JOURNAL OF VIROLOGY, Oct. 1996, p. 6803–6809 Vol. 70, No. 10
0022-538X/96/$04.00+0

A Host Factor That Binds near the Termini of Hepatitis B Virus Pregenomic RNA

SILVIA PERRI[1] AND DON GANEM[1,2,3]*

Departments of Microbiology and Immunology[1] and Medicine[2] and Howard Hughes Medical Institute,[3] University of California, San Francisco, California 94143-0414

Received 15 April 1996/Accepted 24 June 1996

The terminal regions of hepatitis B virus (HBV) pregenomic RNA (pgRNA) harbors sites governing many essential functions in the viral life cycle, including polyadenylation, translation, RNA encapsidation, and DNA synthesis. We have examined the binding of host proteins to a 170-nucleotide region from the 5′ end of HBV pgRNA; a large portion of this region is duplicated at the 3′ end of this terminally redundant RNA. By UV cross-linking labeled RNA to HepG2 cell extracts, we have identified a 65-kDa factor (p65) of nuclear origin which can specifically bind to this region. Two discrete binding sites were identified within this region; in vitro cross-competition experiments suggest that the same factor binds to both elements. One binding site (termed UBS) overlaps a portion of the highly conserved stem-loop structure (ε), while the other site (termed DBS) maps 35 nucleotides downstream of the hexanucleotide polyadenylation sequence. Both binding sites are highly pyrimidine rich and map to regions previously found to be important in the regulation of viral polyadenylation. However, functional analysis of mutant binding sites in vivo indicates that p65 is not involved in the polyadenylation of HBV pgRNA. Potential roles for the factor in viral replication in vivo are discussed.

Hepadnaviruses are small hepatotropic viruses that produce persistent infection of liver cells in a variety of hosts and can cause hepatocellular carcinoma (for a recent review, see reference 14). Upon entering the cell, hepadnaviral capsids are transported to the nucleus, where their 3- to 3.3-kb partially duplex DNA genomes are converted into covalently closed circular molecules (24). The covalently closed circular DNA is used for transcription of subgenomic and pregenomic RNAs (pgRNAs) by host RNA polymerase II. The terminally redundant (2, 12, 13) pgRNA serves two functions: it is used for translation of the viral polymerase (P) and core (C) proteins (3, 4, 31), and it also serves as the template for viral DNA synthesis (36).

The terminal redundancy in pgRNA harbors sequences that are essential for several aspects of the virus life cycle, including polyadenylation, translation, RNA packaging, and DNA synthesis (see Fig. 7). The 5′ copies of these sequences play key roles in viral RNA packaging and translation, while important contributions to viral polyadenylation and reverse transcription are made by sequences in (and adjacent to) both terminal redundancies. Specifically, the 5′ copy of the redundancy contains a highly conserved stem-loop (ε) structure involved in RNA packaging (18, 20, 21, 26) and DNA priming (37, 40); the viral polymerase is required for both processes (1, 17). In duck hepatitis B virus (DHBV), P protein has been shown to bind specifically to this structure in vitro (27, 39). DNA priming occurs at a bulge in the stem-loop, thus generating a P-linked oligonucleotide that is approximately 4 nucleotides (nt) long. This short oligonucleotide is then transferred to a complementary sequence within the 3′ copy of the redundancy, whereupon elongation proceeds (37, 40). Most ε mutations that affect P-ε binding also inhibit both RNA packaging and DNA priming. Some mutations in the apical loop of ε, however, do not significantly alter the P-RNA interaction but do affect RNA packaging and reverse transcription (27). This finding suggests that additional factors, perhaps of host origin, may be involved in RNA encapsidation and DNA priming.

The terminal redundancy also contains the hexanucleotide polyadenylation signal [poly(A) signal] that, in its 3′ representation, allows polyadenylation of the pgRNA. This signal lies downstream from the ε structure and differs from the canonical signal (AAUAAA) by one base (UAUAAA) (33). This variant is known to function poorly; hepadnaviruses possess additional upstream elements (processing signals [PS]) to increase the efficiency of polyadenylation (28), and one of these signals (PS2) is contained within the terminal redundancy. Deletions of the 3′ copy of PS2 in ground squirrel hepatitis B virus (GSHV) significantly lower the polyadenylation efficiency of viral RNA (7, 28). Moreover, downstream of the cleavage site for polyadenylation (and hence just outside of the terminal redundancy) are additional U-rich sequences that are essential for the correct and efficient polyadenylation of human hepatitis B virus (HBV) RNA (33).

Because of the importance of the terminal regions of pgRNA in these many processes, and because of genetic evidence suggesting that host factors may contribute to some of them (27, 28), we looked for host factors able to recognize viral RNA sequences in and around the termini. Here we (i) describe the identification of a 65-kDa host factor that can specifically recognize this region and (ii) characterize its binding sites on the viral genome. Multiple binding sites for the factor are present in pgRNA and map to regions thought to be important in polyadenylation in the related hepadnavirus GSHV. However, mutational analysis suggests that the protein plays no role in polyadenylation, and mapping of its binding sites shows that most fall outside of elements known to function in RNA packaging (26). Models for its potential roles in viral replication are discussed.

MATERIALS AND METHODS

Materials. Restriction enzymes were purchased from New England Biolabs. T3 and T7 polymerases were purchased from Promega. *Taq*I polymerase was purchased from Perkin-Elmer. All of the enzymes and kits were used according to the manufacturers' instructions.

Plasmids and RNA transcripts. The pHPg series constructs were used to

* Corresponding author. Phone: (415) 476-2826. Fax: (415) 476-0939. Electronic mail address: ganem@socrates.ucsf.edu.

