JOURNAL OF VIROLOGY, July 1997, p. 4971–4980 Vol. 71, No. 7
0022-538X/97/$04.00+0

Sequence- and Structure-Specific Determinants in the Interaction between the RNA Encapsidation Signal and Reverse Transcriptase of Avian Hepatitis B Viruses

JÜRGEN BECK AND MICHAEL NASSAL*

Zentrum für Molekulare Biologie, University of Heidelberg, D-69120 Heidelberg, Germany

Received 23 December 1996/Accepted 25 March 1997

Hepatitis B viruses (HBVs) replicate by reverse transcription of an RNA intermediate. Packaging of this RNA pregenome into nucleocapsids and replication initiation depend crucially on the interaction of the reverse transcriptase, P protein, with the *cis*-acting, 5′ end-proximal encapsidation signal ε. The overall secondary structure is similar in all of the hepadnaviral ε signals, with a lower and an upper stem, separated by a bulge, and an apical loop. However, while ε is almost perfectly conserved in all mammalian viruses, the ε signals of duck HBV (DHBV) and heron HBV (Dε and Hε, respectively) differ substantially in their upper stem regions, both in primary sequence and in secondary structure; nonetheless, Hε interacts productively with DHBV P protein, as shown by its ability to stimulate priming, i.e., the covalent attachment of a deoxynucleoside monophosphate to the protein. In this study, we extensively mutated the variable and the conserved positions in the upper stem of Dε and correlated the functional activities of the variant RNAs in a priming assay with secondary structure and physical P protein binding. These data revealed a proper overall structure, with the bulge and certain key residues, e.g., in the loop, being important constraints in protein binding. Many mutations at the evolutionarily variable positions complied with these criteria and yielded priming-competent RNAs. However, most mutants at the conserved positions outside the loop were defective in priming even though they had ε-like structures and bound to P protein; conversely, one point mutant in the loop with an apical structure different from those of Dε and Hε was priming competent. These results suggest that P protein binding can induce differently structured ε RNAs to adopt a new, common conformation, and they support an induced-fit model of the ε-P interaction in which both components undergo extensive structural alterations during formation of a priming-competent ribonucleoprotein complex.

The hepadnavirus family comprises several mammalian members, including the prototypic human hepatitis B virus (HBV), and two less closely related avian representatives, duck HBV (DHBV) and heron HBV (HHBV). All hepadnaviruses are small, enveloped DNA-containing viruses that replicate by reverse transcription (for reviews, see references 18 and 19) of an RNA intermediate, the pregenome, which also serves as mRNA for the capsid protein and the reverse transcriptase (P protein). Specific encapsidation of the RNA pregenome into nucleocapsids (10, 13, 21) and initiation of its reverse transcription (15, 20, 26, 27, 32) are mediated by the interaction of P protein, probably in concert with cellular chaperones (8), with the structured, *cis*-acting ε signal (Fig. 1A).

Human HBV ε adopts a characteristic bulge-and-loop secondary structure (13, 21). Its sequence and structure are highly conserved in all mammalian hepadnaviruses (16), and its 5′ end-proximal ε signal is sufficient for specific RNA encapsidation (10) and replication initiation (24). In the avian viruses, a second, downstream element (region 2; see reference 6) is required, but otherwise the ε signals of DHBV and HHBV (Dε and Hε, respectively) are clearly related to human HBV ε in function and overall structure (Fig. 1B); both contain a bulge (for sake of simplicity we use this term rather than the correct term, asymmetric internal loop; see reference 17) separating the lower and upper stems and an apical loop (4). In Hε, however, the region corresponding to the upper stem in Dε not only diverges in primary sequence from Dε at seven positions but also contains many fewer Watson-Crick base pairs. Nonetheless, Hε interacts productively (33), and with only about twofold lower efficiency than Dε (4), with DHBV P protein. While this promiscuity suggests a rather relaxed specificity, human HBV ε is completely discriminated against by DHBV P protein (22), indicating the existence of specific constraints for productive binding. To elucidate the sequence- and structure-specific determinants underlying the interaction between P protein and ε, we analyzed a series of avian ε mutants in which we simultaneously replaced several nucleotides (nt) in the apical region, either 6 of the 7 residues that are variable between Dε and Hε (v positions) or all 3 residues in this region that are conserved between DHBV and HHBV (k positions); in addition, we altered two positions in the Dε loop sequence that is completely conserved in Hε. In most cases, the new nucleotides did not correspond to the authentic Dε or the Hε sequence (Fig. 1B and Table 1).

First, we tested productive interaction with DHBV P protein of the mutant RNAs using a *trans*-priming assay that measures the ε-dependent covalent transfer to P protein of a deoxynucleoside monophosphate (dNMP) from the corresponding deoxynucleoside triphosphate (dNTP); this assay revealed a variety of activities whose levels ranged from zero to essentially wild-type (wt) levels. We then correlated functional activity with the predicted and, for several representative variants, experimentally determined secondary structures. Finally, we used a direct assay to distinguish physical and productive binding to DHBV P protein. Our experiments revealed a surprising flexibility in primary sequence in particular at the variable positions. The inactive variants, on the other hand, al-

* Corresponding author. Mailing address: Department of Internal Medicine II, University Hospital, University of Freiburg, Hugstetter Strasse 55, D-79106 Freiburg, Germany.

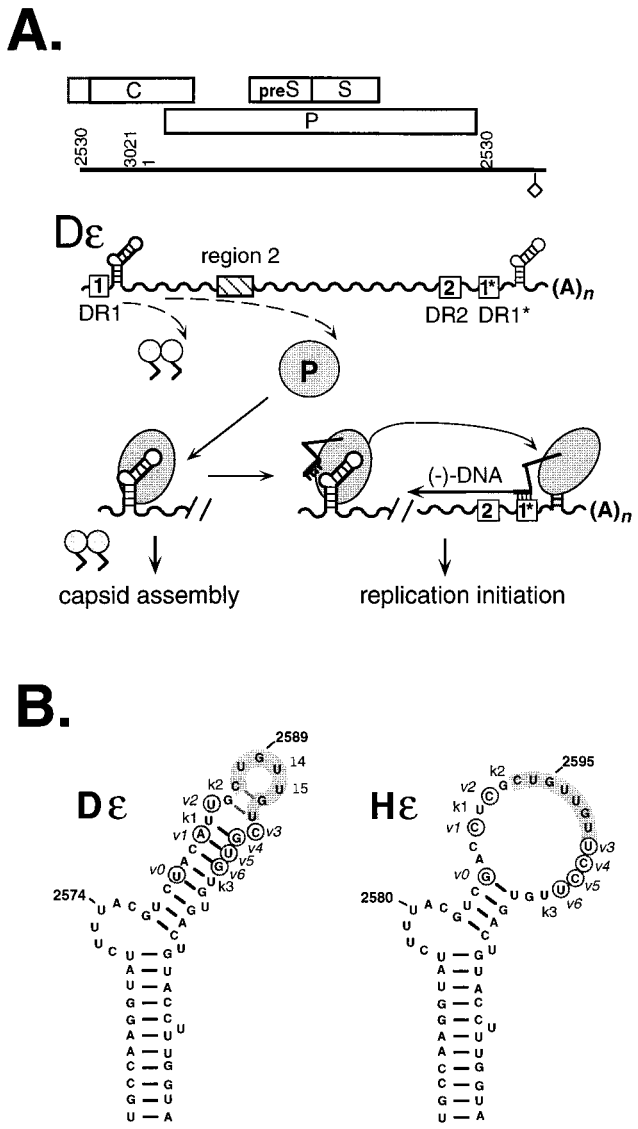

FIG. 1. (A) Functional roles of ε-P interaction. The line symbolizes a linear version of the circular DHBV genome, the diamond represents the polyadenylation signal, and the open rectangles above the line represent the three major ORFs. Numbers are nucleotide positions. The wavy lines represent the terminally redundant RNA pregenome, which also serves as mRNA for the capsid and the P protein (small and large circular and oval objects; the P protein is labeled P). Dε is shown as a symbolic hairpin, and the direct repeat elements DR1, DR2, and DR1* are shown as boxes. Region 2 is an as yet not well-defined second element required for pregenome encapsidation. Binding of P protein to 5′ Dε triggers addition of capsid protein subunits and thus nucleocapsid assembly; in addition, P copies a few nucleotides of the Dε bulge into a DNA primer that is covalently linked to the protein. The entire complex is then translocated to DR1*, and the primer is extended to form minus-strand DNA. (B) Secondary structures of Dε and Hε. The models shown are those in best accord with direct structural data (4). Positions varying between Hε and Dε are encircled (marked v0 to v6); of these, v1 to v6 were simultaneously mutated. The conserved positions in the apical part substituted in this study are marked k1 to k3. The originally proposed 7-nt loop sequence (10) is shaded. The nucleotide numbers of Dε and Hε differ by 6; in this report, DHBV nomenclature is consistently used for all variants.

