JOURNAL OF VIROLOGY, Sept. 1994, p. 5579–5587
0022-538X/94/$04.00+0

Vol. 68, No. 9

Site-Specific RNA Binding by a Hepatitis B Virus Reverse Transcriptase Initiates Two Distinct Reactions: RNA Packaging and DNA Synthesis

JONATHAN R. POLLACK[1] AND DON GANEM[2,3]*

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

Received 31 March 1994/Accepted 28 May 1994

Hepatitis B viruses encode a polymerase (P) protein with key roles in both reverse transcription and genomic RNA encapsidation. Genetic analysis of *cis*-acting signals required for viral replication implicates an RNA stem-loop structure in both RNA packaging and the initiation of reverse transcription, a process in which P protein also serves as the primer. We now show that duck hepatitis B virus (DHBV) polymerase binds specifically and with high affinity to this RNA stem-loop structure. Mutational analysis indicates that all mutations in the RNA target that inhibit the P protein-RNA interaction inhibit both in vivo RNA packaging and in vitro DNA priming to comparable extents. However, certain mutations in the loop region of the RNA have minimal impact on P protein-RNA binding but are nonetheless severely defective for packaging and DNA synthesis. Thus, P protein-RNA complex formation is necessary but not sufficient to initiate these activities. In addition, examination of RNA binding by truncated P proteins indicates that the C terminus of the polymerase, although required for RNA encapsidation in vivo, is dispensable for RNA binding and DNA priming.

RNA-protein recognition reactions function in an enormous variety of important biological processes, including transcriptional elongation (11), mRNA processing (10, 12), mRNA localization (29), mRNA translation (15, 20, 32), and viral RNA replication (1). As befits their functional diversity, RNA-binding proteins display many different structural features and recognize a wide variety of RNA targets. A recent compilation of primary structural motifs associated with RNA-binding proteins indicates that such proteins can be sorted into at least nine different families (25), and there is every reason to believe that this list will continue to grow. Most of these proteins form RNA-protein complexes that are dedicated to effecting a single reaction. Here we present evidence for a viral RNA-protein interaction in hepatitis B viruses (hepadnaviruses) that appears to be required for two distinct processes: RNA encapsidation and the initiation of reverse transcription.

Hepadnaviruses are small DNA viruses that replicate via reverse transcription of an RNA intermediate; this reaction takes place within subviral nucleocapsids (cores) composed of the core (C) protein, the reverse transcriptase (or polymerase [P]), and the RNA template for the reaction. This RNA (termed pregenomic RNA [pgRNA]) is a terminally redundant transcript that is also the message encoding both the polymerase and the core proteins. The first step in viral genomic replication is the selective encapsidation of pgRNA, together with polymerase, into the nucleocapsid. This step is highly selective: only pgRNA, not subgenomic viral mRNAs or host RNAs, are assimilated into cores (13, 17). Encapsidation is mediated by specific *cis*-acting sequences (termed ε) present at the 5′ end of pgRNA (9, 18, 19). The prime functional determinant in ε is an RNA stem-loop structure consisting of a lower stem, a 6-nucleotide (nt) bulge, and an upper stem with a 6- to 7-nt loop (19, 21, 26). In the human hepatitis B virus (HBV), the bulge and stem structures are required for encapsidation, but their specific nucleotide sequences are not critical; functionally important specific primary sequences reside principally in the loop (26). The ε stem-loop structure is conserved phylogenetically among mammalian and avian hepadnaviruses (19). However, in the duck hepatitis B virus (DHBV), ε sequences are not sufficient to direct RNA encapsidation (18); the presence of a second, auxiliary region of RNA located some 900 nt downstream of ε is required to effect encapsidation (6).

Strong genetic arguments have pointed to key roles for the polymerase in both RNA packaging and DNA synthesis. We (17) and others (2) have shown that P-gene null mutants synthesize morphologically normal capsids that are devoid of viral RNA, implicating polymerase in the selection of pgRNA for encapsidation. It has also been shown that encapsidation of the viral polymerase itself requires ε-containing RNA (3). Thus, the packaging of polymerase and the packaging of pgRNA appear to be tightly coupled. Following RNA encapsidation, extensive reverse transcription of the viral genome takes place. Hepadnavirus DNA synthesis is a highly unusual reaction in which the polymerase plays several distinctive roles, the most remarkable of which is to serve as the primer for reverse transcription (as well as the catalyst of chain elongation). As a result, the product minus-strand DNA remains covalently linked to the P-protein chain (35). Recent genetic experiments strongly suggest that this protein-primed reaction is also templated within the ε RNA stem-loop structure (33, 36).

These genetic data are compatible with the simple model that P protein is capable of recognizing ε. Here we demonstrate that DHBV polymerase in fact binds the DHBV encapsidation signal specifically and with high affinity. Mutational analysis of the ε stem-loop reveals that (i) lesions that ablate P protein-RNA binding block both RNA encapsidation and DNA priming and (ii) all mutations that inactivate RNA packaging likewise block DNA priming. These data strongly suggest that this RNA-protein recognition event is necessary to initiate both packaging and priming.

* Corresponding author. Phone: (415) 476-2826. Fax: (415) 476-0939.

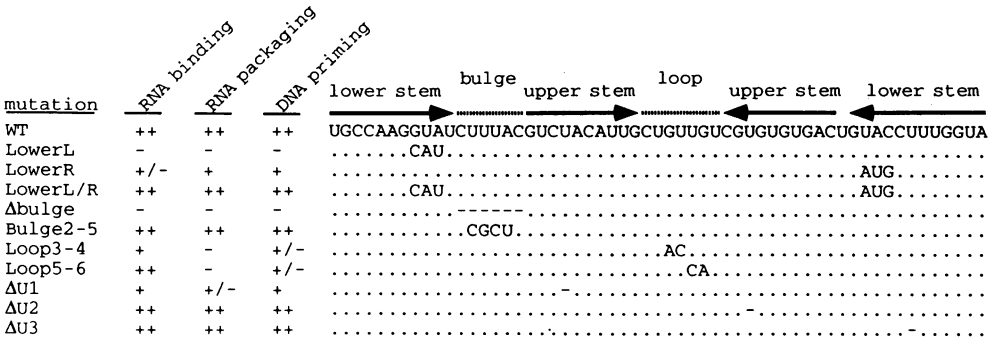

FIG. 1. DHBV ε stem-loop mutations. The primary sequence of the WT DHBV ε stem-loop region is shown at top, annotated above with arrows indicating the sequence components of the upper stem, lower stem, loop, and bulge. Subsequent rows show the mutant designations along with their nucleotide changes. ., no change of nucleotide; -, deletion of a nucleotide. The RNA binding, RNA packaging, and DNA priming efficiencies of mutants are indicated as follows: ++, 50 to 100% of WT levels; +, 10 to 50% of WT levels; +/−, 1 to 10% of WT levels; −, no detectable signal.