transcribe different portions of the HBV pgRNA. In these constructs, the numbering corresponds to the tract of the pgRNA transcribed, where +1 is nt 1815 from the unique site *Eco*RI of HBVadw2 (38). pHPg1-172 is pε-BS, pHPg27-115 is pεc, and pHPg1-92 is pεf described by Pollack and Ganem (26). The corresponding RNAs were transcribed by using the T3 promoter after linearization with *Sma*I. pHPg91-172, pHPg91-172/UAGAAA, pHPg1-92/LowerL, pHPg1-43, pHPg1-30, pHPg1-36, and pHPg19-43 were obtained by cloning the correponding DNA region after PCR amplification from pHPg1-172 with the exception of pHPg1-92/LowerL. In this case, the DNA template used was pε-LowerL (26). The oligonucleotides used for the PCR amplification contained restriction sites to allow the cloning into either *Bam*HI-*Hin*dIII or *Bam*HI-*Kpn*I of the pBS(−) vector (Stratagene). For pHPg91-172/UAGAAA, the 5′ oligonucleotide used for in the PCR amplification contained one mismatch compared with the HBV sequence to introduce one base pair change in the poly(A) hexanucleotide from TATAAA to TAGAAA. The corresponding RNAs were transcribed from the T3 promoter after linearization with either *Eco*RI or *Sma*I. To construct pHPg91-116 and pHPg116-172, the *Hin*dIII-*Bam*HI fragment containing nt 91 to 172 of HBV pgRNA in pHPg91-172 was gel purified and then *Alu*I digested. The two products of this digestion were cloned into either *Hind*III-*Sma*I of pBS(−), generating pHPg91-116, or into *Bam*HI-*Sma*I of pBS(−), generating pHPg116-172. The RNA from pHPg91-116 was transcribed from the T3 promoter after linearization with *Eco*RI. The RNA from pHPg116-172 was transcribed from the T7 promoter after linearization with *Sal*I. HPg116-163, HPg116-146, HPg126-172, and HPg141-172 were obtained by cloning the correponding DNA region after PCR amplification from pHPg116-172 into *Bam*HI-*Kpn*I of pBS(−). Transcription proceeded from the T7 promoter after linearization with *Bam*HI.

Upstream binding site (UBS) constructs UBS1, UBS2, UBS3, and UBS4 were made by PCR amplification using pHPg1-43 as the DNA template. The 5′ oligonucleotide was complementary to the vector, and the 3′ oligonucleotides contained 5-bp mismatches to the wild-type HBV sequence. These mismatches were used to introduce the following mutations: nt 19 to 23, CTAAT to GATTA (UBS1); nt 24 to 28, CATCT to GTAGA (UBS2); nt 29 to 33, CTTGT to GAACA (UBS3); and nt 34 to 38, CACTG to TGTAC (UBS4).

Downstream binding site (DBS) constructs DBS1 and DBS2 were made by PCR amplification using pHPg141-172 as the DNA template. The 3′ oligonucleotide was complementary to the vector, and the 5′ oligonucleotides introduced mutations at nt 158 to 162 (CTTCC to GAAGG; DBS2) and at nt 153 to 157 (CTTTC to GAAAG; DBS2). DBS3 and DBS4 were constructed by using the same DNA template. However, the 5′ oligonucleotide was complementary to the vector, and the 3′ oligonucleotides introduced the mutations at nt 152 to 148 (GACTT to CTGAA; DBS3) and at nt 143 to 147 (CTTCt to GAAGA; DBS4).

The pGPg series constructs were used to transcribe different portions of the GSHV pgRNA. In these constructs, the numbering corresponds to the tract of the pgRNA, where nt +1 corresponds to nt 1 in the sequence determined by Seeger et al. (31a). pGPg1-177 is pG107T described by Cherrington et al. (7). The corresponding RNA was transcribed by using the SP6 promoter after linearization with *Eco*RI. pGPg1-102 was obtained by digesting the gel-purified 200-bp *Hin*dII-*Eco*RI fragment from GPg1-177 with *Bst*Y and cloning it into *Sma*I-*Eco*RI of pBS(−). The corresponding RNA was transcribed from the T3 promoter after linearization with *Eco*RI. pGpg112-177, pGPg120-160, pGPg130-177, and pGPg130-160 were obtained by cloning the corresponding DNA region after PCR amplification from pGPg1-177. The oligonucleotides used contained the restriction sites to allow the cloning into *Hin*dIII-*Eco*RI of the pBS(−) vector. The corresponding RNA were transcribed from the T3 promoter after linearization with *Eco*RI.

pBS(−)T3 control transcript was obtained by using the vector pBS(−) as the template after digestion with *Pvu*II. Transcription from the T3 promoter generates a RNA approximately 150 nt long.

The RNAs were labeled with [α-^{32}P]UTP (3,000 Ci/mmol; Amersham) to a specific activity of 2.94×10^7 cpm/μg. After labeling, the RNAs were gel purified and eluted in 0.5 M ammonium acetate–0.1% sodium dodecyl sulfate (SDS) at 4°C overnight. The unlabeled RNAs were synthesized by using T7 and T3 Megascript kits (Ambion). All of the RNAs, after precipitation, were resuspended in 20 mM KCl, heated to 65°C, and allowed to slowly cool.

srcΔsHBV1636-2602 and srcΔsHBV1800-2400 were used to obtain srcΔsHBV1636-2400. srcΔsHBV1636-2600 was obtained by cloning *Alw*NI-*Apa*I of HBV DNA into *Bam*HI of srcΔs vector (6). Before cloning, *Alw*NI was blunted and ligated with *Bam*HI linkers. srcΔsHBV1800-2400 was obtained by cloning into *Bam*HI of srcΔs vector the PCR-amplified fragment of HBV between nt 1800 and 2400. Finally, since srcΔs has one *Bgl*II site just upstream of *Bam*HI, the 300-bp *Bgl*II fragment from srcΔsHBV1636-2602 was used to substitute for the 100-bp *Bgl*II fragment of srcΔsHBV1800-2400.

The constructs srcΔsHBV1636-2400/UBS2, srcΔsHBV1636-2400/DBS3, and srcΔsHBV1636-2400/UBS2-DBS3 were obtained by site-directed mutagenesis. We modified the vector pAlter (Promega) by replacing the *Pst*I-*Sph*I sites with *Bst*E-*Bgl*II-*Hpa*I sites generating pAlter(M). The 300-bp *Bgl*II fragment from srcΔsHBV1636-2400 was cloned into pAlter(M), and this construct was used to generate single-stranded DNA for site-directed mutagenesis according to the Promega protocol. The oligonucleotide CTTTTTCACCTCTGCCTAATGTAGACTTGTACATGTCCCAC was used to introduce UBS2 (underlined), and the oligonucleotide CTCGTTTTTGCCTTCTCTGAACTTTCCTTCCGTCAG to introduce DBS3 (underlined). For the double mutation, two rounds of site-directed mutagenesis were performed. The fragments containing the desired mutations were reintroduced into srcΔsHBV1636-2400.

All fragments that were either amplified by PCR or subjected to site-directed mutagenesis were fully sequenced to ensure the absence of additional mutations.

Protein extracts. Cytoplasmic extracts were prepared from 50% confluent HepG2 cells as described by Li and Kelly (22). Nuclear extracts from 80% confluent HepG2 and 2.2.15 (32) cells were prepared as described by Dignam et al. (10). The activity of the nuclear extracts was checked by an in vitro transcription assay dependent on RNA polymerase II. The template for transcription was a plasmid containing the adenovirus major late promoter linked to a 400-bp DNA fragment lacking guanosines previously described by Sawadogo and Roeder (30).