lowed us to determine several general requirements for productive binding: one is the formation of an ε-like overall structure and another is the presence of certain key residues that are assumed to have direct contacts with the protein be-

TABLE 1. Structural and functional properties of Dε variants

Variant	Mutated positions[a]	$\Delta\Delta G$[b] (kcal/mol)	Experimental two-dimensional structure	P protein interaction[c]	
				Priming	Binding
V series					
Dε	AUCGUG	0	wt Dε	+++	+++
Hε	CCUCCU	1.2	wt Hε	+++	+++
V1	GAGCAA	<0		+++	ND
V3	UAACAC	0.9	Hε-like	+++	+++
V7	UAGCAC	<0		+++	ND
V28	GGACAA	2.0		+++	ND
V2	GAAGAC	8.8		±	ND
V5	UAAGAC	7.9	non-Dε and non-Hε	±	−
V24	GAAGAA	6.2		±	ND
K series					
wt Dε or Hε	UGU				
K2	AAG	<0		±	++
K8	AAC	2.4		±	ND
K9	CAC	2.4	Hε-like	−	+++
K10	CCA	1.6	Hε-like	±	ND
K11	GAC	2.6		±	ND
K12	GUA	1.6	Hε-like	+	+++
K22	CCG	0.3		±	ND
K23	CUA	1.6		±	ND
K26	CUC	2.4		±	ND
L mutants					
wt	UU				
L4	CU	0	wt Dε	±	−
L5	UA	2.5	non-Dε and non-Hε	+++	+[d]
L4,5	CA	1.4	Dε-like	±	ND

[a] For structural context, see Fig. 1B. The positions mutated in the V series were v1 to v6, those mutated in the K series were k1 to k3, and those mutated in the L mutants were l4 and l5.

[b] $\Delta\Delta G = \Delta G_{\varepsilon} - \Delta G_{\text{non }\varepsilon}$, i.e., the difference in free energy ($\Delta\Delta G$) between the most stable ε-like (ΔG_{ε}) and non ε-like ($\Delta G_{\text{non }\varepsilon}$) structures as calculated by M-FOLD.

[c] Priming activity at 1 μM *trans* RNA was normalized to wt Dε (100%) and the Dε-independent priming signal (0%). Binding efficiency at 80 nM ε RNA was normalized to wt Dε (100%) and P-protein-independent background binding (0%). +++, 50 to 100%; ++, 20 to 50%; +, 10 to 20%; ±, 1 to 10%; −, <1%; ND, not determined.

[d] Variant L5 displayed 50 to 60% of wt binding efficiency at a 1 μM concentration.

cause their substitution impedes protein binding. The conserved residues outside the loop apparently have a special role, as several mutants bound to P failed to support priming. However, we also identified a point variant with a secondary structure whose apical region differs from those of both Dε and Hε but still mediates priming. Together, these data show that a complex but tractable hierarchy of structure and sequence-specific determinants is involved in productive ε-P protein interaction; they also suggest structural flexibility within the RNA as an important criterion for protein binding.

MATERIALS AND METHODS

Chemicals, enzymes, and cells. Reagents for agarose gel electrophoresis were purchased from Serva (Heidelberg, Germany), and polyacrylamide gels were cast from premade acrylamide solutions (AppliChem, Darmstadt, Germany). Enzymes for molecular cloning experiments were obtained from Boehringer (Mannheim, Germany) or New England Biolabs (Bad Schwalbach, Germany). For enzymatic structure probing, RNase A and T_1 from Boehringer were used. In vitro translations were performed with rabbit reticulocyte lysate from Promega (Heidelberg, Germany). Radioactively labeled nucleotides were obtained from Amersham (Braunschweig, Germany). All plasmids were transformed into *Escherichia coli* Top 10 cells (Invitrogen, NV Leek, The Netherlands) and grown

in Standard I medium (Merck, Darmstadt, Germany) supplemented with ampicillin at 50 μg/ml.

Secondary structure predictions. For secondary structure predictions, the program M-FOLD (38) was used. For comparison with experimental secondary structure analyses, all foldings were performed with the complete sequences of the experimentally used 76-nt ε transcripts comprising the 57-nt stem loop (DHBV nt 2560 to 2616; HHBV nt 2566 to 2622) plus additional and, in all constructs, identical terminal sequences as previously described (4). Since these additional sequences did not pair with the central sequences in essentially all optimal and relevant suboptimal foldings, they have apparently no influence on the folding of the stem-loop part. All predictions were carried out for a default temperature of 37°C.

Plasmid constructs. All Dε mutants were obtained by replacing the original *Hin*dIII-to-*Cla*I Dε cassette in plasmid pDε1 (3) with synthetic DNA duplexes that were degenerate at the desired positions. For the V series (mutations at the V positions), these positions were nt 2583 (G/T), 2585 (A/G), 2594 (A/G), 2595 (C/G), 2596 (A/G), and 2597 (A/C), for the K series (mutations at the K positions), these positions were nt 2584 (A/C/G), 2586 (A/C/T), and 2598 (A/C/G), and for the L mutants (mutations in the loop sequence), these positions were nt 2590 (T/C) and 2591 (T/A). Nucleotides at other positions are nucleotides of DHBV 16, except for the K variants, which contain HHBV sequence at positions 2583 (C), 2585 (C), and 2594 to 2597 (T, C, C, and T, respectively). Individual clones obtained after transformation into *E. coli* Top 10 cells were characterized by sequencing; the corresponding variants are listed in Table 1. As the template for in vitro translation of DHBV 16 P protein, we used plasmid pT7AMVpol16H6 (3), which encodes the complete P protein open reading frame (ORF) plus an insert of six His residues between codons 2 and 3; in the in vitro priming assay, this variant was as active as wt P protein.

In vitro transcription. Mutant and wt Dε RNAs were obtained by in vitro transcription of the corresponding plasmids after linearization with *Cla*I. The T7 transcripts are nominally 76 nt in length and contain DHBV sequence from nt 2557 to 2624 plus short terminal vector-derived sequences essentially as described above. Unlabeled RNAs were produced with a MEGAscript kit (Ambion, Austin, Tex.). RNA concentrations were determined photometrically and by comparison with a tRNA standard on ethidium bromide-stained agarose gels. For binding studies and native gel analysis, internally ^{32}P-labeled RNAs (specific activity, about 1.5×10^6 cpm/pmol) were obtained by in vitro transcription in the presence of [α-^{32}P]CTP (specific activity, 800 Ci/mmol) with T7 RNA polymerase (New England Biolabs).

Secondary structure analysis. Enzymatic probing was performed essentially as previously described (12). In vitro transcripts ^{32}P labeled at their 5′ ends were gel purified, precipitated with isopropanol, resuspended in TE buffer (10 mM Tris-Cl$^-$, 1 mM EDTA [pH 7.4]), and renatured by heating to 75°C and slow cooling to room temperature. Aliquots containing 5×10^4 cpm in a volume of 50 μl of TMK buffer (50 mM Tris-Cl$^-$, 40 mM KCl, 10 mM $MgCl_2$ [pH 7.5]) supplemented with 100 μg of total *Saccharomyces cerevisiae* RNA per ml were treated for 20 min at 23°C with RNase A (specific for unpaired pyrimidines; 1 ng per reaction) and RNase T_1 (specific for unpaired G residues; 1 U per reaction). Alkaline hydrolysis ladders were obtained as previously described (3). G-specific RNA sequencing was performed by RNase T_1 digestion under denaturing conditions (14). Reactions were stopped by phenol extraction, and the products, after ethanol precipitation and separation on 8% polyacrylamide gels containing 7 M urea, were visualized with a PhosphorImager (Molecular Dynamics, Krefeld, Germany).