MATERIALS AND METHODS

Plasmid constructions. DHBV nucleotide positions are numbered from the unique *Eco*RI site of DHBV 3 (31). In this numbering scheme, nt 2530 is the transcription initiation site of pgRNA (5).

pT7dpol contains DHBV nt 170 to 3021 cloned into the *Sma*I site of the polylinker of pBS(−) (Stratagene). When linearized with *Afl*II and transcribed in vitro with T7 RNA polymerase, pT7dpol produces a transcript containing the DHBV polymerase coding region but no ε sequences.

pdεBS was constructed by cloning DHBV nt 2526 to 2845 into the *Hin*dIII and *Eco*RI sites of the polylinker of pBS(−). When linearized with *Eco*RV and transcribed in vitro with T3 RNA polymerase, pdεBS produces a 132-nt transcript containing DHBV ε. Stem-loop mutations (Fig. 1) were introduced into pdεBS by site-directed mutagenesis using standard methods (22) and confirmed by DNA sequencing. pE-BS was constructed as previously described (26) and produces a 172-nt transcript containing HBV ε, spanning HBV nt 1815 to 1986 (HBV adw2 [34]; pgRNA initiates at nt 1815).

Overlength stem-loop mutant genomes used for encapsidation assays were constructed in two steps. First, the *Afl*II-*Xba*I (nt 2526 to 2662) DNA fragment of D0.5G [nt 1658 to 3021 in pBS(−)] was replaced with the corresponding fragment (containing the appropriate mutation) of pdεBS. Second, an *Nsi*I (nt 2845) genome monomer was cloned into the *Nsi*I site of D0.5G in a head-to-tail orientation to produce an overlength D1.5G (nt 1658 to 3021/0 to 3021) mutant genome. The presence of the stem-loop mutation in final constructs was confirmed by DNA sequencing.

Polymerase mutant 442 (pol 442 [7]) contains a single nucleotide insertion resulting in the frameshifting and early termination of polymerase at amino acid (aa)442; this mutant is nonfunctional for RNA packaging and DNA synthesis (7, 17). Polymerase mutant YMHA (7) contains a double missense mutation in the conserved reverse transcriptase motif YMDD (altered to YMHA); this mutant packages RNA but is nonfunctional for DNA synthesis (7, 17).

pDHBV-RX was constructed by cloning DHBV nt 2662 to 3021 into the polylinker of pBS(−). When this construct is linearized with *Hin*dIII and transcribed in vitro with T7 RNA polymerase, a riboprobe complementary to DHBV pgRNA can be produced.

In vitro transcription. RNAs for in vitro translation, binding, and DNA primase assays were produced by using a MEGAscript kit (AMBION, Inc.) as instructed by the manufacturer. For production of biotinylated RNAs, 0.75 mM biotin-11-UTP (ENZO Diagnostics, Inc., New York, N.Y.), 1/10 of the nonbiotinylated UTP concentration, was added to the transcription reaction mixture.

In vitro translation and binding. RNA-protein binding reactions were performed cotranslationally. [^{35}S]methionine-labeled DHBV polymerase was translated from 1 μg of T7dpol RNA in a total volume of 25 μl, using nuclease-treated reticulocyte lysate as instructed by the manufacturer (Promega Biotec) in the presence of 100 ng of biotinylated RNA substrate at 30°C for 90 min. All translation/binding reactions were carried out in the presence of at least a 20-fold molar excess of tRNA (from the 50 mg of calf thymus tRNA per ml added to stimulate translation in the reticulocyte lysate, plus the endogenous reticulocyte tRNA pool).

Streptavidin precipitation of polymerase-RNA complexes. Following translation and binding, polymerase-RNA complexes were precipitated by using streptavidin-agarose as described by Scherly et al. (28). Five hundred microliters of Ipp150 (150 mM NaCl, 10 mM Tris-HCl [pH 7.5], 0.1% Nonidet P-40, 10 μg of yeast RNA per ml, 0.04 U of RNasin [Promega] per ml, 1 mM phenylmethylsulfonyl fluoride) was added to the completed translation reaction along with 25 μl packed streptavidin-agarose beads (Bethesda Research Laboratories). The solution was rocked at 4°C for 60 min, pelleted by a 15-s centrifugation, and washed three times with 500 μl of Ipp150. The pellet was resuspended in 50 to 100 μl of 2× sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) sample buffer, vortexed briefly, and boiled for 5 min. After a brief centrifugation, 10 μl of supernatant was loaded onto an SDS–8% polyacrylamide gel. After electrophoresis, the gel was treated with En^{3}Hance (DuPont), dried, and exposed to Kodak X-Omat film at −70°C overnight. Signal intensity was quantitated with a PhosphorImager (Molecular Dynamics).

In vitro DNA priming assays. DNA priming reactions were performed essentially as described by Wang and Seeger (35). Following in vitro translation and binding, 5 μl of the reaction mixture was removed and added to 5 μl of a mixture containing 100 mM Tris-HCl (pH 7.5), 30 mM NaCl, 20 mM MgCl, 28 μM each dATP, dCTP, and dTTP, and [^{32}P]dGTP (2.4 μM, 400 Ci/mmol). Reaction mixtures were incubated at 30°C for 30 min, and reactions were stopped by the addition of 9 volumes 2× SDS-PAGE sample buffer. Samples were then boiled for 3

min, and 10 μl was loaded on an SDS–10% polyacrylamide gel. After electrophoresis, the gel was dried and subjected to autoradiography (−70°C overnight) as described above.

Cell culture and transfections. LMH avian hepatoma cells were grown in H21-F12 (1:1) mix supplemented with 10% fetal calf serum and passaged every 2 to 3 days at a 1:3 dilution. DNA transfections were performed by the calcium phosphate coprecipitation method exactly as described previously (16).

RNA preparation and RNase protection assay. Poly(A)$^+$ total cellular RNA was purified 72 h posttransfection as previously described (16). RNA within cytoplasmic core particles was isolated 72 h posttransfection as previously described (23) by polyethylene glycol precipitation of core particles followed by proteinase K digestion, phenol extraction, and precipitation of nucleic acid. RNase protection analysis was performed as previously described (17) on poly(A)$^+$ RNA or core RNA prepared from equivalent numbers of transfected LMH cells (one-half of a 60-mm-diameter plate). Synthesis of [α-^{32}P]CTP-labeled RNA probes was carried out as previously described (17). Probe DHBV-RX is a 388-nt riboprobe with 27 nonhybridizing polylinker nt and 361 nt of DHBV complementary to pgRNA, spanning nt 2662 to 3021.