RNA-protein cross-linking. Seventeen micrograms of either nuclear or cytoplasmic extract was incubated with 4 ng of uniformly labeled RNA in the presence of an excess of tRNA (2 μg) in binding buffer (5 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES; pH 7.5], 5% glycerol, 25 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.05 mM EDTA). After 10 min of incubation at room temperature, the reaction mixtures were transferred to a microliter dish on ice and UV cross-linked for 30 min, using a UV-Stratalinker 2400 (Stratagene) at a distance from the UV lamps of approximately 5 cm. A mixture of RNases A (Sigma) and T_1 (Sigma) was added at final concentrations of 0.5 μg/μl and 0.5 U/μl, respectively, and each reaction mixture was then incubated for 30 min at 37°C. Finally, protein sample buffer (50 mM Tris [pH 6.8], 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue) was added, and the complexes were analyzed on a an SDS–12.5% polyacrylamide gel. Titration of increasing amounts of the labeled RNA on the 17 μg of nuclear extract revealed that in these binding conditions, there is an excess of p65 to labeled RNA. In the competition experiments, the nuclear extract was incubated as described above, with the additional presence of increasing amounts of unlabeled RNAs; the molar excess of the unlabeled RNAs to labeled RNAs varied in different experiments (see the figure legends).

Cell culture, transfections, and RNA analysis. HepG2 human hepatoma and 293T (11) cells were grown in Dulbecco's modified essential medium supplemented with 10% fetal calf serum. Transfections were performed by calcium phosphate precipitation as previously described (5). In each transfection, we used 7.5 μg of plasmid construct along with 3 μg of plasmid pXGH5 for the growth hormone assay (HGH-TGES Transient Gene Expression; Nichols Institute) used to determine the transfection efficiency. The poly(A)$^+$ mRNA was prepared 48 h posttransfection as previously described (16) and then analyzed by standard Northern (RNA) hybridization (29). The probe was prepared by random priming on a template of the gel-purified *Mlu*I-*Bgl*II fragment of *src* cDNA, using a Rediprime kit (Amersham). Prehybridization and hybridization were performed at 65°C in Church buffer (8).

RESULTS

A 65-kDa host factor that binds to the 5′ end of HBV pgRNA. To determine if any host factor was able to specifically bind to the terminal regions of HBV pgRNA, we prepared cytoplasmic and nuclear extracts from the human hepatoma cell line HepG2. This cell line was chosen because it is fully permissive for viral replication (32). Our first probe was a 172-nt in vitro transcript (HPg1-172) initiating near the cap site of pgRNA and including the entire 5′-terminal redundancy and 45 nt immediately downstream of it. The extracts were incubated with uniformly labeled HPg1-172 RNA at room temperature in the presence of an excess of nonviral RNA to minimize the background from nonspecific interactions. The protein-RNA complexes were UV cross-linked and, after extensive RNase digestion, analyzed by SDS-polyacrylamide gel electrophoresis (PAGE). This analysis revealed that at least three labeled species could be detected (Fig. 1A, lane 4). The faster-migrating band can be resolved in two bands corresponding to two factors with estimated molecular masses of 40.5 and 42.5 kDa. The more slowly migrating band has an estimated molecular mass of 65 kDa. All three species were sensitive to proteinase K digestion, confirming their polypeptide nature (not shown). None of these proteins formed complexes with a uniformly labeled control transcript from the T3 promoter of the pBS(−) vector (Fig. 1A, lane 2) even though this transcript contains a 25-nt region that is 80% U/C rich. These host factors are of nuclear origin: incubation of equivalent amounts of cytoplasmic extract with labeled HPg1-172 revealed only faint signals at the same molecular weight as for the nuclear extract (data not shown). It is possible that the faint

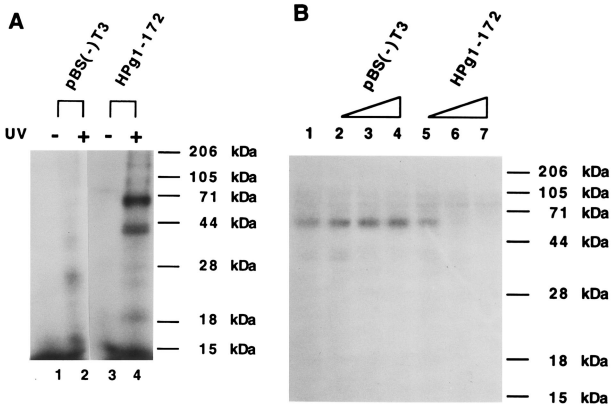

FIG. 1. Nuclear proteins bind to the 5′ 172 nt of HBV pgRNA. (A) SDS-PAGE analysis of nuclear extract from HepG2 cells incubated with either labeled RNA from pBS(−) (lanes 1 and 2) or labeled HPg1-172 (lanes 3 and 4) and then either stored on ice (lanes 1 and 3) or UV cross-linked (lanes 2 and 4). (B) Competition experiment with unlabeled RNAs. The nuclear extract was incubated and UV cross-linked with labeled HPg1-172 only (lane 1) or with that RNA in the presence of increasing amounts of unlabeled RNA from pBS(−) (lanes 2 to 4) or HPg1-172 (lanes 5 to 7). The weight ratios of unlabeled to labeled RNA were 40-fold in lanes 2 and 5, 200-fold in lanes 3 and 6, and 280-fold in lanes 4 and 7.

signals in the cytosolic extracts result from trace contamination by nuclear components.

To confirm the specificity of the interactions, the nuclear extract was incubated with labeled HPg1-172 and with increasing amounts of unlabeled specific (HPg1-172) or nonspecific [pBS(−) T3] competitor RNA. As shown in Fig. 1B, cross-linking of the 65-kDa species to labeled HPg1-172 was specifically competed for by HPg1-172 but not by the nonspecific competitor. These results indicate that this 65-kDa factor binds specifically to this region of HBV pgRNA; hereafter we shall refer to it as p65. In this experiment, the cross-linking of 40.5- and 42.5-kDa species to labeled HPg1-172 is very weak and appears to be abolished by both the specific and the nonspecific competitors. The intensity of the signal for these complexes, however, varied from experiment to experiment, suggesting either that they are degradation products of p65 or that the binding conditions used are not adequate to allow consistent detection of these interactions. In this work, we studied exclusively the more reproducible 65-kDa signal.