Native polyacrylamide gel electrophoresis (PAGE). Internally ^{32}P-labeled Dε RNAs were gel purified to electrophoretic homogeneity, renatured in TE buffer, and adjusted to structure probing buffer conditions as described above. After 15 min of equilibration at room temperature, 1/6 volume of loading buffer (60% sucrose in 10 mM Tris-Cl$^-$, 100 mM NaCl, and 0.1 mM EDTA [pH 7.5]) was added and the samples were electrophoresed for 2 h at 20 W and a surrounding temperature of 4°C on a precooled gel containing 10% polyacrylamide in 0.5× Tris-borate-EDTA buffer. As a control, several RNAs were also renatured in TMK buffer; this treatment did not alter their mobilities from those after renaturation in TE buffer.

***trans*-Priming assay with DHBV P protein and mutant Dε RNAs.** His-tagged DHBV P protein was in vitro translated in rabbit reticulocyte lysate programmed with a Dε-deficient transcript obtained from plasmid pT7AMVpol16H6 linearized with *Afl*II (DHBV position 2526). After 1 h at 30°C, mutant or wt Dε RNAs were added to equal aliquots of the lysate at a 1 μM concentration and incubated for 90 min at 30°C to allow for binding to P protein. Priming was performed by adding [α-^{32}P]dATP (specific activity, 3,000 Ci/mmol) in 1 volume of 2× priming buffer (34) and incubation for 1 h at 37°C. Reactions were stopped by addition of sample buffer, and aliquots were analyzed by sodium dodecyl sulfate-PAGE. Labeled P protein was visualized with a PhosphorImager, and relative priming efficiencies were estimated by quantifying the band intensities with ImageQuant software.

Direct binding assay. P protein-bound and free Dε RNAs were separated as previously described (3). In brief, His-tagged DHBV P protein was in vitro translated for 30 min and then internally ^{32}P-labeled mutant or wt Dε RNA (specific activity, about 1.5×10^6 cpm/pmol) was added to a final concentration of 80 nM and incubated for a further 75 min at 30°C. For metal-affinity purification, the samples were mixed with 1 volume of Ni-nitrilotriacetic acid-agarose (Qiagen, Hilden, Germany) in 400 μl of binding buffer (3) and shaken for 90 min at 4°C. The beads were washed twice with binding buffer and twice with TMK buffer, and the remaining radioactivity was measured by scintillation counting. Bound RNA was isolated by phenol extraction and subsequent ethanol precipitation and analyzed by denaturing PAGE.

Indirect binding assay by priming competition. This assay was performed essentially like the *trans*-priming assay except that instead of a single ε RNA species, a mixture containing 0.2 μM wt Dε and 2 μM mutant Dε RNA was used.

RESULTS

Design of mutant Dε sequences. Six of the seven positions where Hε differs from Dε are clustered at the base of the Dε loop (designated v1 to v6 [Fig. 1B]), while the loop sequence itself and three residues at its boundaries (designated k1 to k3) are identical. The term loop sequence is here defined as the apical 7 nt suggested to be unpaired in the originally proposed Dε structure model (10), although direct data are more compatible with the slightly different arrangement shown in Fig. 1 (4). Assuming that sequence conservation and divergence in the authentic viral RNA elements may reflect different structural and functional contributions from the corresponding positions, we chose to separately mutate the k and the v positions. To this end, we introduced into the previously described pDε plasmid (3) synthetic degenerate DNA cassettes in which either the v or k residues were replaced by, mostly, non-Dε or non-Hε nucleotides. From randomly picked colonies we isolated seven different V and nine K variants. Their sequences are listed in Table 1, together with those of the loop variants L4, L5, and L4,5, in which, in the context of wt Dε, the U residues at loop positions 4 and 5, either alone or in combination, are exchanged for C and A, respectively. L4 and L4,5 have previously been shown to have strongly reduced activities in the priming assay (2).

Functional analysis of Dε and Hε variants by in vitro priming. DHBV P protein, in vitro translated in rabbit reticulocyte lysate, exhibits enzymatic activity (31) as measured in a priming assay (Fig. 2A). A functional Dε sequence, either on the RNA used for P protein in vitro translation (which contains region 2 of the encapsidation signal as part of the P ORF) or added as a separate RNA in *trans*, mediates the covalent transfer to the protein of an appropriate dNMP from the corresponding [^{32}P]dNTP (34). This reaction mimicks the first step of reverse transcription of the RNA pregenome and provides a sensitive test for the functional integrity of variant Dε sequences; a negative result does not, however, discriminate between a lack of physical versus productive binding (see below). In our study, the variant ε sequences were in vitro transcribed, isolated, and subjected to the priming assay at a standard *trans*-RNA concentration of 1 μM (for wt Dε RNA, P protein labeling reaches a maximum at this concentration; see references 4 and 33). Usage of equal aliquots from a single preparation of P protein served to exclude experimental variation in the levels of efficiency of in vitro translation.

The results obtained with the individual variants are shown in Fig. 2B. Most evident were the extensive variation in the signal intensities produced by the V variants, ranging from almost background to wt levels (compare, e.g., Fig. 2B, lanes V3 and V5), and the general strong reduction in signal intensities observed for the K variants (e.g., lanes K2 to K26). Of the loop mutants, L4 and L4,5 were essentially inactive while L5 produced a relatively strong signal (L lanes). Semiquantitative values were obtained by measuring the signal intensities with the PhosphorImager (Table 1). Hence, on one hand, a surprisingly large number of nucleotide exchanges were tolerated without functional consequences, in particular at the v positions, while other combinations, even of a few foreign nucle-

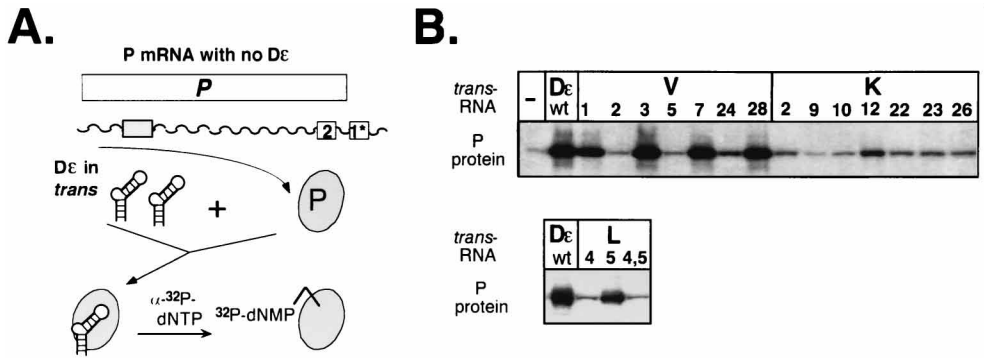

FIG. 2. Effects of mutations on the priming reaction. (A) Principles of the priming assay. P protein is in vitro translated from an mRNA lacking Dε, which is instead supplied in *trans* as a separate RNA molecule (see legend to Fig. 1 for clarification of symbols). Upon productive binding and provision of a suitable ^{32}P-labeled dNTP, P protein becomes covalently linked to the corresponding dNMP. (B) Priming efficiencies of various V, K, and L mutants. The indicated *trans* RNAs were added at a concentration of 1 μM to equal aliquots from a P protein in vitro translation reaction mixture, and the labeled P protein, after sodium dodecyl sulfate-PAGE, was visualized with a PhosphorImager. Nucleotide sequences of the individual V, K, and L variants are summarized in Table 1, which also contains an estimate of their priming efficiencies as measured by the extent of P protein labeling. Note that the weak bands obtained with some variants do not exceed the background priming signal in the lane without Dε (−).

otides as in the K variants or a single substitution as in variant L4, led to reduction if not complete loss of function. With the following experiments, we sought to link these functional differences with the secondary structures of the individual RNAs.