RESULTS

DHBV polymerase binds to the viral encapsidation signal. To demonstrate binding of DHBV polymerase to the encapsidation signal (ε), we performed a streptavidin-biotin-mediated coprecipitation assay (28). [^{35}S]methionine labeled DHBV polymerase was translated in rabbit reticulocyte lysate in the presence of synthetic DHBV ε RNA transcripts into which biotinylated uridine had been incorporated. P protein-RNA complexes were then detected by precipitation of biotinylated RNA substrates with streptavidin-linked agarose beads. The RNA target was present in the reaction mix during translation to allow for the possibility of cotranslational binding. (Pilot experiments verified that none of the RNA substrates had an inhibitory effect on polymerase translation in vitro [data not shown].) ^{35}S-labeled DHBV polymerase translated in reticulocyte lysate appears as an 80- to 85-kDa doublet (Fig. 2B, lane 2). The higher-molecular-weight species corresponds to the full-length 768-aa protein. The lower-molecular-weight species corresponds to in vitro translation initiation at a second in-frame methionine codon, 43 aa downstream (data not shown).

To demonstrate specific P protein-ε interaction, we performed a binding reaction with either a DHBV ε or an HBV ε RNA substrate. Although DHBV ε and HBV ε share phylogenetically conserved RNA structures (Fig. 2A) (19), DHBV and HBV polymerases in vivo direct encapsidation only of their own respective encapsidation signals (reference 26 and data not shown). DHBV polymerase was precipitated in the presence of biotinylated DHBV ε (Fig. 2B, lane 3) but not biotinylated HBV ε (lane 4) or nonbiotinylated DHBV ε (lane 5) RNA substrate. The faint background signal that often appears with negative-control RNA substrates is due to nonspecific binding of polymerase to the streptavidin-agarose beads (data not shown). Since the amount of polymerase protein loaded in lane 2 (Fig. 2B) represents 1/10 of the amount analyzed in each binding reaction, and since binding occurs in RNA substrate excess, we estimate that 5 to 10% of the total polymerase protein produced in vitro is functional for RNA binding. We have also found that polymerase-ε interaction will occur with the same efficiency and specificity when the RNA target is added posttranslationally in the presence of 0.5 mM cycloheximide (not shown).

To determine the dissociation constant of the DHBV polymerase-ε interaction, we performed an RNA titration (Fig. 2C, inset) using increasing amounts of DHBV ε RNA substrate. Polymerase binding was saturated at an RNA concentration of 177 nM, and half-maximal binding occurred at 14 nM RNA. Since a 10-fold reduction in the polymerase concentration did not significantly lower the RNA concentration at half-maximal binding (indicating that the polymerase concentration is significantly below the RNA concentration at half-maximal binding; data not shown), this value represents a reasonable estimate of the dissociation constant (K_d) for the P protein-ε interaction, assuming that all input RNAs are competent for binding. This value has ranged from 8 to 26 nM in replicate experiments.

Binding of polymerase requires the stem-loop structure. We next determined the RNA structure and sequence specificity of polymerase-ε interaction by assaying P-protein binding to a series of RNA stem-loop mutants (Fig. 3; summarized in Fig. 1).

To test the importance of the lower stem, we introduced three consecutive nucleotide changes on either the left (LowerL) or right (LowerR) side of this stem in order to disrupt potential base pairing. In the doubly mutated construct (LowerL/R), potential for base pairing in this region is restored. Polymerase did not bind the LowerL substrate (Fig. 3, lane 5), while introduction of the compensatory changes in the right side of the stem restored binding (LowerL/R; lane 7). A small amount of binding (ca. 10% of the wild-type [WT] level) was observed with the LowerR RNA substrate (lane 6). Closer examination of LowerR revealed that the nucleotide substitutions introduced to disrupt the lower stem can be accommodated in an alternative lower stem structure by base pairing to alternative partners in this region; such accommodation is not possible for LowerL (data not shown).

We next determined the consequences of mutating the bulge and loop for binding. Deletion of the 6-nt bulge (Δbulge; Fig. 3, lane 8) eliminated binding, while changing four consecutive residues in the bulge (Bulge2-5; lane 9) did not reduce binding. Altering residues 3 and 4 of the 7-nt loop (Loop3-4; lane 14) reduced polymerase binding approximately fourfold, while a mutant altering residues 5 and 6 of the loop (Loop5-6; lane 15) bound P protein with WT efficiency. In other RNA-protein interactions, single unpaired nucleotides often contribute to binding specificity (14). DHBV ε contains two unpaired U residues in the upper stem and one unpaired U residue in the lower stem. Deleting one of the unpaired U residues in the upper stem (ΔU1; lane 16) reduced binding four- to fivefold, while deleting either of the two other unpaired U residues (ΔU2 and ΔU3; lanes 17 and 18) had no effect.

ε mutations that eliminate P binding abolish RNA packaging. As RNA encapsidation in HBV requires both structural and sequence-specific contributions of the HBV encapsidation signal (26), we wished to determine ε structure and sequence requirements for DHBV RNA encapsidation. We therefore introduced the stem-loop mutations described above into the 5′ copy of ε in overlength DHBV genomes, which were then transfected into LMH avian hepatoma cells. Three days posttransfection, we harvested from equivalent numbers of cells either total cellular poly(A)$^+$ RNA or the RNA contained within purified cytoplasmic core particles. The RNA present in equal portions of each preparation was quantified by RNase protection using a riboprobe complementary to pgRNA sequences; the ratio of encapsidated to total poly(A)$^+$ RNA is a measure of the packaging efficiency.

The results of an encapsidation assay are shown in Fig. 4. The WT DHBV RNA pregenome was encapsidated, as a