To determine whether active viral replication would cause a change in the pattern of proteins bound to this region, a nuclear extract was also prepared from 2.2.15, a HepG2 cell line that is stably transfected with HBV DNA and thus produces HBV (32). Incubation of this extract with labeled HPg1-172 gave the same UV cross-linking pattern as with the nuclear extract prepared from nontransfected HepG2 cells (not shown). Although HBV P protein is thought to interact with ε sequences in this region, we did not detect this interaction by UV cross-linking, a result which is likely attributable to the extremely low abundance of P protein in infected cells and/or to its preferential sequestration in *cis* by endogenous HBV RNA in the extract.

p65 binds to two discrete sites in the 5′ region of HBV pgRNA. It has previously been shown that some mutations in the apical loop of DHBV ε do not significantly affect the ability of P protein to bind ε. However, in both HBV and DHBV, RNA packaging levels are severely affected by such loop mutations (26, 27), suggesting that other factors, possibly of host origin, may be involved in these processes. To determine whether p65 might be one such factor, we tested the ability of the nuclear extract to bind to HPg1-172 sequences carrying the loop mutation Loop1-4, which changes 4 of the 6 nt in the apical loop and strongly impairs RNA encapsidation in vivo (26). As shown in Fig. 2A, p65 binds to this mutant RNA. This result was also confirmed by competition experiments in which the extract was incubated with labeled wild-type HPg1-172 in the presence of increasing amounts of unlabeled Loop1-4 RNA as the competitor. In these experiments, Loop1-4 competed efficiently with wild-type HPg1-172 for p65 binding (data not shown).

To determine the location of the p65 binding site(s) in this region, the extract was incubated with labeled RNAs corresponding to either the 5′ half of pgRNA1-172 (HPg1-92) or the 3′ half of this region (HPg91-172). The result was that p65 binds to both fragments, indicating the presence of at least two discrete binding sites in the 5′ 172 nt of HBV pgRNA (Fig. 2B). Cold competition experiments were performed to exclude the possibility that the RNA-protein complex detected at 65 kDa was actually a mixture of two different proteins binding to

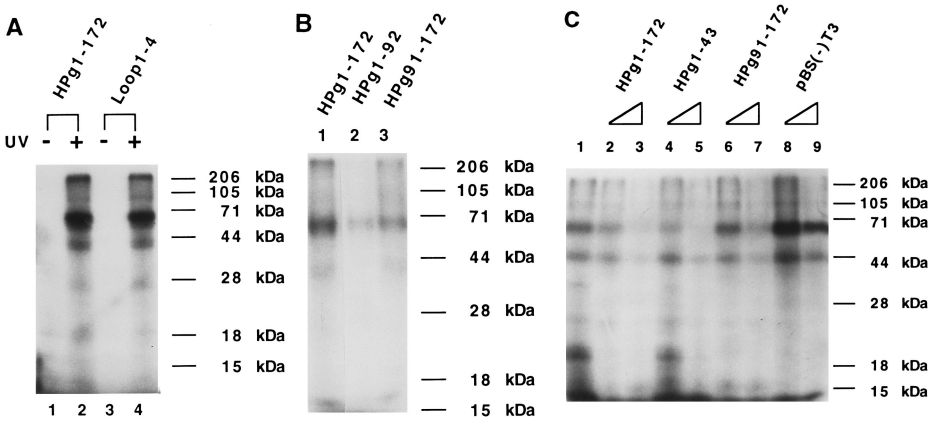

FIG. 2. p65 binds to two discrete binding sites. (A) SDS-PAGE analysis of nuclear extract incubated (all lanes) and UV cross-linked (lanes 2 and 4) with either labeled HPg1-172 (lanes 1 and 2) or labeled HPg1-172 containing loop mutation Loop1-4 (lanes 3 and 4). (B) SDS-PAGE analysis of nuclear extract incubated and UV cross-linked with labeled HPg1-172 (lane 1), labeled HPg1-92 (lane 2), or labeled HPg91-172 (lane 3). (C) Competition experiment with unlabeled RNAs. The nuclear extract was UV cross-linked to labeled HPg1-172 only (lane 1) or to this RNA in the presence of increasing amounts of unlabeled HPg1-172 (lanes 2 and 3), HPg1-43 (lanes 4 and 5), HPg91-173 (lanes 6 and 7), or RNA from pBS(−) (lanes 8 and 9). The molar excesses of unlabeled to labeled RNA were 50-fold in lanes 2, 4, 6, and 8 and 250-fold in lanes 3, 5, 7, and 9.

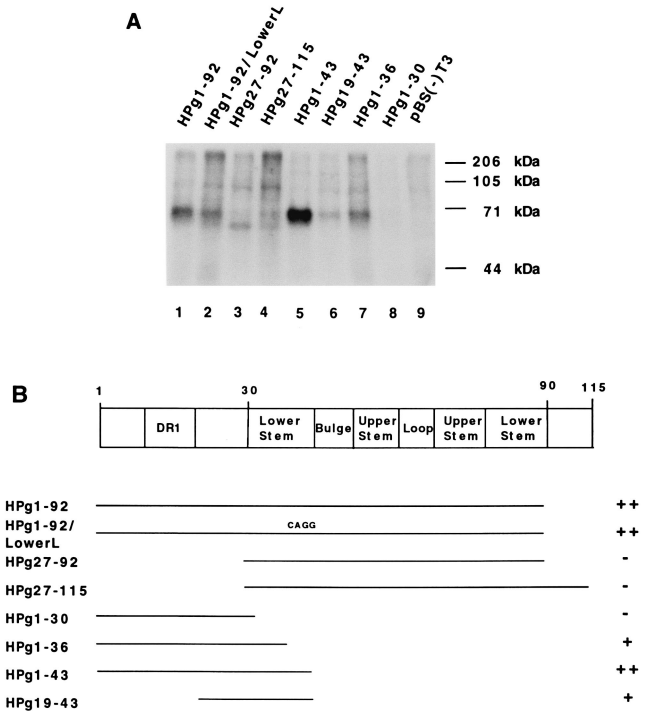

FIG. 3. Delimitation of UBS. (A) SDS-PAGE analysis of nuclear extract UV cross-linked to HPg1-92 (lane 1) and its derivatives containing deletions from the 5′ end (lane 3), from the 3′ end (lanes 5, 6, and 8), or from both ends (lane 6). In lane 2, the extract was UV cross-linked to HPg1-92 containing a mutation that destroyed the pairing in the lower stem of ε, HPg1-92/LowerL (26). In lane 3, the labeled RNA, HPg27-115, contained a 5′ deletion but included additional 3′ sequence containing the poly(A) signal. (B) Scheme of the RNAs tested for their ability to be recognized by p65 with respect to the indicated sequence landmarks present in the first 92 nt of HBV pgRNA. At the right is shown a scoring of the UV cross-linking results: ++, 25 to 100% of wild-type binding; +, binding reduced to 10 to 25% of the wild-type level; −, less than 5 to 10% of wild-type binding.

