the predicted RNase H region, displayed substantially reduced incorporation in the endogenous polymerase assay (Table 2), with an altered pattern of labeled nucleic acids, the upper signals that represent circular viral genomes being absent (lanes 5 and 7). Finally, of the three deletions in domain II (Fig. 3B), two abolished polymerase activity; the mutation in HBVΔ1, although deleting 10% of the P protein sequence, displayed nearly normal levels of DNA synthesis on a slightly shortened HBV DNA genome (Fig. 3B, lane 3).

Taken together, these results demonstrate that missense mutations and in-phase insertions in any one of the three conserved P gene domains that are assumed to encode the terminal protein, the DNA polymerase, and the RNase H are critical for catalytic activity, as determined in the endogenous polymerase assay. In contrast, a large segment can be removed from domain II without abolishing the synthesis of viral genomes as detected in this assay.

Mutants P3 and P5 selectively abolish RNase H activity. As shown above, there were two mutants with missense mutations in domain IV, HBV P3 and HBV P5, whose labeling pattern of core-associated nucleic acids differed from that of the HBV wild type in that they lacked the signals representing circular double-stranded HBV DNA. Since these mutations change amino acids that are highly conserved in RNase H sequences of various origin (Fig. 2D), it appeared likely that the observed DNA products might be related to the persistence of hybrid molecules consisting of genomic RNA and minus-strand DNA, a result expected to accompany a loss of RNase H activity. To test this assumption, the labeled DNA products from all three RNase H mutants were analyzed for their strandedness by annealing to HBV DNA single strands cloned into M13 phage vectors in minus- or plus-strand orientation (Fig. 4). As expected, the labeled DNA products from mutants P3 and P5 hybridized selectively with the plus-strand HBV probe but not with the HBV minus strand; mutant P4, which gave no DNA signal in the original sizing analysis (Fig. 3C, lane 6), was again negative in this more sensitive assay in which all labeled DNA products are concentrated into a single band. In the control with HBV wild type, the labeled DNA annealed equally well to both HBV DNA strands. These results demonstrate that, in contrast to the mutant P4, which is totally defective in viral DNA synthesis, HBV mutants P3 and P5 are still capable of synthesizing viral DNA minus strands, although at a reduced level, but not of synthesizing plus strands.

Further evidence for this notion was obtained by more detailed structural analysis of the RNA-DNA hybrid molecules produced by the RNase H mutants and labeled in the endogenous polymerase assay. To demonstrate the presence of an RNA moiety, these molecules were subjected to RNase H digestion using either *E. coli* RNase H or AMV-RT in the absence of deoxynucleoside triphosphates. Additional samples were treated with RNase A, which does not degrade RNA-DNA hybrids, or were heat denatured to obtain the labeled DNA as single strands. As a control, the same treatments were carried out with the labeled DNA products produced by the HBV wild type.

The results obtained in these experiments for the DNA products from mutant P5 are shown in Fig. 5. They demonstrate that incubation with either RNase H activity (lanes 8 and 11) altered the electrophoretic mobility of the labeled DNA products such that it became comparable to that of the DNA single strands obtained after heat denaturation (lane 7). In contrast, the pattern of labeled DNA did not change after repair DNA synthesis with AMV-RT in the presence of deoxynucleotides (lane 9) or after incubation with RNase A (lane 10). Thus, the RNA pregenome had been converted into RNA-DNA duplex structures of heterogeneous sizes