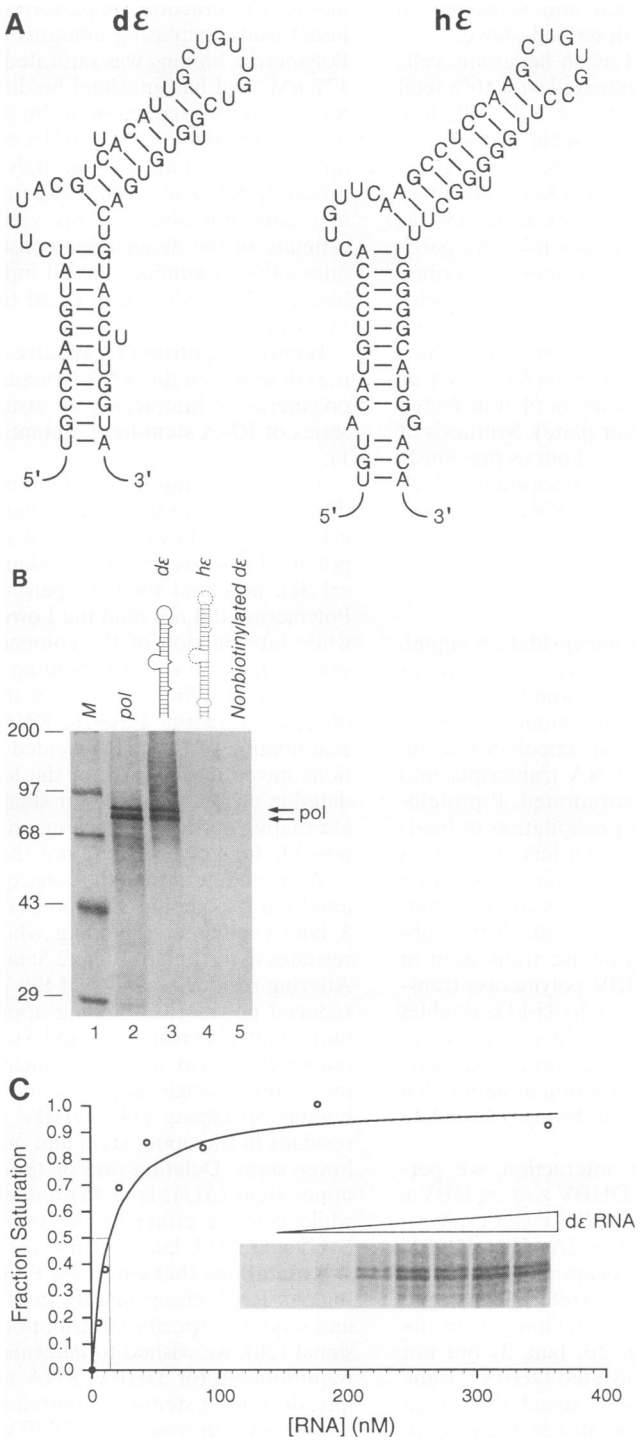

FIG. 2. DHBV polymerase-RNA binding assay. (A) DHBV (DHBV 3) ε (dε) and HBV (HBV adw2) ε (hε) RNA structures based on phylogenetic analysis and enzymatic probing (19, 21, 26). (B) ^{35}S-labeled DHBV polymerase protein was translated in reticulocyte lysate in the presence of substrate RNAs containing biotinylated U residues. Polymerase-RNA complexes were precipitated with streptavidin-agarose beads, released by boiling in SDS-PAGE sample buffer, and analyzed by SDS-PAGE. Lanes: 1, molecular size standards; 2, 1/10 of the total P protein analyzed per binding assay; 3 to 5, polymerase binding to DHBV ε (dε), HBV ε (hε), and nonbiotinylated DHBV ε, respectively. Sizes are indicated in kilodaltons. (C) RNA titration of polymerase binding with increasing amounts of biotinylated DHBV ε RNA, ranging from 0 to 355 nM. Fraction saturation is defined as the ratio of protein bound at each RNA concentration to the maximal protein bound. The curve drawn was determined by the best fit of the equation fraction saturation = [RNA]/(K_d + [RNA]), using the nonlinear least-squares method furnished in Sigmaplot (Jandel Scientific, Corte Madera, Calif.).

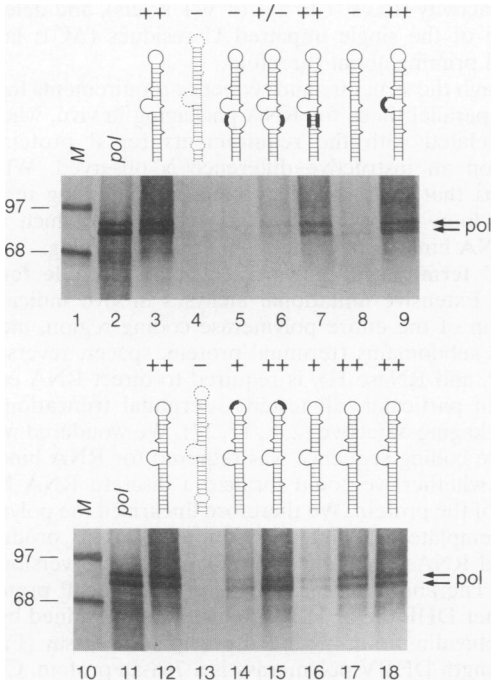

FIG. 3. RNA binding assay of DHBV ε stem-loop mutants. Steptavidin-biotin-based precipitation assays on mutant RNA substrates were performed as described in the legend to Fig. 2B. Lanes: 1 and 10, molecular size standards; 2 and 11, 1/10 of the total P protein analyzed per binding assay; 3 and 12, polymerase binding to DHBV ε; 4 and 13, polymerase binding to HBV ε; 5 to 9 and 14 to 18, polymerase binding to LowerL, LowerR, LowerL/R, Δbulge, Bulge2-5, Loop3-4, Loop5-6, ΔU1, ΔU2, and Δ3, respectively. The relative binding efficiency of each mutant is indicated as follows: ++, 50 to 100% of WT levels; +, 10 to 50% of WT levels; +/−, 1 to 10% of WT levels; −, binding level not above that of the negative control (HBV ε) RNA. Sizes are indicated in kilodaltons.