High sequence flexibility at the variable positions in the context of an ε-like overall structure. First, we used the M-FOLD program (38) to predict the secondary structures of the variant ε RNAs, although the validity of this theoretical approach is limited in some respects; for instance, the program slightly favors non-ε-like secondary structures for Dε and Hε over the experimentally confirmed bulge-and-loop arrangement (4). For wt Dε and Hε, the calculated differences in free energy between ΔG values of the most stable ε-like and non-ε-like structures (termed $\Delta\Delta G$ values and calculated for ε RNAs of 76 nt in length) are rather small (0 kcal/mol for Dε and 1.2 kcal/mol for Hε). Many variants exhibited similarly low $\Delta\Delta G$ values ranging from <0 to 2.6 kcal/mol, which would not allow for unambiguous statements about their structures. However, for the V variants we noted a striking correlation between inactivity in the priming assay and a large positive $\Delta\Delta G$ value ($>$6.0 kcal/mol) (Table 1). Such a large unfavorable $\Delta\Delta G$ plausibly reflects the possibility that the corresponding RNAs adopt a stable non-ε-like structure that is incompatible with P protein binding. To test this notion, we selected a representative pair of V mutants with drastically different activities, namely V3 and V5, for direct enzymatic secondary structure analysis. Both variants contain the same non-Dε and non-Hε nucleotides at five of the six mutated v positions and differ only at v4. Mutant V3 carries a C residue, displays near wt priming activity, and has a small $\Delta\Delta G$ value (0.9 kcal/mol). V5, with a G residue at the corresponding position, is essentially inactive and exhibits a large $\Delta\Delta G$ (7.9 kcal/mol). This pair of mutants was therefore particularly attractive for distinguishing between structural and sequence-specific contributions to productive P protein binding.

While we have not completely determined the structures of these RNAs, we took advantage of the fact that the existence of an ε-like structure can be easily monitored by its sensitivity to digestion with RNase A at U2574 in the Dε bulge. The absence of this signal, in turn, indicates the absence of the ε-like bulge; similarly, G2589 is a hallmark residue in the wt Dε loop (Fig. 1; these residues are also explicitly marked in Fig. 4 to 7). For simpler comparison, we use the DHBV numbering system also to indicate the corresponding residues in Hε (authentic HHBV numbers may be easily obtained by adding 6 to the given DHBV position number).

As the presence of more than one conformer can lead to superposition of cleavage signals from different structures, we first analyzed Dε, Hε, and variants V3 and V5 as well as variant L5 (see below) by nondenaturing gel electrophoresis. Because of the known 3′ end heterogeneity of in vitro transcripts, RNAs of homogeneous lengths were isolated by electrophoresis on denaturing gels and renatured and their mobilities were compared under denaturing and native conditions (Fig. 3). Expectedly, all transcripts migrated identically in the presence of urea, while distinct mobility differences were revealed on the nondenaturing gel. All variants tested in this assay produced only one major band, suggesting the presence of one major conformer. Identical mobility patterns were obtained regardless of whether renaturation of the transcripts was performed in TE buffer or in the TMK buffer (not shown) used in the subsequent probing experiments. These were essentially performed as previously described for wt Dε and Hε (4).

The results for the active variant V3 and the inactive variant V5 are shown in Fig. 4A. Most evident was the very small number of accessible nucleotides in V5, all in a contiguous region partially overlapping the Dε loop; in particular, the bulge-specific signal at U2574 was virtually absent. V3, by

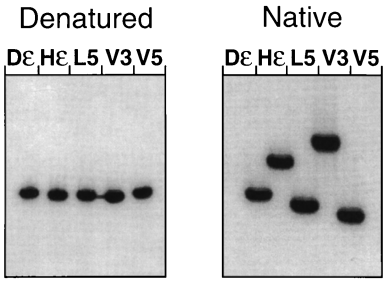

FIG. 3. Electrophoretic mobilities of selected ε RNAs under native and denaturing conditions. The indicated in vitro transcripts were gel purified under denaturing conditions, renatured, and analyzed by electrophoresis in polyacrylamide gels containing no urea (native) or 7 M urea (denatured). The appearance of a single major band for each RNA on the native gel suggests the presence of only one major conformer. The high mobility of V5 versus that of V3 RNA indicates a compact, largely base-paired structure, while the low mobility of Hε versus that of Dε is in accord with its relatively open apical conformation.

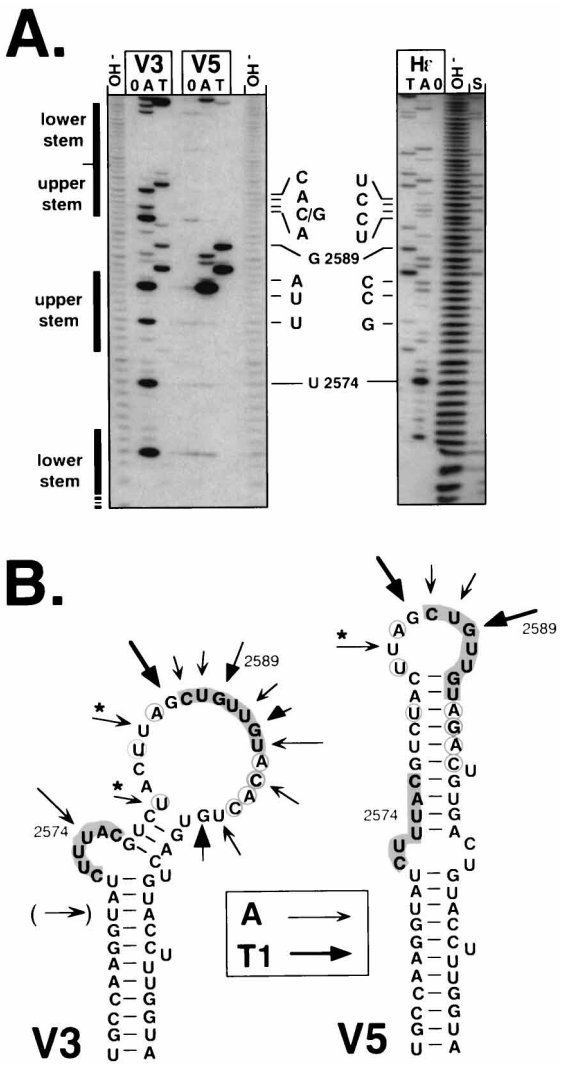

FIG. 4. Secondary structures of variants V3 and V5. (A) Enzymatic structure probing. 5′-end-labeled V3 and V5 RNAs were denatured, renatured, and treated with RNase A, T_1, or no enzyme (lanes A, T, and 0, respectively). The products were analyzed on denaturing polyacrylamide gels; as a marker, RNA fragments obtained by partial alkaline hydrolysis were used (lanes OH^-). Hε lanes show the products obtained with wt Hε RNA, and lane S shows the G-specific fragments generated from wt Dε RNA by RNAse T_1 under sequencing conditions. The bars on the left show the positions of the stem-loop elements in wt Dε. The bulge- and loop-specific indicator residues U2574 and G2589, respectively, as well as those residues that differ between Hε and the variants, are explicitely indicated. C/G indicates the only sequence variation between V3 and V5. (B) Secondary structure models for V3 and V5 RNAs. The prominent cleavage positions derived from RNAs treated with RNase A and T_1 shown in panel A are superimposed (arrows) on the structures which they fit best. The sizes of the arrows correspond approximately to the signal intensities. The bulge and loop nucleotides of wt Dε are shaded in both models. The arrow in parentheses corresponds to a product that in wt Dε is due to secondary cleavage (4). Asterisks mark residues that may be hypersensitive to RNAse A digestion, because the following A creates a preferred substrate for the enzyme (12).

contrast, produced a series of signals, both at U2574 and in the apical region. Many of these coincided with nucleotides also accessible in Hε (Fig. 4A, Hε lanes). These data indicate a largely base-paired, non-ε-like arrangement for V5 that is fully compatible with the calculated most stable model (Fig. 4B), and that arrangement is further substantiated by the significantly increased mobility of V5 RNA on the nondenaturing gel (compare lanes V3 and V5 in Fig. 3). The active variant V3, by contrast, shares many structural features with Hε, especially the bulge and a relatively open apical region. We conclude that a stable non-ε-like structure interferes with productive binding to P, that even a single nucleotide exchange can bring about this type of drastic structural change, and that formation of such improper structures is predictable by a large $\Delta\Delta G$ value as defined above; this latter assumption is substantiated by the fact that all V variants with $\Delta\Delta G$-values of >6.0 kcal/mol (Table 1) were inactive. Hence, the only major restriction for sequence divergence at the v positions seems to be that new nucleotides must not, in the given sequence context, induce formation of stable non-ε-like structures.