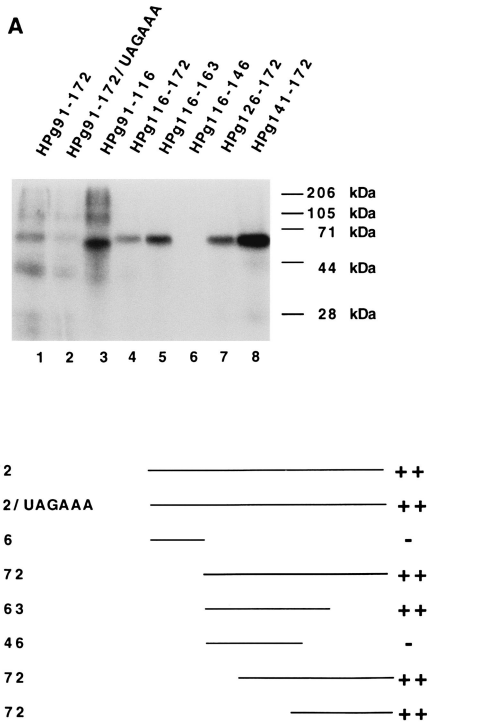

FIG. 4. Delimitation of DBS. (A) SDS-PAGE analysis of nuclear extract UV cross-linked to HPg91-172 (lane 1) and its derivatives containing deletions from the 5′ end (lanes 4, 7, and 8), from the 3′ end (lane 3), or from both ends (lanes 5 and 6). The RNA used in lane 2, HPg91-172, contained a mutated poly(A) site. (B) Scheme of the RNAs tested for the ability to be recognized by p65. At the right is a semiquantitiative scoring of the binding results as in Fig. 3B.

the two sites. Figure 2C shows the extract incubated with labeled full-length HPg1-172 in the presence of increasing amounts of unlabeled competitor RNAs corresponding to either the full-length RNA (HPg1-172 [lanes 2 and 3]), the upstream binding site UBS (HPg1-43 [lanes 4 and 5]), the downstream binding site DBS (HPg91-172 [lanes 6 and 7]), or a nonspecific RNA [pBS(−)T3 (lanes 8 and 9)]. Both UBS and DBS can fully compete with HPg1-172 for p65 binding. Furthermore, one aliquot of the extract was incubated with labeled UBS and one was incubated with labeled DBS, both in presence of increasing amounts of unlabeled UBS or DBS. In both cases, p65 binding to the labeled RNA was competed for by both unlabeled RNAs. These results strongly suggest that identical or closely related factors bind to both UBS and DBS.

To map each binding site more precisely, we constructed a series of plasmids that allowed expression of smaller portions of the 5′ region of HBV pgRNA. RNA transcribed in vitro from each of these constructs was then tested for its ability to be cross-linked to p65 in nuclear extracts. Figure 3 shows that UBS is fully contained within nt 1 to 43, with fragments as short as those spanning nt 19 to 43 still displaying detectable (albeit reduced) binding. UBS therefore overlaps a portion of one arm of the lower stem of ε, but clearly the ε secondary structure cannot be required for p65 binding. UBS also overlaps the HBV region corresponding to PS2, the sequence that was found to be important for efficient usage of the poly(A) signal in the closely related hepadnavirus GSHV (see Fig. 8) (7, 28). Figure 4 shows that the minimal region encompassing all DBS sequences competent to bind p65 falls between nt 141 and 163. This site is 35 nt downstream of the poly(A) signal and coincides with the U-rich region implicated in the polyadenylation of HBV RNAs (see Fig. 8) (33).

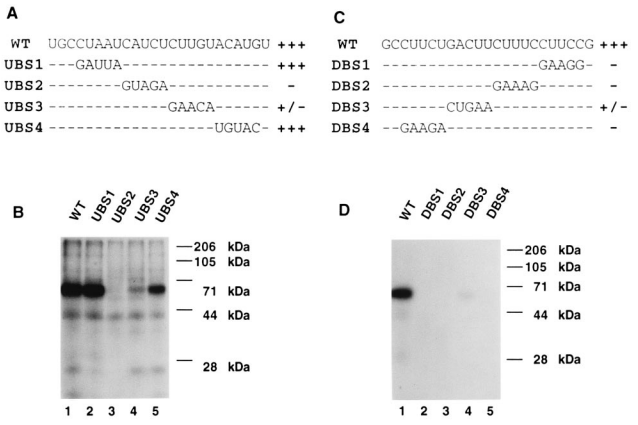

FIG. 5. Mutational analysis of p65 binding sites. (A) Sequence of the minimal region defining UBS and depiction of the nucleotide sequence changes introduced in each mutant RNA; at the right is a scoring of the binding assay as in Fig. 3B. WT, wild type. (B) SDS-PAGE analysis of nuclear extract UV cross-linked either to labeled HPg1-43 (lane 1) or to each of the indicated labeled mutant RNAs. (C) Sequence of the minimal region defining DBS and scheme of the changes made in each mutant RNA, with schematic scoring of the binding reaction at the right. (D) SDS-PAGE analysis of nuclear extract UV cross-linked either to labeled HPg141-172 (lane 1) or to each of the indicated labeled mutant RNAs.

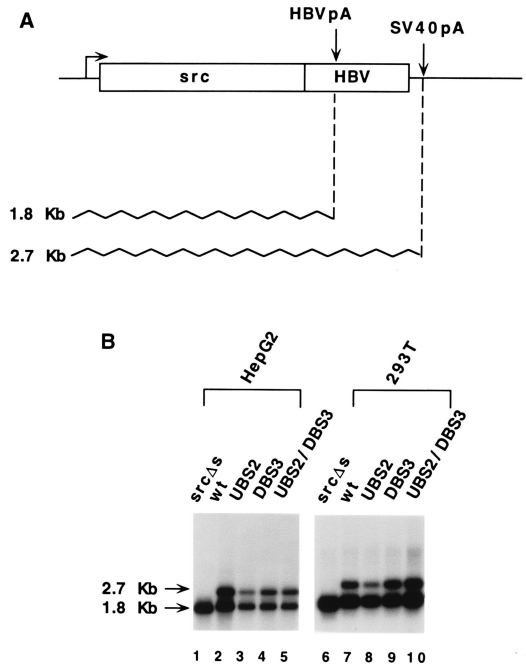

FIG. 6. Measurement of the efficiency of the HBV poly(A) site in the presence or absence of mutations interfering with p65 binding. (A) Scheme describing the tandem poly(A) site assay. The wavy lines represent the RNA products after polyadenylation at either the HBV poly(A) site (HBVpA) or the SV40 poly(A) site (SV40pA). (B) Northern analysis of poly(A)$^+$ RNAs from HepG2 cells (left) and 293T cells (right) transfected with vector alone (srcΔs: lanes 1 and 6), vector with wild-type (wt) HBV sequence between nt 1636 and 2400 (lanes 2 and 7), or the HBV sequence from nt 1636 to 2400 containing the UBS2 mutation (lanes 3 and 8), the DBS3 mutation (lanes 4 and 9), or the UBS2/DBS3 double mutation (lanes 5 and 10).