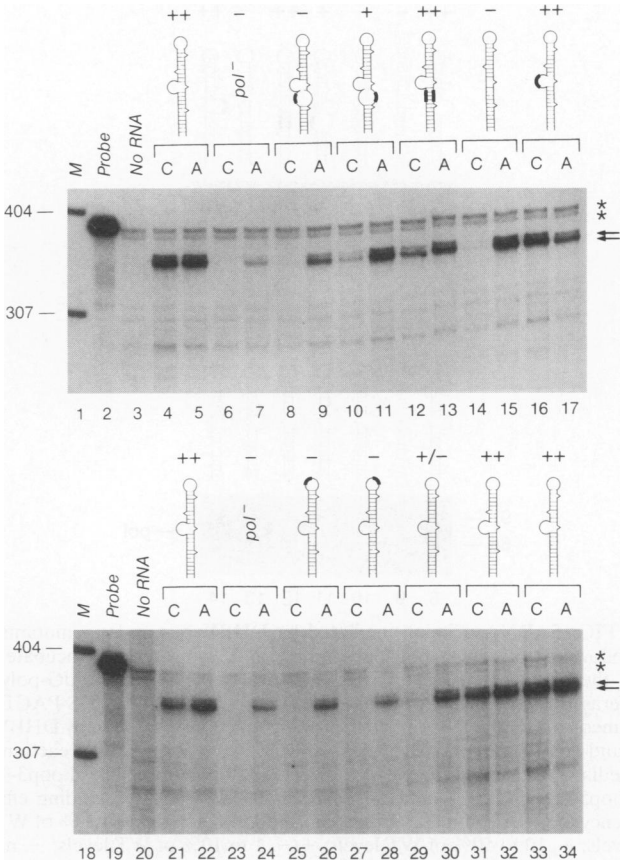

FIG. 4. Encapsidation assay of DHBV ε stem-loop mutants. From equivalent numbers of transfected cells, we isolated either total poly(A)$^+$ RNA or RNA extracted from cytoplasmic core particles. RNA from equal portions of each preparation was quantified by RNase protection using a DHBV antisense riboprobe that, when annealed to the complementary DHBV pgRNA sequence and digested with RNase, protects a 361-nt fragment (which appears as a doublet). Lanes 1 and 18, molecular size standards; lanes 2 and 19, undigested probe; lanes 3 and 20, probe digested in the absence of input RNA; lane pairs 4-5, 6-7, 8-9, 10-11, 12-13, 14-15, 16-17, 21-22, 23-24, 25-26, 27-28, 29-30, 31-32, and 33-34, core RNA (C)-poly(A)$^+$ RNA (A) from cells transfected with WT, pol442, LowerL, LowerR, LowerL/R, Δbulge, Bulge2-5, WT, pol442, Loop3-4, Loop5-6, ΔU1, ΔU2, and Δ3, respectively. Arrows indicate relevant protected fragments; asterisks indicate incompletely digested input probe. The encapsidation efficiency of each mutant is indicated as follows: ++, 50 to 100% of WT levels; +, 10 to 50% of WT levels; +/−, 1 to 10% of WT levels; −, no detectable encapsidated RNA. Sizes are indicated in nucleotides.

protected band of the expected size (361 nt) appears in sample lanes of both poly(A)$^+$ RNA (lane 5) and RNA isolated from core particles (lane 4). The efficiency of encapsidation (ca. 50%) was comparable to our previously reported results (17). As expected, the reaction is P-protein dependent: a viral genome carrying an early frameshift mutation in polymerase (pol442; lanes 6 and 7) was unable to encapsidate pgRNA.

For the most part, the packaging efficiency of our stem-loop mutants exhibited a strong correlation with the ability of these mutants to bind P protein. LowerL was not encapsidated (Fig. 4, lanes 8 and 9), while the compensatory mutations introduced in LowerL/R restored RNA encapsidation (lanes 12 and 13) to near WT levels; therefore, we conclude that the lower stem is required for encapsidation. LowerR (lanes 10 and 11) was encapsidated, although at fivefold-reduced levels, again the likely result of alternative base pairing which retains a lower stem structure (as discussed above). Deletion of the bulge (Δbulge; lanes 14 and 15) abolished RNA packaging, while alteration of residues within the bulge (Bulge2-5; lanes 16 and 17) was tolerated. Both mutants containing substituted loop residues (Loop3-4 [lanes 25 and 26] and Loop5-6 [lanes 27 and 28]) were not packaged, and of the single unpaired U residue deletion mutants, only ΔU1 (lanes 29 and 30) showed reduced packaging efficiency (ca. 5 to 10% of the WT level).

The features of ε important for DHBV RNA encapsidation are nearly identical to those previously reported for HBV encapsidation (26). The RNA structure and sequence requirements for DHBV packaging also closely parallel those for DHBV P protein-RNA binding, with the notable exception of the loop mutants. This is particularly evident in Loop5-6, which bound polymerase at WT levels but was not encapsidated (see Discussion).

ε mutations blocking polymerase binding also block DNA priming. It has been shown recently that P-protein-mediated priming of DNA synthesis occurs at RNA sequences within the encapsidation signal (33, 36). The sixth nucleotide of the bulge (C-2576) within the motif 5′-UUAC-3′ templates the covalent linkage of a dG to DHBV polymerase. Up to three more nucleotides are polymerized (pol-GTAA), followed by a trans-

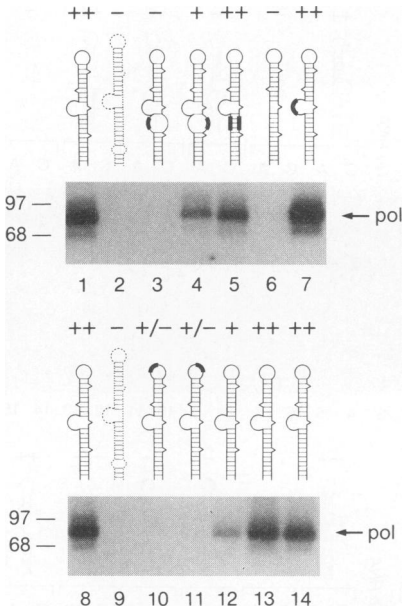

FIG. 5. DNA priming mediated by DHBV ε stem-loop mutants. Following translation, polymerase and RNA template were incubated in the presence of [^{32}P]dGTP. The covalently linked [^{32}P]dG-polymerase product of the priming reaction was analyzed by SDS-PAGE. Lanes 1 and 8 and lanes 2 and 9, priming reactions mediated by DHBV ε and HBV ε, respectively; lanes 3 to 7 and 10 to 14, priming reactions mediated by LowerL, LowerR, LowerL/R, Δbulge, Bulge2-5, Loop3-4, Loop5-6, ΔU1, ΔU2, and Δ3, respectively. The relative binding efficiency of DNA priming is indicated as follows: ++, 50 to 100% of WT levels; +, 10 to 50% of WT levels; +/−, 1 to 10% of WT levels; −, no detectable priming. Sizes are indicated in kilodaltons.

location to the 3′ end of pgRNA where DNA synthesis continues. The fact that DNA priming occurs at RNA sequences within the encapsidation signal suggests that polymerase might use a common RNA recognition event for both DNA priming and RNA encapsidation. We therefore tested the ability of our stem-loop mutants to mediate DNA priming in vitro. Following cotranslational binding of P protein to the mutant RNAs, a portion of the reaction mixture was removed and incubated with [α-^{32}P]dGTP; covalent [^{32}P]dG-polymerase linkage was then detected by SDS-PAGE (Fig. 5).

WT DHBV ε mediated DNA priming: an 80- to 85-kDa ^{32}P-labeled species (actually a doublet) corresponding to P protein appears in Fig. 5, lane 1. HBV ε did not support DNA priming by the DHBV polymerase (lane 2). ^{32}P linkage clearly results from the polymerase activity of P protein, since DHBV ε is unable to mediate DNA priming when the conserved reverse transcriptase motif YMDD of DHBV polymerase is mutated to YMHA (data not shown); this mutant is also nonfunctional for DNA synthesis in vivo (7).

