A Small 2′-OH- and Base-dependent Recognition Element Downstream of the Initiation Site in the RNA Encapsidation Signal Is Essential for Hepatitis B Virus Replication Initiation*

(Received for publication, August 20, 1999, and in revised form, October 1, 1999)

Susanne G. Schaaf, Jürgen Beck, and Michael Nassal‡

From the Department of Internal Medicine II/Molecular Biology, University Hospital Freiburg, Hugstetter Strasse 55, D-79106 Freiburg, Germany

Hepatitis B viruses replicate through reverse transcription of an RNA intermediate. In contrast to retroviral reverse transcriptases, their replication enzyme, P protein, does not use a nucleic acid primer but initiates DNA synthesis de novo from within an RNA stem-loop structure called e. A short DNA oligonucleotide is copied from e and covalently attached to P protein, and then synthesis is arrested. The information for initiation site selection and synthesis arrest must be contained in the structure of the P protein/e complex. Because P protein activity depends on cellular chaperones this complex can as yet only be generated by in vitro translation of duck hepatitis B virus P protein in rabbit reticulocyte lysate; functional interaction with its cognate RNA element De can be monitored by the covalent labeling of P protein during primer synthesis. Combining this in vitro priming reaction and a set of chimeric RNA-DNA DNA analogues, we found that only five ribose residues in the 57-nucleotide stem-loop were sufficient to provide a functional template; these are a single residue in the template region and the two base pairs at the tip of the lower stem. The base identities in the very same region are essential as well. The presence of this 2′-OH- and base-dependent determinant shortly downstream of the initiation site suggests a mechanism that can account for both initiation site selection and programmed primer synthesis arrest.

Hepatitis B virus (HBV),1 the causative agent of B-type hepatitis in humans, is the type-member of the hepadnaviruses. These small DNA-containing viruses replicate by reverse transcription of an RNA intermediate, the pregenomic RNA (pgRNA). Their replication mechanism is unique among the retroid elements (1–4). It involves selective packaging of the pgRNA by the reverse transcriptase, called P protein, into nucleocapsids (5), and initiation of reverse transcription without a nucleic acid primer (6). Specific recognition of pgRNA over other viral or cellular RNAs is effected by the interaction between P protein and a 5′-proximal stem-loop structure on pgRNA that is unique to this RNA species (7, 8). This RNA

1 The abbreviations used are: HBV, hepatitis B virus; DHBV, duck hepatitis B virus; pgRNA, pregenomic RNA; DR1*, direct repeat 1*; e, encapsidation signal; De, encapsidation signal of DHBV; wt, wild-type.

* This work was supported by Grant DFG Na154/4-1 from the Deutsche Forschungsgemeinschaft, by the Center for Clinical Research I of the University Hospital Freiburg, and by the Fonds der Chemischen Industrie. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 49-761-270-3507; Fax: 49-761-270-3507; E-mail: nassal2@ukl.uni-freiburg.de.
Fig. 1. A, the mechanism of replication initiation in hepatitis B viruses. The straight black line symbolizes a linearized version of the circular DHBV DNA genome; numbers represent nucleotide positions; diamond represents the polyadenylation signal. The positions of the open reading frames encoding the core protein (C), the surface proteins (preS and S), and P protein (P) are indicated by the open boxes. The wavy line depicts the pgRNA, a capped and polyadenylated transcript comprising a unit length genome plus a terminal redundancy, because of which a second copy of the sequence containing the encapsidation signal (De) and the DR1 is present at the 3′-end. pgRNA serves first as bicistronic mRNA for core (small spherical objects) and P protein (sphere marked P). P protein binding to the 5′-copy of De initiates nucleocapsid assembly (left) and initiation of (−)-DNA synthesis (right). Part of a bulged region within De is copied into a short DNA oligonucleotide that becomes covalently linked to the enzyme. After translocation of the complex to the 3′-proximal DR1⁎, the oligonucleotide is used as primer for a complete (−)-DNA strand. The associated chaperones that are essential for these reactions and (−)-DNA synthesis are not shown. The general process is essentially the same in human HBV replication. B, secondary structures of HBV and DHBV encapsidation signals. The overall structures of HBV e and De are similar, with a characteristic central bulge separating the lower and upper stem. Encircled nucleotides indicate positions that were deleted or changed as indicated in the synthetic RNA-DNA De analogues. The template region is denoted by larger lettering; the DNA primer is shown on the right (De bound), covalently bound via its 5′-end to a Tyr residue in P protein. For De, it is known that primer synthesis requires a major structural alteration by P protein binding that involves opening of the upper stem. Gray ovals show RNA regions that are protected from Pb²⁺ cleavage in the protein complex, whereas the black oval indicates a region that becomes hypersensitive to single-strand specific reagents.

Fig. 2. In vitro reconstitution of (−)-DNA initiation. DHBV P protein is in vitro translated in reticulocyte lysate programmed with an artificial mRNA that only contains the P open reading frame but neither the 5′- nor the 3′-copy of De. Instead, De RNA or appropriate analogues are used as templates for DNA primer synthesis. As a consequence of a productive RNA-protein interaction, this trans-priming reaction will result in covalent labeling of the P protein if one of the dNTPs is labeled in the α-position. Labeling of the protein can subsequently be monitored by SDS-polyacrylamide gel electrophoresis and autoradiography.

particular of the bulge region, prevent binding of P protein and consequently priming. Alternatively, point mutations at some strategic positions, e.g., U-2590 in the loop, inhibit binding despite conservation of the authentic structure. Finally, formation of a functional P protein/De initiation complex is accompanied by a major structural alteration in the apical stem, induced by P protein binding (Fig. 1B) (23), and RNA variants unable to undergo this conformational change are not accepted as templates for initiating reverse transcription (23). Hence, overall structure, base identities at specific positions in the upper stem, and RNA flexibility contribute in a complex fashion to a productive interaction between De and P protein.