Characterization of p65 binding sites. Both UBS and DBS are about 20 nt long but do not show any significant primary sequence homology to one another (Fig. 5A and C). Both are, however, highly pyrimidine rich: UBS is 66% U/C rich, and DBS is 90% U/C rich. To characterize the two sites in more detail, blocks of 5 adjacent nt were converted to their complementary sequences in each binding site, and the resulting mutant RNAs were tested for p65 binding by UV cross-linking in nuclear extracts (Fig. 5). The mutations were examined in the context of the shortest RNAs that were efficiently bound by p65; accordingly, the UBS mutations were introduced into pHPg1-43, and the DBS mutations were introduced into pHPg141-172. Figure 5A shows that only the central core of UBS is critical for p65 binding (mutants UBS2 and UBS3). Strikingly, this central core is 80% U/C. In contrast, a somewhat larger region seems to be necessary for p65 binding to DBS, as all four clusters of mutations in DBS severely affected the binding (Fig. 5B). These results were confirmed by competition experiments in which the extract was incubated with either labeled HPg1-43 or labeled HPg141-172 in the presence of increasing amounts of unlabeled UBS mutants or DBS mutants, respectively. Only those mutants which scored as binding competent in Fig. 5 effectively competed with wild-type sequences for p65 binding under these conditions (data not shown).

p65 is not involved in HBV polyadenylation. As noted above, both UBS and DBS overlap regions that have been previously found to be important for polyadenylation of hepadnaviral RNAs. This finding suggested that p65 may play a role in this process. To test this possibility, we examined the effects of wild-type and mutant UBS and DBS sequences on polyadenylation in vivo, using a tandem poly(A) site assay previously characterized (6). In this assay (Fig. 6A), an HBV region (nt 1636 to 2400) containing all information required for efficient usage of the HBV poly(A) signal was cloned downstream a 1.6-kb c-*src* cDNA driven by a simian virus 40 (SV40) promoter. In this construct, 3′ of the HBV DNA there is an SV40 early poly(A) signal. When the HBV poly(A) signal is used, the resulting RNA is 1.8-kb long. When the HBV signal is not used, the RNA is polyadenylated at the SV40 signal, producing a larger RNA of 2.7 kb. The relative ratio of the two RNAs is a measure of the efficiency of the usage of the HBV poly(A) signal. For wild-type HBV sequences, 65% of the transcripts were processed at the HBV poly(A) site; the remaining 35% read through and were processed at the SV40 poly(A) site. Into this construct we then introduced the UBS2 mutation, the DBS3 mutation, or the UBS2/DBS3 double lesion and analyzed the efficiency of polyadenylation at the HBV site. Figure 6B shows that in HepG2 cells, mutations that abolish p65 binding at DBS have no effects on polyadenylation at the HBV poly(A) site (lane 4). Mutations abolishing binding to UBS slightly improve the utilization of the HBV site (lane 3). However, since this effect is small and is not conserved in the mutant with both binding sites altered (lane 5), the effects of UBS mutation on polyadenylation do not seem to be significant. Identical results were also obtained in a nonhepatic cell line, 293T (Fig. 6B, lanes 6 to 10).

p65 specifically binds to the termini of another hepadnaviral pgRNA. If p65 plays an important role in hepadnaviral infection, the binding sites for this factor would be expected to be conserved in other mammalian hepadnaviruses. To explore this possibility, we examined whether p65 would interact with the terminal region of the pgRNA of the related virus GSHV. Similarly to what we observed with the HBV pgRNA, a UV cross-linking assay on labeled 5′ 177 nt of GSHV pgRNA incubated with HepG2 nuclear extract revealed a 65-kDa complex (Fig. 7A, lane 1). Moreover, competition experiments in which the extract was incubated with labeled GSHV pgRNA (GPg1-177) in the presence of increasing amounts of unlabeled specific GSHV RNA (GPg1-177) or unlabeled nonspecific RNA [pBS(−)] showed that this complex results from a highly specific interaction, since only the GSHV RNA competed with the probe for p65 binding (Fig. 7A, lanes 1 to 5). To assess if this complex was formed by the same 65-kDa protein binding to HBV pgRNA, the extract was incubated with labeled GSHV pgRNA in the presence of unlabeled HBV RNA containing either UBS (HPg1-92) or DBS (HPg91-177). Both HBV RNAs efficiently competed with the GSHV RNA for p65 binding (Fig. 7A, lanes 1 and 6 to 9), indicating that the same factor specifically binds to both hepadnaviral pgRNAs.

p65, however, binds to only a single site on GSHV pgRNA terminal region. Incubation of the extract with the labeled 5′ half of nt 1 to 177 of GSHV pgRNA (GPg1-102) did not produce the UV cross-linked product of 65 kDa (Fig. 7B, lane 2). This result shows that GSHV does not contain a p65 binding site overlapping PS2 and corresponding to UBS of HBV; this finding is consistent with the data showing that p65 does not function in the polyadenylation processes. The 3′ half of the nt 1 to 177 region of GSHV (GPg112-177) was efficiently bound by p65 (Fig. 7B, lane 3). The binding site is located between nt 120 and 160, and significant binding activity is retained by GPg130-177 (data not shown). This p65 binding site on the GSHV pgRNA termini maps in roughly the same position as DBS of HBV. The sequence does not exhibit a high primary sequence homology to either HBV UBS or HBV DBS; nevertheless, like UBS and DBS, it is 83% pyrimidine rich.