The ability of stem-loop mutants to support DNA priming correlated exactly with their in vivo RNA packaging function. LowerL (Fig. 5, lane 3) was unable to mediate DNA priming, while this activity was restored in LowerL/R (lane 5). LowerR (lane 4) functioned in priming, though at threefold-reduced levels; again, this activity is the probable result of an alternate lower stem formation. The presence of the bulge (lane 6) but not its specific sequence (lane 7) was required for primase activity. (Note that we intentionally left intact residue C-2576 of the bulge, which templates [^{32}P]dG linkage.) Nucleotide substitutions in the loop (lanes 10 and 11) drastically reduced priming activity (to ca. 1 to 5% of WT levels), and deletion of only one of the single unpaired U residues (ΔU1; lane 12) impaired priming about threefold.

Although the structure and sequence requirements for DNA priming parallel those for RNA packaging in vivo, when they are correlated with the requirements for P protein-RNA interaction an instructive difference is observed. While all mutations that inactivate P protein-RNA binding inactivate priming, loop mutations (especially Loop5-6) which do not affect RNA binding nonetheless inactivate priming.

The C terminus of polymerase is dispensable for RNA binding. Extensive mutational analyses in vivo indicate that expression of the entire polymerase coding region, including all of its subdomains (terminal protein, spacer, reverse transcriptase, and RNase H), is required to direct RNA encapsidation; in particular, all tested C-terminal truncations have been packaging defective (2, 8, 17, 27). We wondered whether the entire coding sequence was required for RNA binding in vitro or whether we could localize a discrete RNA binding domain of the protein. We therefore linearized the polymerase cDNA template in different locations so as to produce 3′-truncated RNAs encoding C-terminally deleted versions of P protein. The ability of each of these truncated P proteins to bind either DHBV ε or HBV ε was then determined by using the streptavidin-biotin-mediated precipitation assay (Fig. 6).

Full-length DHBV polymerase is a 768-aa protein. C-terminal truncations of polymerase at aa 728 (Fig. 6, lanes 8 to 10) and aa 568 (lanes 12 to 14) were competent to bind DHBV ε. Truncations at aa 472 (lanes 15 to 17) and aa 371 (lanes 18 to 20) did not bind DHBV ε. Truncations at aa 728 and 568, but not at aa 472, are also competent to prime DNA synthesis (data not shown). Therefore, amino acid residues beyond position 568 appear dispensable for RNA binding and DNA priming.

The lower-migrating species of the doublet in each of these binding reactions represents in vitro translation initiation from the second methionine residue (aa 44) of polymerase. Because this species binds RNA and primes DNA synthesis, we conclude that the N-terminal 43 aa of polymerase are also dispensable for these activities.

Encapsidation in vivo is not dependent on the polymerase activity of P protein, since P protein containing mutations that alter the conserved reverse transcriptase motif YMDD to YMHA (aa 511 to 514) directs RNA encapsidation but not DNA synthesis in vivo (7, 17). In vitro, this same polymerase mutant was competent to bind RNA (Fig. 6, lanes 5 to 7) but does not prime DNA synthesis (data not shown). Thus, the polymerase activity of P protein is dispensable for both ε binding and RNA encapsidation.

DISCUSSION

The polymerase-RNA interaction. Here we have demonstrated that DHBV polymerase binds specifically and with high affinity ($K_d \cong 14$ nM) to DHBV ε. That DHBV polymerase does not bind the structurally similar HBV ε is a strong indication of the specificity of the interaction. Mutational analysis of the DHBV ε stem-loop reveals that the requirements for binding include the lower stem and bulge structures, a single unpaired U residue (U1), and to a lesser extent loop nt 3 and 4. Binding is not dependent on the polymerase activity of P protein. Since the ability of the polymerase to covalently bind a G residue templated by the bulge indicates an intimate interaction between the P polypeptide and the RNA, we strongly favor the view that P protein binds directly to ε. However, since all binding assays were carried out in a

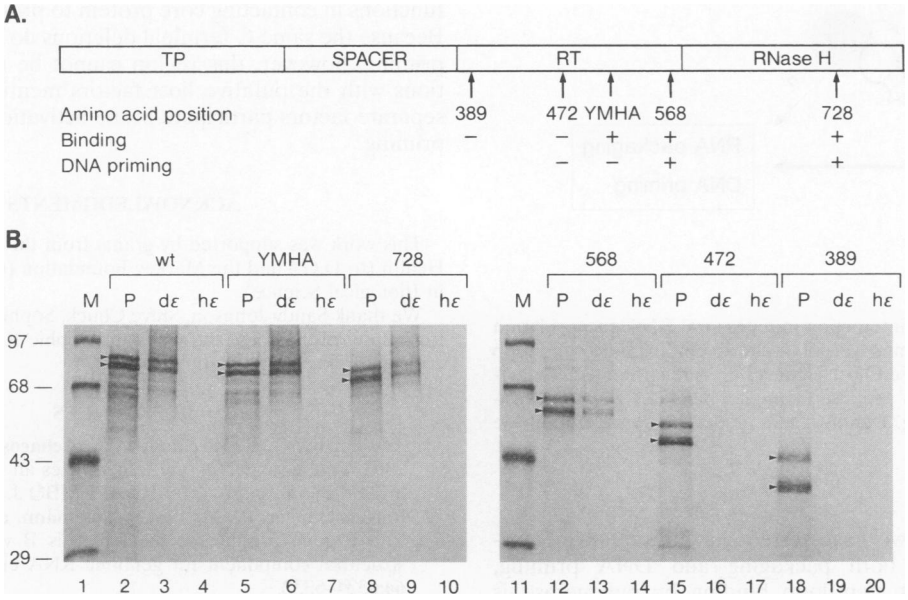

FIG. 6. RNA binding of C-terminally truncated polymerase proteins. (A) Schematic representation of the domain structure of the 786-aa DHBV polymerase protein. Polymerase protein was truncated at aa 728, 568, 472, and 389, as indicated by arrows. YMHA (aa 513/514) is a double missense mutation in the conserved reverse transcriptase motif YMDD. Below each truncation is indicated its ability to bind RNA and prime DNA synthesis. TP, terminal protein; RT, reverse transcriptase. (B) Truncation products were tested for binding to DHBV ε and HBV ε, using the streptavidin-biotin-based precipitation assay described in the legend to Fig. 2B. Lanes 1 and 11, molecular size standards; lane pairs 2-4, 5-7, 8-10, 12-14, 15-17, and 18-20, binding by WT polymerase, polYMHA, pol728, pol568, pol472, and pol389, respectively. P lanes indicate 1/10 of the total protein analyzed per binding reaction; dε and hε lanes indicate binding to DHBV ε and HBV ε, respectively. Major truncation products are indicated by arrowheads in each P lane. Sizes are indicated in kilodaltons.

reticulocyte lysate, we cannot formally exclude the possibility that the polymerase-ε interaction is indirect, i.e., that a cellular factor may participate in the binding. Experiments are under way to investigate this possibility. Since DHBV polymerase contains no currently known RNA binding motif (e.g., ribonucleoprotein motif, KH motif, RGG box, zinc finger, Arg-rich motif [25]), it is of great interest to characterize further the nature of this RNA-protein interaction.