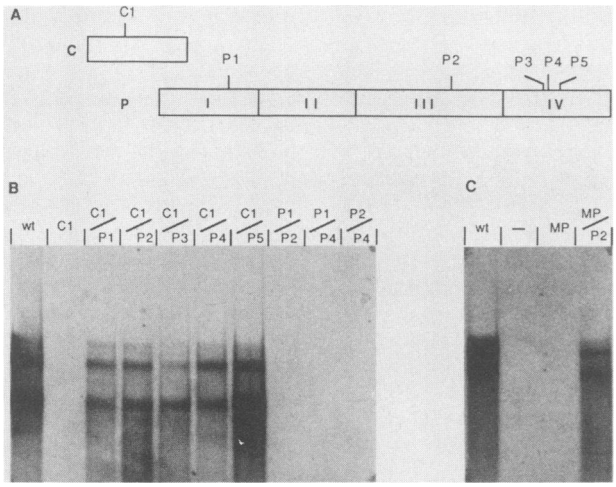

FIG. 6. Complementation of P gene mutants in *trans* by constructs producing the wild-type P gene product. (A) Positions of the individual mutations in the HBV P gene; mutant C1 carries the wild-type P gene but a defective core gene. (B) Results obtained in the endogenous polymerase assay after cotransfection of two defective HBV genomes. (C) Complementation of mutant P2 with a construct expressing the P gene under foreign promoter control (pMT-HBVpol, MP; for plasmid structure, see Fig. 1). The core particle-associated polymerase activity was detected as described in the legend to Fig. 3.

and of less than full genome length. Surprisingly, in the control experiment with the wild type, reaction products with added AMV-RT plus deoxynucleoside triphosphates resulted in more slowly migrating DNA bands, indicating that gap filling had been incomplete during the endogenous polymerase assay. Results similar to those described here for mutant P5 were also obtained with mutant P3 (data not shown).

Taken together, the results described above are all consistent with the notion that HBV mutants P3 and P5 produce core particles containing a P gene product that can function as RT but not as RNase H. Therefore, we conclude that the most 3′-terminal conserved domain of the P gene (domain IV) encodes an hepadnavirus RNase H.

P gene mutations can be compensated for by intergenic, but not intragenic, complementation. The results of the mutational analysis described above indicate that two of the three missense mutations in domain IV (P3 and P5) inactivated the RNase H function residing in this domain without significantly affecting the priming and chain-elongation reaction required for minus (first) strand synthesis. Although no first DNA strand was synthesized by the third mutant (HBV P4), this result may be taken to indicate that domains I and III function independently of domain IV. It also raises the general question of whether P protein segments can function autonomously, i.e., can complement each other in *trans* after being released by proteolysis from the primary P gene product. To test this possibility experimentally, P gene mutants were cotransfected in pairs and tested in the endogenous polymerase assay for their ability to reconstitute all P gene functions. As a positive control, each P gene mutant was cotransfected with a construct carrying an intact P gene but also carrying a frameshift mutation in the core gene, and therefore unable by itself to synthesize DNA polymerase-containing core particles (mutant HBV C1; Fig. 6B). Representative results of this analysis are shown in Fig. 6. Cotransfection of mutant C1 with any one of the P gene mutants resulted in core particles containing polymerase activity and not different in their labeling patterns from cores produced from an HBV wild-type genome. In contrast, no DNA polymerase activity could be detected by using any pair of P gene mutants, as exemplified for mutants P1, P2, and P4 in Fig. 5B (further data in G. Radziwill, Ph.D. thesis, University of Heidelberg, 1989). These results demonstrate that the recovery of core-associated DNA polymerase activity is due to a free mixing of core gene products and P gene products in the cotransfected cell and not to recombinational events shown to occur after mixed transfection in vivo (24). They also demonstrate that the mutations introduced affect polymerase activity directly, not indirectly, by interfering with a packaging signal in the RNA template.

The complementation data described above also indicate that P protein synthesis does not depend on a translational coupling of core gene and P gene expression. This conclusion is supported further by complementation experiments with plasmid pMT-HBVpol, a construct in which all HBV sequences upstream of the P gene were removed and replaced by the human metallothionein promoter (Fig. 1). As exemplified in Fig. 6C by using mutant HBV P2, this construct is able to complement polymerase-negative HBV genomes in *trans* so that the normal pattern of labeled (circular and linear) DNA molecules is obtained in the endogenous polymerase assay. Similar results were obtained with the polymerase-defective deletion mutant HBVΔ12, demonstrating that construct pMT-HBVpol produces, in the absence of C gene translation, a fully functional P gene product which can reverse transcribe pregenomic RNA molecules that by themselves are defective in P gene function. These results again support the notion that the three components forming the viral core, namely the genomic RNA, the core protein, and the P protein, exist in a freely diffusible state and that they thus may originate from separate, and potentially defective, viral genomes.