Sequence-specific constraints to priming competence at the conserved k positions despite an ε-like overall structure. All of the nine K variants displayed drastically reduced activities in the *trans*-priming assay (Fig. 2), suggesting that at least one of the k positions contributes in a sequence-specific way to productive P binding. In contrast to those of the V variants, the calculated $\Delta\Delta G$ values were small and gave no clear indications for stable non-ε-like structures. We therefore directly analyzed the secondary structure of the best-performing variant, K12, which had about 15% of the activity of wt Dε, and that of K9, which was essentially inactive; their primary sequences differ at k positions 1, 2, and 3 (K12, GUA; K9, CAC). The nuclease sensitivity patterns of both variants (Fig. 5A) were similar to that of wt Hε; the bulge indicator at nt U2574 was clearly detected, and the series of further products indicated a similarly nonpaired apical region. One apparent difference was the relatively strong signal, in K12, at position U2597, which was weak in wt Hε and virtually absent from K9. How this relates to the partial priming proficiency of K12 is currently not clear. However, while subtle structural differences cannot be monitored by enzymatic probing, the data provide direct evidence that it is not overall structure, as with V5, that reduces their priming competence but rather the nature of the replaced nucleotides. Surprisingly, however, the lowered or abolished priming competence of the K mutants did not correlate with an inability to bind to P protein (see below).

Drastically different consequences of mutations at individual positions in the Dε loop sequence. All of the variants described above contained multiple mutations in comparison to Dε and Hε. The loop sequence (positions 2587 to 2593 in Dε and positions 2593 to 2599 in Hε) is completely conserved between the two viruses, and several previously published data had suggested that the loops in Dε (22, 28) and human HBV ε (21, 23) contain important sequence-specific determinants for P binding. The three loop variants we analyzed differ from wt Dε only at one or two positions (in L4, U2590 is replaced by C; in L5, U2591 is replaced by A; and L4,5 contains both mutations). In accord with previous data (2), L4 and L4,5 exhibited drastically reduced priming activities, but L5, at a concentration of 1 μM, gave a signal of about two-thirds of that of wt Dε (Fig. 2B, L lanes).

Surprisingly, the experimental structure analysis of the variants (Fig. 6) revealed that the inactive mutant L4 produced a nuclease sensitivity pattern that was absolutely indistinguishable from that of wt Dε (Fig. 6A; compare Dε and L4 lanes); a similar result was obtained with L4,5 (L4,5 lanes). Hence, despite having a Dε-like secondary structure, these mutants are unable to productively interact with P protein; the single nucleotide exchange in L4 acts on the sequence-specific rather than on the overall structural level. The active variant L5 showed clear signals around U2574, which, though they were weaker than in wt Dε, indicated that the bulge nucleotides were exposed; however, a whole series of new fragments was

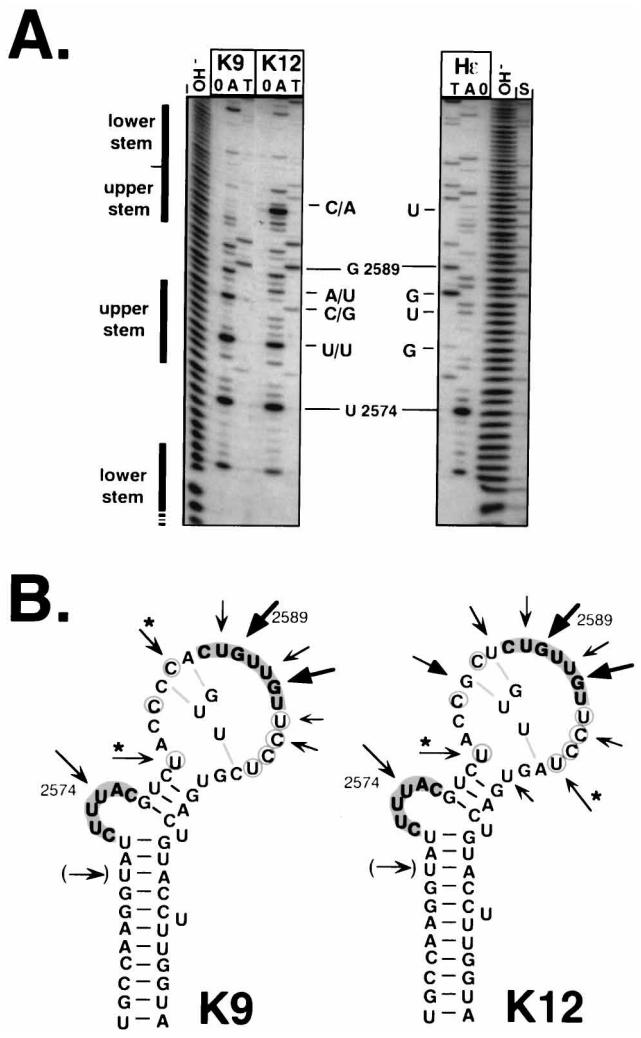

FIG. 5. Secondary structures of variants K9 and K12. (A) Enzymatic probing. Probing data were obtained as described for variants V3 and V5 in the legend to Fig. 4. (B) Secondary structure models. Cleavage positions derived from the results of experiments shown in panel A are marked by arrows (see the legend to Fig. 4B for clarification); the structures shown are those that fit best the experimental data. Authentic residues at the k positions are indicated inside the loop; the asterisks indicate cleavage sites containing the preferred RNase A target pyrimidine A.

generated in a region corresponding to the upper-right half-stem of Dε (Fig. 6A, L5 lanes) and the strongest loop-specific signal was shifted from G2589 to G2586. This nuclease pattern does not accurately correspond to any of the stable conformations predicted by M-FOLD. To test whether L5 RNA adopts two or more similarly stable conformations producing overlapping patterns, we also subjected this variant to nondenaturing gel electrophoresis; as only a single major band was observed (Fig. 3), we have no positive indication for the presence of more than one major conformer. Because of these difficulties, we do not propose a specific secondary structure model for L5 (in Fig. 5B, the nuclease pattern is simply superimposed on the wt Dε structure). However, the additional nuclease-sensitive sites in L5 clearly demonstrate the structural differences in its apical region from Dε; notably, this rather drastic structural change is caused by a single exchange in a nonpaired region. Furthermore, the structure of free L5 RNA is not identical to that of Hε, yet L5 rather efficiently supports the priming reaction with DHBV P protein (see above). Hence, L5, in addition to Hε, provides another example of a distinctly structured RNA that interacts productively with the same protein.

Physical interaction with P protein is required but not sufficient for productive binding of ε RNA. Production of signals in the priming assay beyond a low background is dependent on the presence of ε. On the other hand, evidence for the existence of binding-competent but priming-inactive ε RNAs has been previously presented (22). We therefore applied a direct interaction assay to the variants described above to distinguish between physical and productive binding that leads to priming. The assay relies on immobilizing an N-terminally His-tagged DHBV P protein to an Ni^{2+}-containing matrix; binding of labeled ε RNA to P protein is monitored by scintillation counting and reanalysis of the bound RNA by denaturing PAGE. RNA immobilization to the matrix is P protein dependent and specific, as Dε but not human HBV ε RNA is retained by DHBV P protein (3).

In one set of experiments, we compared the amounts of various internally labeled ε RNAs which, after incubation at an average RNA concentration of about 80 nM with in vitro-translated P protein, were immobilized to the solid phase (Fig. 7A). A semiquantitative evaluation of these data is shown in Table 1. The priming-incompetent variants V5 and L4 did not significantly bind to P protein, confirming our conclusion that an improper overall structure (V5) or mutation of a key residue (such as U2590 in L4) prevents binding. However, not only the priming-competent wt RNAs and variants V3 and L5 but also K12, K9, and, to a lesser extent, K2 (not shown) bound to P protein, despite their severely reduced or even completely negative priming phenotype.

These results were further corroborated in priming competition experiments. Here, a normal *trans*-priming reaction with 200 nM wt Dε RNA was supplemented with a 10-fold molar excess of several K mutants and the nonbinding L4 variant. In this format, binding-competent variants were expected to compete with wt Dε RNA and hence to reduce the priming signal. In accord with their inactivity at a concentration of 80 nM (Fig. 2B), the tested K variants gave signals only slightly above background when they were present alone at a 2 μM concentration (Fig. 7B, lanes without wt Dε). However, all K variants, but not L4, substantially diminished wt-Dε-mediated labeling of P protein (Fig. 7B, lanes with wt Dε) and therefore acted as competitive inhibitors. Hence, many K variants are able to interact with P protein despite their severe priming defects. This clearly attests to the existence of requirements for priming that go beyond mere physical binding.