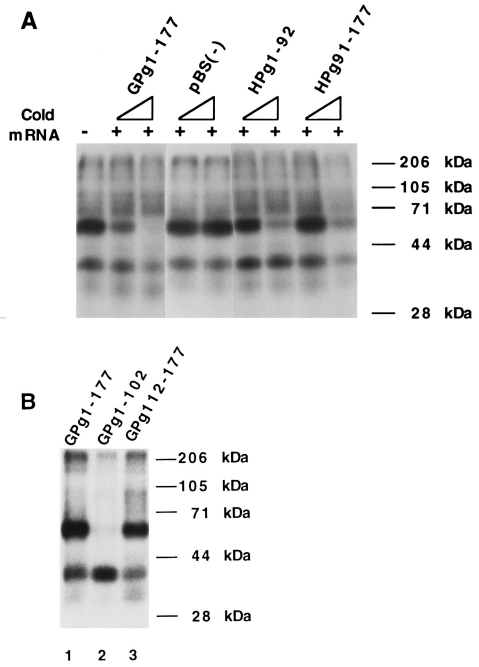

FIG. 7. p65 binds to the terminal region of GSHV pgRNA. (A) SDS-PAGE analysis of nuclear extract incubated with labeled GPg1-177 alone (lane 1) and in the presence of increasing amounts of unlabeled GPg1-177 (lanes 2 and 3), unlabeled RNA from pBS(−)T3 [pBS(−); lanes 4 and 5), and unlabeled HBV RNAs containing either UBS (HPg1-92; lanes 6 and 7) or DBS (HPg91-172; lanes 8 and 9). The molar excesses of unlabeled to labeled RNAs were 50-fold in lanes 2, 4, 6, and 8 and 250-fold in lanes 3, 5, 7, and 9. (B) SDS-PAGE analysis of nuclear extract incubated with labeled full-length GPg1-177 (lane 1), with labeled 5′ half of the RNA (GPg1-102; lane 2), or with labeled 3′ half of the RNA (GPg112-177; lane 3).

DISCUSSION

This study indicates that a 65-kDa host factor, present in the nuclei of uninfected cells, can interact with two sites (UBS and DBS) present in the first 170 nt of HBV pgRNA. Figure 8 shows a schematic depiction of the locations of UBS and DBS, with reference to several other well-known landmarks in this region of HBV or GSHV. Since a portion of this 170-nt sequence is also duplicated at the 3′ end of pgRNA, another copy of UBS is present at the opposite end of the transcript. The 3′ copy of DBS lies just downstream of the cleavage site for polyadenylation and hence is excluded from mature pgRNA, but it is included within the nuclear pre-mRNA prior to polyadenylation.

The binding sites for p65 are at least partially coextensive with elements known to function in both polyadenylation and encapsidation of pgRNA. The present data, however, indicate that p65 interaction with these sites probably plays no major role in either process. Explicit attempts to detect effects on polyadenylation by using a sensitive in vivo assay known to detect other regulators of hepadnaviral polyadenylation (7, 28) showed no effect of mutational ablation of UBS, DBS, or both (Fig. 6). Furthermore, GSHV pgRNA has not conserved the p65 binding site (UBS) overlapping the polyadenylation upstream element (PS2). As regards RNA packaging, DBS and the 3′ copy of UBS lie outside sequences previously shown to be sufficient to allow encapsidation of heterologous RNAs (26). The 5′ UBS element partially overlaps sequences involved in packaging, but it does not include the ε stem-loop structure known to be the principal determinant of this activity. Certainly p65 is not the putative host factor proposed to account for the behavior of ε mutations which block packaging without affecting P protein binding (27).

Many interesting and important potential functions for the p65-pgRNA interaction remain to be explored. Since p65 appears to be a nuclear protein, it could be involved in functions such as the regulation of transcriptional pausing or elongation, RNA stability, or the nuclear-cytoplasmic transport of pgRNA. Its mainly nuclear localization, though, does not exclude a role for this protein in processes that occur in the cytoplasm. The 5′ copy of UBS is in the 5′ nontranslated region of the RNA and could play a role in regulating the efficiency of translation initiation at the C gene AUG. The 5′ copy of DBS is within the C open reading frame, where it could affect translational elongation of C protein or even initiation of P protein biosynthesis. Certainly, p65 loaded onto the RNA in the nucleus could remain associated with the transcript in the cytosol to regulate translation.

Other important potential functions of p65-RNA interactions could involve viral DNA synthesis. Since p65 molecules can be bound to both ends of the RNA, it is tempting to speculate that they may be involved in bringing the two ends of the RNA template together, for example, by protein-protein interactions either with each other or with additional polypeptides. Such an activity could facilitate the transfer of nascent minus-strand DNA from the 5′ copy of ε to the 3′ copy of DR1. The binding of p65 to the 5′ copy of UBS, which is just downstream of DR1, could also play a role in the generation or translocation of the RNA primer for plus-strand DNA synthe-

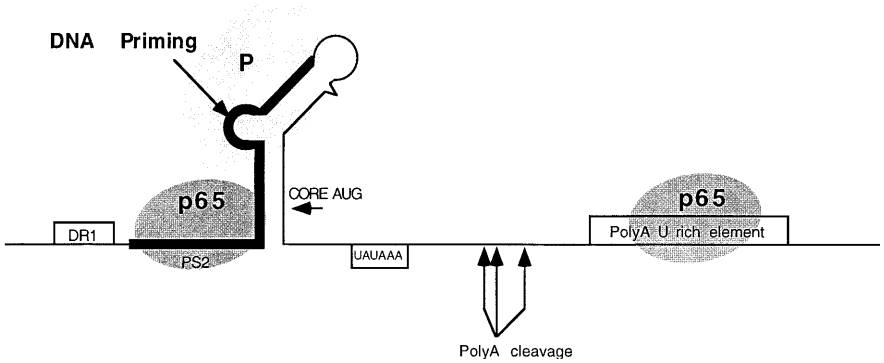

FIG. 8. Schematic depiction of landmarks within the 5′ 172 nt of HBV pgRNA. The locations of p65 binding sites are noted with respect to elements known to influence various functions in which this region plays an important role in several hepadnaviruses (see the text for more details). Black bar, PS2 polyadenylation-enhancing element; DR1, extent of direct repeat sequence involved in DNA strand transfers; UAUAAA, HBV polyadenylation signal; arrows, poly(A) addition sites; long box, U-rich region required for accurate and efficient poly(A) addition.

sis (23, 35). Since p65 appears to be a nuclear factor, and since viral DNA synthesis is largely cytoplasmic, any role for p65 in this process would most likely require that the factor be loaded onto the nuclear pre-RNA. There is ample precedent for important viral regulators interacting with nuclear representations of viral transcripts (for instance, *tat* and *rev* of human immunodeficiency virus [9, 15]). We are now carrying out a detailed analysis of the effects of UBS and DBS mutations on viral RNA synthesis and genomic replication in HBV-transfected HepG2 cells. Preliminary results reveal striking effects on viral DNA synthesis whose nature and extent depend on which p65 binding site is ablated. Such studies strongly suggest that p65 binding is not simply a curiosity of in vitro experimentation and indicate that p65 may play multiple roles in replication. The exact nature of these roles is under investigation.