DISCUSSION

The results of the present mutational analysis of the HBV P gene lead to several conclusions regarding the functional and structural organization of the gene. First, we provide strong experimental evidence to support the long-standing prediction that the P gene encodes the hepadnavirus DNA polymerase/RT in that mutations throughout the P frame drastically reduced the enzymatic activity detected in the endogenous polymerase assay. This was the case for several in-phase insertions and in-phase deletions, as well as for five missense mutations affecting highly conserved amino acid sequence motifs (Fig. 2 and 3). Of the latter, mutation P2 at amino acid 540 is particularly noteworthy since the target sequence (YMDD) is present in many retrovirus RTs (2) and has been identified as being essential for RT activity in the case of human immunodeficiency virus by changing YMDD to YMHD (18) as was done in creating HBV P2. Therefore, we conclude that the domain containing the YMDD motif (domain III) encodes the HBV DNA polymerase/RT.

Second, point mutations near the C terminus of the P gene product (plus strand) (domain IV) led to the complete loss of second strand DNA synthesis in the endogenous polymerase assay. This confirmed experimentally the RNase H function which we predicted for this region from a low degree of sequence similarity to RNase H sequences from retroviruses and *E. coli* (Fig. 2D). Domain IV seems to contribute to some extent to the DNA polymerase activity residing in the neighboring domain III since all three RNase H mutations

lead, in addition to abolishing plus-strand DNA synthesis, to a significant reduction (P3, P5) or even complete loss (P4) of minus-strand DNA synthesis (Table 2 and Fig. 4).

Third, an extended spacer which is to a large extent nonessential for polymerase function was identified in the second quarter of the HBV P gene (domain II). This spacer separates the RT/RNase H region from a domain located in the N terminus region that is well conserved between hepadnaviruses, but not related to any retroviral protein, and that was recently shown to encode the protein that is covalently linked to the viral genome (DNA terminal protein [5]).

Because quite similar results have been obtained in parallel studies with mutated DHBV genomes (5, 21; unpublished results from our laboratory), we conclude that the hepadnavirus P gene product is a multifunctional protein with several distinct domains that catalyze the major steps in hepadnavirus reverse transcription, i.e., priming, DNA synthesis, and removal of the RNA template. Compared with the retrovirus *pol* gene product (Fig. 2B), the hepadnavirus P protein lacks an endonuclease, which is not surprising since there is no specific integration into the host genome. Furthermore, the retrovirus protease is replaced by the HBV DNA terminal protein, which is consistent with the absence of core-polymerase fusions in hepadnaviruses (9, 21). RT and RNase H remain arranged in the same order but are located in HBV near the end of the P gene. Another distinct difference is the presence of the extended spacer/tether that links the HBV terminal protein to the catalytic domains but that is absent in retroviruses, in which a protease precursor is directly adjacent to the RT/RNase H precursor (19). This HBV tether does not correspond to the tether predicted to connect RT and RNase H in the retrovirus *pol* gene product (14) but meanwhile shown to be part of the polymerase domain (27). Finally, it should be noted that the HBV DNA polymerase (domain III) appears to be divided into two subdomains by sequences around amino acid 460 which are poorly conserved between mammalian hepadnaviruses and totally removed by a 54-amino-acid deletion in avian viruses (Fig. 2A) and which tolerate mutational changes without affecting DNA polymerase activity (R. Bartenschlager, unpublished data).

The model organizing the P gene in several functional domains as proposed above raises the questions of whether these domains are able to function independently of each other in vitro, or more importantly, in vivo, and whether they are always synthesized as part of the full-length P gene product. As discussed further below, no satisfactory assays are presently available that would allow reliable monitoring of P gene functions in vitro with exogenous substrates, as can be done with retrovirus *pol* gene products. Therefore, in the present study we were forced to rely on the endogenous polymerase assay. This assay, however, monitors only DNA polymerase activity in core particles that arose intracellularly in a complex process involving core assembly, RNA packaging, DNA priming, and DNA synthesis, and therefore primarily only questions regarding the overall process could be addressed. However, by using our mutants also in complementation experiments, additional conclusions regarding P gene function could be drawn. First, since P gene mutations residing in separate functional domains did not complement each other in *trans*, it seems most likely that the P gene is expressed as a single translational unit from its first translational start codon, and not additionally from internal starts, to yield subgenic products (as occurs in pre-S/S gene expression). The absence of intragenic complementation strongly suggests, in addition, that the primary P gene product remains intact, at least until core assembly, and that it is not processed into freely diffusible subunits with partial replication functions. It also argues against processing after assembly, unless there is only a single P protein molecule in the core particle. Thus, our data are consistent with a model in which the P protein remains intact while performing its replication functions (5). This seems to be more likely, since detailed studies with purified DHBV cores support the notion that the P gene product is tightly, probably covalently, linked to its endogenous DNA product/template (5, 20).