DISCUSSION

The ε sequence is highly conserved in all mammalian hepatitis B viruses. Not unexpectedly, therefore, the ε signals and P proteins of human HBV and woodchuck HBV (WHV) can functionally substitute for each other (37). In this evolutionary light, the substantial sequence and structure divergence in the apical part of the avian hepadnavirus ε signals is per se remarkable; even more surprising is the similarly efficient interaction of Dε and Hε with DHBV P protein (4, 33).

The combined structural and functional analysis of the Dε variants reported in this paper allows for several generalizations that explain this apparent promiscuity: (i) in accord with their divergence in Dε and Hε, the v positions tolerate many mutations and hence are not essential for sequence-specific contacts with protein; (ii) even here, however, single nucleotide exchanges can be deleterious if they induce stable, nonbulge-containing structures; (iii) such nonproductive structures

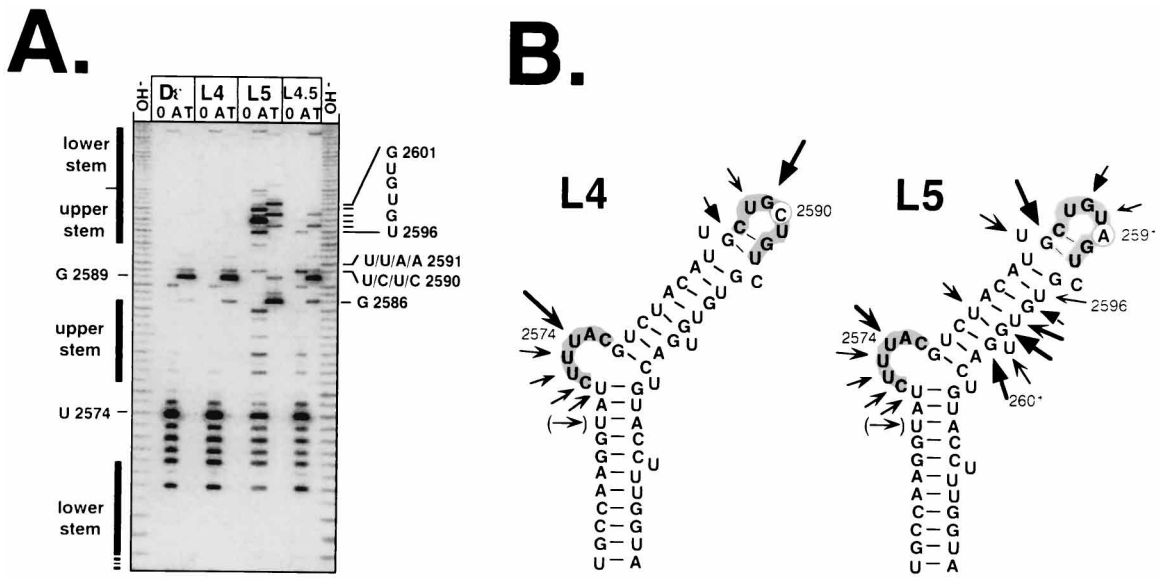

FIG. 6. Secondary structures of variants L4, L5, and L4,5. (A) Enzymatic probing. Experimental conditions were as described in the legend to Fig. 4. The specific nucleotides noted for positions 2590 and 2591 reflect the order of the samples as they were applied to the gel. wt Dε and L4 generated indistinguishable patterns, while in L5 a stretch of nucleotides corresponding to the upper-right half-stem in Dε was clearly accessible; the exact positions of the most prominent signals are marked on the right. (B) Secondary structure models. The structural model for L4 is identical to that of wt Dε and in accord with the data; for L5, none of the stable predicted structures fits perfectly to the probing data; for better orientation, the cleavage positions are superimposed on the wt Dε model.

can be predicted by a large unfavorable $\Delta\Delta G$ value; (iv) certain key residues act as sequence-specific determinants, as demonstrated by the loop variant L4, in which replacement of a single U prevents protein binding despite preservation of the wt Dε secondary structure; (v) physical binding is insufficient for functional activity as illustrated by many k-position mutants; and (vi) not only the natural Hε RNA but also an artificial Dε variant, L5, with a substantially differing apical structure can productively interact with DHBV P protein. Based on these results, we propose a dynamic model for the ε-P interaction that accounts for the phenotypes of all analyzed RNAs.

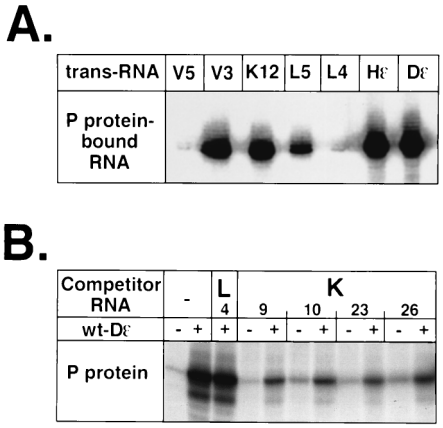

FIG. 7. Physical binding of ε RNAs to DHBV P protein. (A) Direct binding assay. The indicated internally ^{32}P-labeled ε RNAs were incubated for 75 min with in vitro-translated, N-terminally His_6-tagged DHBV P protein; then, Ni^{2+}-loaded nitrilotriacetic acid-agarose beads were added and the RNAs that remained bound after extensive washing were analyzed by denaturing PAGE (3). (B) Competitive binding assay. Priming reactions were performed as described in the legend to Fig. 2 with 2 μM mutant RNA, either alone (−) or in the presence of 200 nM wt Dε RNA (+). The lanes without competitor RNA (−) show the background signal produced with no *trans* RNA and the level of protein labeling with 200 nM wt Dε alone. Note that all K variants strongly reduced the priming signal and that mutant L4, which did not bind in the direct assay, also had no measurable effect in this assay format.

Sequence flexibility and constraints in the apical part of avian hepadnavirus encapsidation signals. The initial functional screen by the *trans*-priming assay revealed drastically different activities for the V variants, while most mutations at the conserved k positions and a single mutation at one position in the loop (L4), but not at another position (L5), markedly reduced or abolished productive binding (Table 1).

The straightforward conclusion from the activity of many V variants with non-Dε and non-Hε nucleotides at five of six positions is that many different sequences fulfil all requirements for productive interaction with P protein. Without consideration of the role of the Dε or Hε sequence as part of the pre-C ORF, this result indicates that there is little selective pressure on conserving the primary sequence at the v positions; this conclusion plausibly explains the divergence between Dε and Hε. The v positions also show the highest variability when the ε signals of all DHBV sequences deposited in the GenBank database are compared (not shown). In addition, these data argue strongly against sequence-specific contributions of the v positions to P protein binding, since active RNAs containing each of the four possible nucleotides were found.

Nonetheless, some V mutants were essentially inactive; our data indicate that this defect is caused by the formation of stable, non-ε-like structures. The inactive variant V5 showed very few nuclease-sensitive sites, and the bulge-specific RNase A cleavage products were virtually absent. The existence of the largely base-paired structure lacking the bulge derived from these experiments is further supported by the high mobility of V5 RNA on native polyacrylamide gels (Fig. 3); bulges, as present in wt Dε and Hε RNAs and in the active variant V3, have been demonstrated to induce kinks that slow down migration (5).

The proposed V5 structure also conforms to the most stable structure predicted by the M-FOLD program. While it is difficult to judge the relevance of alternative predicted structures

with similar calculated free energies, as with wt Dε and Hε (4), the differences in stability (Table 1) between ε-like and non-ε-like structures for V5 and all other inactive V variants were drastically higher ($\Delta\Delta G$ = 6.2 to 8.8 kcal/mol) than for the active variants ($\Delta\Delta G$ = ≤2.0 kcal/mol); also, the absolute ΔG values for these improper structures were much larger than for wt Dε and the active variants (ΔG = −16 to −18 kcal/mol versus −10 to −13 kcal/mol). Hence, a large positive $\Delta\Delta G$ value is most likely a valid criterion to predict the inactivity of ε mutants. The low intrinsic stability of the ε-like arrangement in the avian signals is probably the reason for the dramatic influence that sometimes even single nucleotide substitutions had on overall RNA structure. For human HBV ε, by contrast, the experimentally confirmed bulge-and-loop arrangement is also calculated to be much more stable than any non-ε-like structure ($\Delta\Delta G$ = −6.0 kcal/mol for a 57-nt ε RNA).