Polymerase-RNA binding initiates both encapsidation and DNA synthesis. Mutational analysis of the ε stem-loop reveals that all lesions that ablate P protein-RNA binding block both RNA encapsidation and DNA priming and that all mutations that diminish or abolish RNA packaging have parallel effects on DNA priming. These data strongly suggest that the same RNA recognition event is necessary to initiate both reactions.

The precise molecular events that follow P-ε complex formation are unknown for both of these reactions. For RNA encapsidation, the simplest model would be that the P-RNA complex is recognized by assembling C-protein subunits via noncovalent interactions between C and P polypeptides. Such interactions could be direct or could be mediated by accessory host proteins. But in either case, such interactions must not be possible with free polymerase, since mutant viruses with ε deletions do not encapsidate P protein chains (3). (Another possibility is that P-ε interactions alter the ε RNA structure such that it can be directly recognized by C protein; however, the sequence nonspecificity of C protein-RNA interactions [4] makes such a model unlikely.) Irrespective of their molecular details, models that envision the P-ε complex as the target of the encapsidation machinery are satisfying because they solve the problem of ensuring the encapsidation of the polymerase. This solution is different from that employed by retroviruses, which resolve this dilemma by synthesizing their polymerases as Gag-Pol fusion proteins; these are then incorporated into the capsid through interactions between the Gag portion of the fusion and other Gag subunits.

In HBV, ε sequences alone are sufficient to mediate encapsidation. This is not the case in DHBV, in which a second RNA region some 900 nt downstream of ε is also required for RNA encapsidation (6, 18). A similar requirement for auxiliary encapsidation sequences has been reported for several retroviruses (24). How these downstream sequences are involved in the encapsidation process is uncertain. Clearly, they are not required for pgRNA recognition by polymerase, since DHBV ε alone can bind P protein. However, these downstream sequences may provide additional P-protein binding sites, be involved in the recruitment of core protein or cellular factors, or even participate in RNA-RNA interactions that generate structures required for efficient encapsidation.

The packaging process thus delivers to the virion a pgRNA molecule on which active polymerase is already bound to sequences known to be functional in initiation of reverse transcription. While it is clear that extensive DNA synthesis does not occur prior to encapsidation, it is not known whether this is also true for the synthesis of the 3- to 4-nt, P-linked DNA primer in vivo. Because in vitro DNA priming and synthesis can occur in the absence of core protein (35), these events could occur either prior to or concomitantly with RNA packaging in vivo. The *cis* preference that polymerase exhibits for encapsidating (and therefore replicating) the pgRNA from which it has been translated suggests that all of these events might in fact occur cotranslationally in vivo. However, in vitro we observe no strict requirement for cotranslational binding of P protein to ε sequences.

Polymerase-RNA interaction is necessary but not sufficient for RNA packaging and DNA priming. As discussed above, all

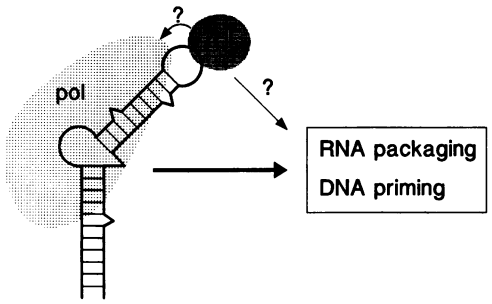

FIG. 7. Polymerase-ε interaction initiates viral RNA encapsidation and DNA synthesis. Polymerase (pol) is shown binding the ε stem-loop RNA structure. A proposed host factor (HF) is diagrammed contacting the loop region. This proposed complex mediates RNA encapsidation and DNA priming. Details of the model are considered in the text.

RNA mutations that we have tested that block polymerase-RNA binding block both packaging and DNA priming, strongly implying that stem-loop binding by polymerase is required for these reactions. However, the phenotype of DHBV loop mutations reveals that while P protein-RNA binding is necessary for these steps, it is not sufficient. The Loop3-4 and Loop5-6 mutants, which bind polymerase at 25 and 100% of WT levels, respectively, demonstrate no detectable RNA packaging activity and severely reduced (1 to 5% of WT levels) DNA priming. A critical role for loop nucleotides in HBV encapsidation has also been described (26).

Why might this be so? That loop mutants inhibit RNA packaging might seem at first to implicate a possible role of the loop in C-protein recognition. However, loop mutations are equally disruptive to DNA priming, which occurs independently of core protein. We therefore propose that some additional factor(s), perhaps of host origin, may be required to effect both the RNA packaging and DNA priming functions. It is attractive to speculate that such a factor(s) might contact the loop region (nt 3 to 6) of ε, thereby being brought into close proximity with bound P protein (Fig. 7). Such a putative factor(s) might then activate the RNA packaging and DNA priming activities of P protein, for example by inducing a conformational change or by catalyzing a covalent modification (or both). It is not unreasonable to expect that such factors could be present in the rabbit reticulocyte lysate; although hepadnaviral replication is highly species and tissue restricted in vivo, most of these restrictions can be bypassed if the promoter driving pgRNA expression, which is normally liver specific, is replaced by a more broadly active heterologous promoter (30). Thus, most cells appear to have whatever host machinery is required to support pgRNA packaging and reverse transcription.

The C terminus of polymerase is dispensable for RNA binding but not RNA packaging. By producing C-terminally truncated versions of polymerase in vitro, we have demonstrated that the C terminus of polymerase (beyond aa 568) is not required for RNA binding and DNA priming. In vivo, however, several deletion mutations and even single missense mutations introduced in the C terminus of polymerase (beyond aa 568) abolish RNA packaging (2, 8, 17, 27). It is a formal possibility that all of these mutations merely destabilize the P polypeptide in vivo. However, we favor a model in which the C terminus of polymerase, encompassing the RNase H domain of the protein, functions at a step in RNA encapsidation subsequent to RNA binding. Perhaps this region of polymerase functions in contacting core protein to promote viral assembly. Because the same C-terminal deletions do not inactivate DNA priming, however, this region cannot be involved in interactions with the putative host factors mentioned above, unless separate factors participate in the activation of packaging and priming.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health (to D.G.) and the Markey Foundation (to the UCSF Program in Biological Science).