This model seems difficult to reconcile with recent reports from Laub and co-workers (3, 4) which describe two HBV-related RT activities detected in activity gels, a major and a minor species migrating at about 70 and 90 kilodaltons, respectively, and seemingly corresponding to a processed and nonprocessed P gene product, respectively. In repeating these experiments, we have also been able to detect RT activities of similar sizes in about equal amounts in extracts from HepG2 cells transfected with replication-competent HBV genomes (Radziwill, Ph.D. thesis). However, the experimental procedure was poorly reproducible in our hands, and in addition, two observations made an interpretation of our results even more difficult. (i) The HBV-related RT was lost during further purification of the core virus fraction used as starting material (as was observed earlier with DHBV [20]), and (ii) surprisingly, there was always a low level of RT activity that migrated the same distance released from mock-transfected HepG2 cells. In the light of these conflicting results, we feel that additional experiments using better assays and better purification procedures will be needed to reach more conclusive answers.

Finally, it should be noted that the demonstration of intergenic complementation between P gene mutants and C gene mutants from HBV indicates that these two genes are expressed independently and act in *trans* and therefore that the HBV P gene product is synthesized uncoupled from the HBV core gene. This conclusion is in keeping with results obtained recently from detailed analyses of the mechanism of P gene expression in DHBV (9, 21) which demonstrate that the DHBV P gene is most probably translated from its own start codon and not, as has been shown for several retroviruses (13), by ribosomal frameshifting from the overlapping core gene.

ACKNOWLEDGMENTS

We thank J. Salfeld for constructing the insertion mutants, O. Steinau for computer analyses, M. Nassal for oligonucleotide synthesis, and H. J. Schlicht and R. Bartenschlager for useful discussions. We are also grateful to N. Schek for critically reading the manuscript, C. Kuhn for help with the figures, and K. Coutinho for typing the manuscript.

This work was supported by grants from the Deutsche Forschungsgemeinschaft (SFB 229) and the Bundesministerium für Forschung und Technologie (BCT 0318/5).

LITERATURE CITED

1. **Acs, G., M. A. Sells, R. H. Purcell, P. Price, R. Engle, M. Shapiro, and H. Popper.** 1987. Hepatitis B virus produced by transfected HepG2 cells causes hepatitis in chimpanzees. Proc. Natl. Acad. Sci. USA **84**:4641–4644.
2. **Argos, P.** 1988. A sequence motif in many polymerases. Nucleic Acids Res. **16**:9909–9916.
3. **Bavand, M., 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.