For the K variants, as well as the loop mutants, the $\Delta\Delta G$ criterion was not applicable, as all calculated stability differences were low (less than 2.6 kcal/mol); this suggested in turn that grossly aberrant overall structures were not responsible for their greatly reduced or abolished activities. Of the nine k-position mutants tested in the priming assay, only K12 produced a reasonable signal (about 15% of that of wt Dε). Enzymatic probing of variants K12, K9, and K10 (not shown) produced signals compatible with all RNAs having ε-like features, in particular, the typical bulge. In the apical region, relatively many signals were observed, mostly at the same positions as in Hε. Currently, we cannot rule out the possibility that minor structural alterations compared with Dε or Hε contribute to the reduced priming activities of the K variants; also, we have not explicitely tested whether these RNAs are present in one or more different conformations. However, our data suggest that the K variants have similar overall ε-like structures and that their priming defects relate to sequence rather than structure. This argues for an involvement of one or more of the k positions in sequence-specific interactions with the protein that do not, however, act on the level of physical but of productive binding (see below). The functional defects of the K variants are mirrored by the evolutionary conservation of the k positions in Dε and Hε and also in the other known DHBV sequences.

A clear-cut example for the importance of RNA sequence in P protein recognition was provided by variant L4, in which the replacement of U2590 by C, in accord with previous observations (2; mutant SLM13 in reference 29), drastically reduced functional activity but did not induce any detectable structural differences from wt Dε. This striking dependence of priming (and binding to the protein [see below]) on the chemical nature of the base in the context of an unaltered overall structure suggests U2590 as a direct contact point between RNA and protein.

In this light, the substantial priming activity of mutant L5 was unexpected, as in this mutant U is replaced by the chemically less related A. With the caveat that the observed nuclease products arose from more than one RNA conformer present in the test solution (although we detected only one major band on native polyacrylamide gels), the presence of the nuclease products indicated the accessibility of the characteristic bulge nucleotides but also the single strandedness of a region that, in Dε, forms the right half of the upper stem. Establishing an explicit secondary structure model for L5 will require further investigations; however, the available data clearly demonstrate that free L5 RNA, or at least a major portion of it, adopts a structure that is distinct from those of Dε and Hε yet is able to support priming. Hence, the naturally occurring Hε sequence represents probably but one of many alternatives to authentic Dε RNA that can productively interact with P protein.

Physical versus productive RNA binding to P protein. Application of a direct binding assay (3) allowed us to distinguish physical from productive RNA-P protein interactions. The priming-defective variants V5 and L4 did not detectably bind to immobilized P protein, which for L4 is in accord with an independently obtained result by Tavis and Ganem (mutant SLM13 in reference 29; this RNA apparently displayed some activity in a yeast expression system [28]). Thus, both an improper overall structure (V5) and the absence of a key residue like U2590 (L4) prevent efficient physical binding to P protein and, as a consequence, priming. Drastic effects of single nucleotide exchanges on protein binding similar to those in L4 have also been observed in some other protein-binding RNAs (for a review, see reference 7), for instance, in human immunodeficiency virus type 1 TAR (25), the iron response element (9), and RNA phage operators (35). Recent structural data indicate that often these mutationally defined residues are in direct contact with their cognate proteins (e.g., reference 30). However important such single nucleotides may be, their role in binding is conceivably modulated by, in particular, the local environment (36). This role of the local environment may explain why simultaneous substitution (G→A) of the residue preceding the L4 mutation was reported to partially restore binding (mutant de-loop3,4 in reference 22). On the other hand, our variant L5 was clearly priming competent while the double-mutant de-loop5,6 (containing the nucleotides CA instead of the authentic UG) apparently bound to P but was priming deficient (22). Hence, context is important, and conclusive interpretations for variants with multiple nucleotide exchanges require analyses of several related mutants.

It is therefore noteworthy that all tested K mutants rather efficiently bound to immobilized P protein; the partially active variant K12 did not bind better than the priming-incompetent mutant K9, as estimated from the similar amounts of ^{32}P-labeled RNAs retained on the matrix at two different concentrations (80 nM and 1 μM). Limited titration data indicate three- to fourfold-reduced but easily detectable binding for variant K2. That binding occurs at the authentic ε binding site on P was confirmed in the priming competition experiments where all binding-proficient variants, but not the binding-incompetent mutant L4, led to clear reductions in the intensities of the wt Dε-mediated priming signal. These data confirm that RNA binding is an essential prerequisite but not per se sufficient for priming.

The existence of binding-competent but priming-deficient Dε loop mutants has led to the proposal that, as with the TAR-Tat interaction in human immunodeficiency virus type 1 (for a review, see reference 11), a cellular loop-binding factor may be required for priming (22). The high frequency of the same phenotype in the K mutants makes it equally likely that the priming defect is directly related to the ε-P interaction, which suggests that the k positions contain determinants not only involved in initial P recognition but in a subsequent structural rearrangement (see below).

A model for productive ε-P protein interaction. Several, but not all, of the above-described results may be explained by a simple lock-and-key recognition mechanism. RNAs having a grossly aberrant shape (V5), or containing improper bases at one or more key positions (L4), do not stably bind to the protein (Fig. 8A). However, the fact that three RNAs with differing secondary structures not only physically but productively interact with DHBV P protein, as well as the fact that binding-competent K variants are unable to support priming, deserves further attention. The rather moderate functional

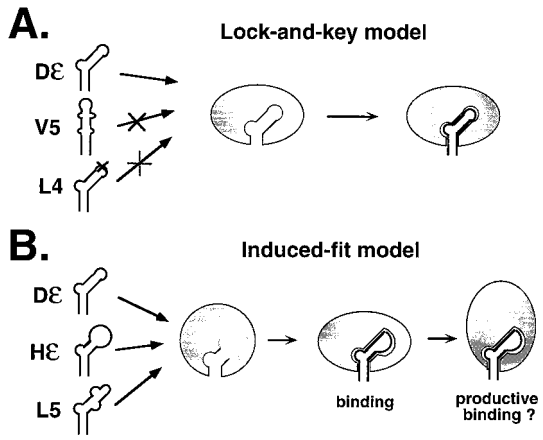

FIG. 8. Models for the interaction between avian hepadnavirus ε signals and DHBV P protein. (A) Lock-and-key model. P protein has a binding pocket whose shape determines whether an RNA can bind or not; binding is prevented by an improper overall RNA structure (e.g., that of variant V5) or the absence of one or more key residues, even in the context of a correct overall structure (variant L4; ×). This simple model, however, does not distinguish between physical and productive binding, and it does not explain productive binding by RNAs with differently structured apical regions. (B) Induced-fit model. As in panel A, overall structure and key residues are important, but in addition, structural flexibility in the apical ε part allows for a conformational change in the RNA (and, possibly, the protein). All binding-competent RNAs meet this criterion, but priming competence requires additional structural alterations in the RNA and the protein (see reference 29; the structural alterations are symbolized by the different shape of the P protein); in binding-competent but priming-deficient RNAs, such as several K variants, the ability to undergo such secondary conformational changes appears to be defective.

consequences of structural diversity in the apical stems of the Dε, Hε, and L5 RNAs could be accounted for if, upon binding to P protein, all three were forced into a new common, or at least similar, conformation (Fig. 8B). There is accumulating evidence for the general nature of such induced-fit mechanisms in RNA-protein interactions affecting both the RNA and the protein (e.g., reference 1). Sufficient flexibility in an RNA's structure to accommodate such a rearrangement may therefore be another criterion for binding to P protein.

A mechanistic alternative is that of several preexisting RNA conformers, only the proper one is recruited into the complex with P protein. While our data do not explicitly exclude such an interpretation, we consider this possibility less likely. First, the native PAGE assays give no positive evidence for the existence of stable alternative structures; this result is particularly evident from the markedly different mobilities of wt Dε and the priming-competent variant V3 and from the absence of any detectable band that might correspond to a common, active conformation. Second, a rapid equilibrium between substantially different alternative structures giving rise to a single band is unlikely for kinetic reasons. While minor structural changes may take place with fast kinetics, rearrangements involving the breaking and reformation of several hydrogen bonds occur very slowly under conditions of low concentrations of salt and low temperature, as were used in the native PAGE assays (7a). Furthermore, we have recently obtained direct evidence that the apical structure of wt Dε RNA bound to P protein is substantially different from that in free RNA (3a).