We thank Sandy Johnson, Steve Chuck, Sophie Roy, and Ian Taylor for helpful discussions, and we thank Sophie Roy and Ian Taylor for critical reading of the manuscript.

REFERENCES

1. **Andino, R., G. E. Rieckhof, P. L. Achacoso, and D. Baltimore.** 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5′-end of viral RNA. EMBO J. **12**:3587–3598.
2. **Bartenschlager, R., M. Junker-Niepmann, and H. Schaller.** 1990. The P gene product of the hepatitis B virus is required as a structural component for genomic RNA encapsidation. J. Virol. **64**:5324–5332.
3. **Bartenschlager, R., and H. Schaller.** 1992. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. EMBO J. **11**:3413–3420.
4. **Birnbaum, F., and M. Nassal.** 1990. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. J. Virol. **64**:3319–3330.
5. **Buscher, M. W., W. Reiser, H. Will, and H. Schaller.** 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription. Cell **40**:717–724.
6. **Calvert, J., and J. Summers.** 1994. Two regions of an avian hepadnavirus RNA pregenome are required in *cis* for encapsidation. J. Virol. **68**:2084–2090.
7. **Chang, L.-J., R. Hirsch, D. Ganem, and H. E. Varmus.** 1990. Effects of insertional and point mutations on the functions of the duck hepatitis B virus polymerase. J. Virol. **64**:5553–5558.
8. **Chen, Y., W. S. Robinson, and P. L. Marion.** 1992. Naturally occurring point mutation in the C terminus of the polymerase gene prevents duck hepatitis B virus RNA packaging. J. Virol. **66**:1282–1287.
9. **Chiang, P.-W., K.-S. Jeng, C.-P. Hu, and C. Chang.** 1992. Characterization of a *cis* element required for packaging and replication of the human hepatitis B virus. Virology **186**:701–711.
10. **Cullen, B. R., and M. H. Malim.** 1991. The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional regulators. Trends Biochem. Sci. **16**:346–350.
11. **Das, A.** 1993. Control of transcription by RNA-binding proteins. Annu. Rev. Biochem. **62**:893–930.
12. **Dreyfuss, G., M. J. Matunis, S. Pinol-Roma, and C. G. Burd.** 1993. hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. **62**:289–321.
13. **Enders, G. H., D. Ganem, and H. E. Varmus.** 1987. 5′-Terminal sequences influence the segregation of ground squirrel hepatitis virus RNAs into polyribosomes and viral core particles. J. Virol. **61**:35–41.
14. **Frankel, A. D., I. W. Mattaj, and D. C. Rio.** 1991. RNA-protein interactions. Cell **67**:1041–1046.
15. **Hellen, C. U., G. W. Witherell, M. Schmid, S. H. Shin, T. V. Pestova, A. Gil, and E. Wimmer.** 1993. A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. Proc. Natl. Acad. Sci. USA **90**:7642–7646.
16. **Hirsch, R., R. Colgrove, and D. Ganem.** 1988. Replication of duck hepatitis B virus in two differentiated human hepatoma cell lines after transfection with cloned viral DNA. Virology **167**:136–142.
17. **Hirsch, R., J. Lavine, L. Chang, H. 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.

18. **Hirsch, R., D. D. Loeb, J. R. Pollack, and D. Ganem.** 1991. *cis*-acting sequences required for encapsidation of duck hepatitis B virus pregenomic RNA. J. Virol. **65:**3309–3316.
19. **Junker-Niepmann, M., R. Bartenschlager, and H. Schaller.** 1990. A short *cis*-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. EMBO J. **9:**3389–3396.
20. **Katze, M. G., M. Wambach, M.-L. Wong, M. Garfinkel, E. Meurs, K. Chong, B. R. G. Williams, A. G. Hovanessian, and G. N. Barber.** 1991. Functional expression and RNA binding analysis of the interferon-induced, double-stranded RNA-activated, 68,000-M_r protein kinase in a cell-free system. Mol. Cell. Biol. **11:**5497–5505.
21. **Knaus, T., and M. Nassal.** 1993. The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop structure that is critical for its function. Nucleic Acids Res. **21:**3967–3975.
22. **Kunkel, T. A., J. D. Roberts, and R. A. Zabour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**376–382.
23. **Lavine, J., R. Hirsch, and D. Ganem.** 1989. A system for studying the selective encapsidation of hepadnaviral RNA. J. Virol. **63:** 4257–4263.
24. **Linial, M. L., and A. D. Miller.** 1990. Retroviral RNA packaging: sequence requirements and implications. Curr. Top. Microbiol. Immunol. **157:**125–152.
25. **Mattaj, I. W.** 1993. RNA recognition: a family matter? Cell **73:**837–840.
26. **Pollack, J. R., and D. Ganem.** 1993. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. J. Virol. **67:**3254–3263.
27. **Roychoudhury, A. F. Faruqui, and C. Shih.** 1991. Pregenomic RNA encapsidation analysis of eleven missense and nonsense polymerase mutants of human hepatitis B virus. J. Virol. **65:**3617–3624.
28. **Scherly, D., W. Boelens, W. J. van Venrooij, N. A. Dathan, J. Hamm, and I. W. Mattaj.** 1989. Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. EMBO J. **8:**4163–4170.
29. **Schwartz, S. P., L. Aisenthal, Z. Elisha, F. Oberman, and J. K. Yisraeli.** 1992. A 69-kDa RNA-binding protein from Xenopus oocytes recognizes a common motif in two vegetally localized maternal mRNAs. Proc. Natl. Acad. Sci. USA **89:**11895–11899.
30. **Seeger, C., B. Baldwin, and B. C. Tennant.** 1989. Expression of infectious woodchuck hepatitis virus in murine and avian fibroblasts. J. Virol. **63:**4665–4669.
31. **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.
32. **Tang, C. K., and D. E. Draper.** 1989. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. Cell **57:**531–536.
33. **Tavis, J., S. Perri, and D. Ganem.** 1994. Hepadnavirus reverse transcription initiates within the stem-loop RNA packaging signal and employs a novel strand transfer. J. Virol. **68:**3536–3543.
34. **Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter.** 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. ICN/UCLA Symp. Mol. Cell. Biol. **18:**57–70.
35. **Wang, G.-H., and C. Seeger.** 1992. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. Cell **71:**663–670.
36. **Wang, G.-W., and C. Seeger.** 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J. Virol. **67:**6507–6512.