4. **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.
5. **Bartenschlager, R., and H. Schaller.** 1988. The aminoterminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. EMBO J. **7**:4185–4192.
6. **Cattaneo, R., H. Will, N. Hernandez, and H. Schaller.** 1983. Signals regulating hepatitis B surface antigen transcription. Nature (London) **305**:336–338.
7. **Challberg, M. D., S. V. Desiderio, and T. J. Kelly, Jr.** 1980. Adenovirus DNA replication in vitro: characterisation of a protein covalently linked to nascent DNA strands. Proc. Natl. Acad. Sci. USA **77**:5105–5109.
8. **Chang, C., K. Jeng, C. P. Hu, K. J. Lo, T. S. Su, L.-P. Ting, C. K. Chon, S. H. Han, E. Pfaff, J. Salfeld, and H. Schaller.** 1987. Production of hepatitis B virus in vitro by transient expression of cloned HBV DNA in a hepatoma cell line. EMBO J. **6**:675–680.
9. **Chang, L. J., P. Pryciak, D. Ganem, and H. E. Varmus.** 1989. Biosynthesis of the reverse transcriptase of hepatitis B virus involves de novo translational initiation not ribosomal frameshifting. Nature (London) **337**:364–367.
10. **Galibert, F., T. N. Chen, and E. Mandart.** 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. J. Virol. **41**:51–65.
11. **Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. Nature (London) **281**:646–650.
12. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis-B viruses. Annu. Rev. Biochem. **56**:651–693.
13. **Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, and H. E. Varmus.** 1988. Characterization of ribosomal frameshifting in HIV-I gag-pol expression. Nature (London) **331**: 280–283.
14. **Johnson, M. S., M. A. McClure, D.-F. Feng, J. Gray, and R. F. Doolittle.** 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. USA **83**:7648–7652.
15. **Junker, M., P. Galle, and H. Schaller.** 1987. Expression and replication of the hepatitis-B virus genome under foreign promoter control. Nucleic Acids Res. **15**:10117–10132.
16. **Kaplan, P., R. Greenman, J. Gerin, R. Purcell, and W. Robinson.** 1973. DNA polymerase associated with human hepatitis B antigen. J. Virol. **12**:995–1005.
17. **Kozak, M.** 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J. Mol. Biol. **196**:947–950.
18. **Larder, B. A., D. J. M. Purifoy, K. L. Powell, and G. Darby.** 1987. Site-specific mutagenesis of AIDS virus reverse transcriptase. Nature (London) **327**:716–717.
19. **Mous, J., E. P. Heimer, and S. F. J. Le Grice.** 1988. Processing protease and reverse transcriptase from human immunodeficiency virus type I polyprotein in *Escherichia coli*. J. Virol. **62**:1433–1436.
20. **Radziwill, G., H. Zentgraf, H. Schaller, and V. Bosch.** 1988. The duck hepatitis-B virus DNA polymerase is tightly associated with the viral core structure and unable to switch to an exogenous template. Virology **163**:123–132.
21. **Schlicht, H. J., G. Radziwill, and H. Schaller.** 1989. Synthesis and encapsidation of the duck hepatitis B virus reverse transcriptase does not require the formation of core/polymerase fusion proteins. Cell **56**:85–92.
22. **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.
23. **Sprengel, R., C. Kuhn, H. Will, and H. Schaller.** 1985. Comparative sequence analysis of duck and human hepatitis-B virus genomes. J. Med. Virol. **15**:323–333.
24. **Sprengel, R., H. E. Varmus, and D. Ganem.** 1987. Homologous recombination between hepadnaviral genomes following in vivo DNA transfection: implications for studies of viral infectivity. Virology **159**:454–456.
25. **Summers, J., and W. Mason.** 1982. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell **29**:403–415.
26. **Sureau, C., J. L. Romet-Lemonne, J. I. Mullins, and M. Essex.** 1986. Production of hepatitis B virus by a differentiated human hepatoma cell line after transfection with cloned circular HBV DNA. Cell **47**:37–47.
27. **Tanese, N., and S. P. Goff.** 1988. Domain structure of the Moloney murine leukemia virus reverse transcriptase: mutational analysis and separate expression of the DNA polymerase RNase H activities. Proc. Natl. Acad. Sci. USA **85**:1777–1781.
28. **Toh, H., H. Hayashida, and T. Hiyata.** 1983. Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis-B virus and cauliflower mosaic virus. Nature (London) **305**:827–829.
29. **Tsurimoto, T., A. Fujiyama, and K. Matsubara.** 1987. Stable expression and replication of hepatitis B virus genome in an integrated state in a human hepatoma cell line transfected with the cloned viral DNA. Proc. Natl. Acad. Sci. USA **84**:444–447.
30. **Will, H., J. Salfeld, E. Pfaff, C. Manso, L. Theilmann, and H. Schaller.** 1986. Putative reverse transcriptase intermediates of human hepatitis B virus in primary liver carcinomas. Science **231**:594–596.
31. **Yaginuma, K., Y. Shirakata, M. Kobayashi, and M. Koike.** 1987. Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBV DNA. Proc. Natl. Acad. Sci. USA **84**:2678–2682.
32. **Zoller, M. J., and H. Smith.** 1984. Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. DNA **3**:479–488.