Finally, productive RNA binding has recently been reported to alter the protease sensitivity of the DHBV P protein (29), suggesting a structural alteration in the protein. From these and the above-described results, we envisage that productive ε-P protein interaction is a multistep process. First, binding and nonbinding RNAs are discriminated against because of their overall shape and certain key residues. While the structure of binding-competent RNAs (and protein) may already change at this stage, a second structural alteration of the RNA, and the protein, is required to make the entire complex priming competent (Fig. 8B). In this scenario, the binding-competent K variants fulfil the first criterion but are unable to support the switch to a priming-competent conformation. A direct comparison of priming-competent and incompetent ε RNAs in the complex with P protein should reveal the postulated structural differences and hence allow us to test this model.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft through SFB 229 and the Fonds der Chemischen Industrie.

We thank H. Schaller for critical discussions and support and H. Bartos for providing plasmids pT7AMVpol16, pDεL4, pDεL5, and pDεL4,5.

REFERENCES

1. **Allain, F. H., C. C. Gubser, P. W. Howe, K. Nagai, D. Neuhaus, and G. Varani.** 1996. Specificity of ribonucleoprotein interactions determined by RNA folding during complex formation. Nature **380:**646–650.
2. **Bartos, H.** 1995. Ph.D. thesis. University of Heidelberg, Heidelberg, Germany.
3. **Beck, J., and M. Nassal.** 1996. A sensitive procedure for mapping the boundaries of RNA elements binding in vitro translated proteins defines a minimal hepatitis B virus encapsidation signal. Nucleic Acids Res. **24:**4364–4366.

3a. **Beck, J., and M. Nassal.** Unpublished data.

4. **Beck, J., H. Bartos, and M. Nassal.** 1997. Experimental confirmation of a hepatitis B virus (HBV) ε-like bulge-and-loop structure in avian HBV RNA encapsidation signals. Virology **227:**500–504.
5. **Bhattacharyya, A., A. I. H. Murchie, and D. M. J. Lilley.** 1990. RNA bulges and the helical periodicity of double-stranded RNA. Nature **343:**484–487.
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. **Draper, D. E.** 1995. Protein-RNA recognition. Annu. Rev. Biochem. **64:**593–620.

7a. **Grüne, M., J. P. Fürste, S. Klußmann, V. A. Erdmann, and L. R. Brown.** 1996. Detection of multiple conformations of the E-domain of 5S rRNA from Escherichia coli in solution and in crystals by NMR spectroscopy. Nucleic Acids Res. **24:**2592–2596.

8. **Hu, J., and C. Seeger.** 1996. Hsp90 is required for the activity of a hepatitis B virus reverse transcriptase. Proc. Natl. Acad. Sci. USA **93:**1060–1064.
9. **Jaffrey, S. R., D. J. Haile, R. D. Klausner, and J. B. Harford.** 1993. The interaction between the iron-responsive element binding protein and its cognate RNA is highly dependent upon both RNA sequence and structure. Nucleic Acids Res. **21:**4627–4631.
10. **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.
11. **Karn, J., M. J. Gait, M. J. Churcher, D. A. Mann, I. Mikaelian, and C. Pritchard.** 1995. Control of human immunodeficiency virus gene expression by the RNA binding proteins Tat and Rev, p. 192–220. *In* K. Nagai and I. Mattaj (ed.), Frontiers in molecular biology: RNA-protein interactions. Oxford University Press, Oxford, United Kingdom.
12. **Knapp, G.** 1989. Enzymatic approaches to probing of RNA secondary and tertiary structure. Methods Enzymol. **180:**192–212.
13. **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.
14. **Kuchino, Y., and S. Nishimura.** 1989. Enzymatic RNA sequencing. Methods Enzymol. **180:**154–163.
15. **Lanford, R. E., L. Notvall, and B. Beames.** 1995. Nucleotide priming and reverse transcriptase activity of hepatitis B virus polymerase expressed in insect cells. J. Virol. **69:**4431–4439.
16. **Laskus, T., J. Rakela, and D. H. Persing.** 1994. The stem-loop structure of the cis-encapsidation signal is highly conserved in naturally occurring hepatitis B virus variants. Virology **200:**809–812.
17. **Mattaj, I. W.** 1993. RNA recognition—a family matter? Cell **73:**837–840.
18. **Nassal, M.** 1996. Hepatitis B virus morphogenesis. Curr. Top. Microbiol. Immunol. **214:**297–337.
19. **Nassal, M., and H. Schaller.** 1996. Hepatitis B virus replication—an update. J. Viral Hepatitis **3:**217–226.
20. **Nassal, M., and A. Rieger.** 1996. A bulged region of the hepatitis B virus RNA encapsidation signal contains the replication origin for discontinuous first-strand DNA synthesis. J. Virol. **70:**2764–2773.

21. **Pollack, J. R., and D. Ganem.** 1993. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. J. Virol. **67:**3254–3263.
22. **Pollack, J. R., and D. Ganem.** 1994. Site-specific RNA binding by a hepatitis B virus reverse transcriptase initiates two distinct reactions: RNA packaging and DNA synthesis. J. Virol. **68:**5579–5587.
23. **Rieger, A., and M. Nassal.** 1995. Distinct requirements for primary sequence in the 5′- and 3′-part of a bulge in the hepatitis B virus RNA encapsidation signal revealed by a combined *in vivo* selection/*in vitro* amplification system. Nucleic Acids Res. **23:**3309–3315.
24. **Rieger, A., and M. Nassal.** 1996. Specific hepatitis B virus minus-strand DNA synthesis requires only the 5′ encapsidation signal and the 3′-proximal direct repeat DR1*. J. Virol. **70:**585–589.
25. **Sumner-Smith, M., S. Roy, R. Barnet, L. S. Reid, R. Kuperman, U. Delling, and N. Sonenberg.** 1991. Critical chemical features in *trans*-acting-responsive RNA are required for interaction with human immunodeficiency virus type 1 Tat protein. J. Virol. **65:**5196–5202.
26. **Tavis, J. E., and D. Ganem.** 1993. Expression of functional hepatitis B virus polymerase in yeast reveals it to be the sole viral protein required for correct initiation of reverse transcription. Proc. Natl. Acad. Sci. USA **90:**4107–4111.
27. **Tavis, J. E., S. Perri, and D. Ganem.** 1994. Hepadnavirus reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer. J. Virol. **68:**3536–3543.
28. **Tavis, J. E., and D. Ganem.** 1995. RNA sequences controlling the initiation and transfer of duck hepatitis B virus minus-strand DNA. J. Virol. **69:**4283–4291.
29. **Tavis, J. E., and D. Ganem.** 1996. Evidence for activation of the hepatitis B virus polymerase by binding of its RNA template. J. Virol. **70:**5741–5750.
30. **Valegard, K., J. B. Murray, P. G. Stockley, N. J. Stonehouse, and L. Liljas.** 1994. Crystal structure of an RNA bacteriophage coat protein-operator complex. Nature **371:**623–626.
31. **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.
32. **Wang, G. H., and C. Seeger.** 1993. Novel mechanism for reverse transcription in hepatitis B viruses. J. Virol. **67:**6507–6512.
33. **Wang, G. H., F. Zoulim, E. H. 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.
34. **Weber, M., V. Bronsema, H. Bartos, A. Bosserhoff, R. Bartenschlager, and H. Schaller.** 1994. Hepadnavirus P protein utilizes a tyrosine residue in the TP domain to prime reverse transcription. J. Virol. **68:**2994–2999.
35. **Witherell, G. W., J. M. Gott, and O. C. Uhlenbeck.** 1991. Specific interactions between RNA phage coat proteins and RNA. Prog. Nucleic Acid Res. Mol. Biol. **40:**185–220.
36. **Wu, H. N., and O. C. Uhlenbeck.** 1987. Role of a bulged A residue in a specific RNA-protein interaction. Biochemistry **26:**8221–8227.
37. **Ziermann, R., and D. Ganem.** 1996. Homologous and heterologous complementation of HBV and WHV capsid and polymerase functions in RNA encapsidation. Virology **219:**350–356.
38. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.