We do not yet know the identity of p65 or understand its normal role in uninfected cells. Certainly, other host proteins are known to preferentially interact with pyrimidine-rich RNA sequences. Examples of these include PTB, which was initially defined as a splicing factor but also appears to play a role in translational control in some picornaviruses (19), and U2AF, a factor involved in RNA splicing (41). Both of these proteins share with p65 the ability to bind to multiple different U/C-rich primary sequences (34). Although PTB is described as a 57-kDa protein, U2AF was initially defined as a 65-kDa RNA-binding protein, making it a potential candidate for p65. Our preliminary evidence, however, suggests that p65 is not identical to U2AF. Using RNA mobility shift asays, we find that purified U2AF will, not surprisingly, bind to UBS and DBS, but this binding is not ablated by all the mutations that block UV cross-linking to p65 (25). Clearly, purification and characterization of p65 will be required to unambiguously determine its structure and clarify its function.

REFERENCES

1. **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 the genomic RNA encapsidation. J. Virol. **64:**5324–5332.
2. **Buscher, M., 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.
3. **Chang, L.-J., D. Ganem, and H. E. Varmus.** 1990. Mechanism of translation of the hepadnaviral polymerase (P) gene. Proc. Natl. Acad. Sci. USA **87:** 5158–5162.
4. **Chang, L.-J., P. Pryciak, D. Ganem, and H. E. Varmus.** 1989. Biosynthesis of the reverse transcriptase of the hepatitis B viruses involves de novo translational initiation not ribosomal frameshifting. Nature (London) **337:**364–368.
5. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7:**2745–2752.
6. **Cherrington, J., and D. Ganem.** 1992. Regulation of polyadenylation in human immunodeficiency virus (HIV): contributions of promoter proximity and upstream sequences. EMBO J. **11:**1513–1524.
7. **Cherrington, J., R. Russnak, and D. Ganem.** 1992. Upstream sequences and cap proximity in the regulation of polyadenylation in ground squirrel hepatitis virus. J. Virol. **66:**7589–7595.
8. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81:**1991.
9. **Constantoulakis, P., M. Campbell, B. K. Felber, G. Nasioulas, E. Afonina, and G. N. Pavlakis.** 1993. Inhibition of Rev-mediated expression by an RNA binding protein encoded by the interferon-inducible 9-27 gene. Science **259:**1314–1318.
10. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11:**1475–1489.
11. **DuBridge, R. B., P. Tang, H. C. Hsia, P.-M. Long, J. H. Miller, and M. P. Calos.** 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. **7:**379–387.
12. **Enders, G. H., D. Ganem, and H. E. Varmus.** 1985. Mapping the major transcripts of ground squirrel hepatitis virus: the presumptive template for reverse transcriptase is terminally redundant. Cell **42:**297–308.
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. **Ganem, D.** 1996. *Hepadnaviridae* and their replication, p. 2703–2737. *In* B. N. Fields, D. M. Knipe, P. M. Howley, et al., Fields virology, 3rd ed. Lippincott-Raven, Philadelphia.
15. **Gatignol, A., A. Bucler-White, B. Berkhout, and K.-T. Jeang.** 1991. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. Science **251:**1597–1600.
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. E. Varmus, and D. Ganem.** 1990. Polymerase gene products of hepatitis B viruses are required for genomic RNA packaging as well as for reverse transcription. Nature (London) **344:**552–555.
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. **Jang, S. K., and E. Wimmer.** 1990. Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. Genes Dev. **4:**1560–1572.
20. **Junker-Niepmann, M., R. Bartenschlager, and H. Schaller.** 1992. 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.
21. **Knauss, T., and M. Nassal.** 1993. The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop that is critical for its function. Nucleic Acids Res. **21:**3967–3975.
22. **Li, J. J., and T. J. Kelly.** 1984. Simian virus 40 DNA replication *in vitro*. Proc. Natl. Acad. Sci. USA **81:**6973–6977.
23. **Loeb, D., R. C. Hirsch, and D. Ganem.** 1991. Sequence-independent RNA cleavages generate the primers for plus-strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. EMBO J. **10:** 3533–3540.
24. **Mason, W. S., M. S. Halpern, J. M. England, G. Seal, J. Egan, L. Coates, C. Aldrich, and J. Summers.** 1983. Experimental transmission of duck hepatitis B virus. Virology **131:**375–384.
25. **Perri, S., R. Singh, M. Green, and D. Ganem.** Unpublished data.
26. **Pollack, J., and D. Ganem.** 1993. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. J. Virol. **67:**3254–3263.
27. **Pollack, J., and D. Ganem.** 1994. Site-specific RNA binding by a hepatitis B virus reverse transcriptase initiates two reactions: RNA packaging and DNA synthesis. J. Virol. **68:**5579–5587.
28. **Russnak, R., and D. Ganem.** 1990. Sequences 5′ to the polyA signal mediate differential polyA site use in hepatitis B. Genes Dev. **4:**764–776.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
30. **Sawadogo, M., and R. G. Roeder.** 1985. Factors involved in specific transcription by human RNA polymerase II: analysis by a rapid and quantitative in vitro assay. Proc. Natl. Acad. Sci. USA **82:**4394–4398.
31. **Schlicht, H., G. Radziwill, and H. Schaller.** 1989. Synthesis and encapsidation of duck hepatitis B virus reverse transcriptase do not require formation of core-polymerase fusions proteins. Cell **56:**85–92.

31a. **Seeger, C., D. Ganem, and H. E. Varmus.** 1984. Nucleotide sequence of an infectious molecularly cloned genome of ground squirrel hepatitis virus. J. Virol. **51:**367–375.

32. **Sells, M. A., M.-L. Chen, and G. Acs.** 1987. Production of hepatitis B virus particles in HepG2 cells transfected with cloned hepatitis B virus DNA. Proc. Natl. Acad. Sci. USA **84:**1005–1009.
33. **Simonsen, C. C., and A. D. Levinson.** 1983. Analysis of processing and polyadenylation signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. Mol. Cell. Biol. **3:**2250–2258.
34. **Singh, R., J. Valcarel, and M. Green.** 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science **268:**1173–1179.
35. **Staprans, S., D. Loeb, and D. Ganem.** 1991. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. J. Virol. **65:**1255–1262.
36. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of a hepatitis B like virus by reverse transcription of an RNA intermediate. Cell **29:**403–415.
37. **Tavis, J., S. Perri, and D. Ganem.** 1994. Hepadnaviral reverse transcription initiates within the RNA stem-loop of the viral encapsidation signal and employs a novel strand transfer. J. Virol. **68:**3536–3542.
38. **Valenzuela, P., M. Quiroga, J. Z. 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.
39. **Wang, G., F. Zoulim, E. Leber, J. Kitson, and C. Seeger.** 1994. Role of RNA in enzymatic activity of the reverse transcriptase of hepatitis B viruses. J. Virol. **68:**8437–8442.
40. **Wang, G.-H., and C. Seeger.** 1993. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. Cell **71:**663–670.
41. **Zamore, P. D., and M. Green.** 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc. Natl. Acad. Sci. USA **86:**9243–9247.