Given the size of the RNA element and the probably multiple protein components, direct structural analyses of the hepadnaviral initiation complex are not in sight. We therefore exploited the in vitro translation system to further characterize the De/P protein interaction. As a reverse transcriptase, P protein can use both RNA and DNA as templates. Hence, we first asked whether (and if so to what extent) the RNA nature of De is important for its interaction with the P protein. To this end, we used the trans-priming assay to functionally analyze synthetic chimeric De molecules in which increasing portions of the ribose phosphate backbone were substituted by their deoxyribose counterparts. Surprisingly little of the entire structure, i.e., only the template region and the two base pairs at the tip of the lower stem, was required to consist of RNA. Based on these results, we then tested whether the nature of the bases at these positions was essential. For this, we used in vitro transcribed all-ribo molecules in which increasing parts of the lower stem of De were replaced by either the corresponding regions from HBV e, or simply by G-C pairs. Convergent with the first data set, we found that conservation of two to three authentic base pairs at the tip of the lower stem was necessary and sufficient for a productive interaction with P protein. Hence, whereas most of the lower stem simply needs to be base paired, the presence of 2′-hydroxyls, and of specific bases at the tip of the lower stem underlying the bulge, is of utmost importance. This indicates that this small region contains a principal determinant for the interaction with the hepadnaviral reverse transcriptase, most likely because it forms, together with the bulge, a distinct three-dimensional structure and/or is involved in direct protein contacts. The structural and functional implications of these findings are discussed in the light of recently solved structures of other RNA-protein complexes.

EXPERIMENTAL PROCEDURES

In Vitro Synthesis of De RNA—All-ribo De molecules were obtained by in vitro transcription (T7 MEGAscript kit; Ambion) according to the manufacturer’s recommendations. After in vitro transcription, template DNA was digested with DNase I, and the RNA was extracted with phenol, precipitated with isopropanol, and resuspended in TE buffer (10 mM Tris·Cl, 1 mM EDTA, pH 7.5). Plasmid pDel (25) served as template for the synthesis of wild-type De RNA used as reference in experi-
ments with RNA/DNA chimeras. The resulting transcript is 76 nucleotides in length and contains DHBV sequence from position 2557 to 2624. Wild-type De RNA used as reference in experiments with sequence variants of De was transcribed from plasmid pBDwt containing DHBV16 sequence from position 2520 to 2652 cloned into the SalI and EcoRV sites of pBluescript II SK(+) (Stratagene). Plasmids coding for De variants H/D0, H/D1, H/D2, H/D2.5, H/D3, and synD were obtained by polymerase chain reaction-based mutagenesis using pBDwt as template, and their identities were confirmed by DNA sequencing. Plasmids were linearized with EcoRV and transcribed with T7 RNA polymerase yielding RNA transcripts of 170 nucleotides in length.

Chemical Synthesis and Purification of RNA/DNA Chimeras—Oligonucleotides were synthesized on an ABI synthesizer using standard phosphoramidite chemistry; for ribonucleotides, 2'-O-t-butyldimethylsilyl ether protection was used (28). Thymidines in DNA moieties within chimeric De molecules were substituted by 2'-deoxyuridines to avoid negative effects of the additional 5'-methyl group. After removal of the protecting groups, the chimeric oligonucleotides were precipitated using n-butanol and subsequently purified by electrophoresis through denaturing 12% polyacrylamide gels. Full-length products were detected by UV-shadowing, cut out, eluted, precipitated with isopropanol, and resuspended in TE buffer. Preparations were analyzed for purity and concentration by electrophoresis on 2% agarose gels, and photometrically assuming that 1 A260 nm equals 40 μg/ml RNA. All chimeric De analogues, 53 nucleotides in length, contain the authentic De sequence with a shortened lower stem from DHBV16 position 2561 to 2615 (Fig. 1B). Typical yields of a 200-nmol synthesis were in the range of 100–150 μg of oligonucleotide.

In Vitro Trans Priming Assay—DHBV P protein was in vitro synthesized in a coupled in vitro transcription/translation system (TNT T7 quick coupled transcription/translation system; Promega) from pT7AMVpolL3′e. This plasmid contains the complete P protein open reading frame and was obtained by deletion of the 3′-proximal copy of De in pT7AMVpol (12) in order to exclude cis priming events. Transcription/translation reactions, typically in a total volume of 50–100 μl, were performed according to the manufacturer’s protocol. After 60 min of incubation at 30 °C, the samples were split into 10-μl aliquots, De RNA or synthetic analogues were added, and the samples were incubated for further 60 min at 30 °C to form De RNA/P protein complexes. For the detection of protein priming activity 2.5 μCi of α-[32P]dATP (3000 Ci/mmol) in 1 volume of 2× priming buffer (12) was added, and the samples were incubated for 60 min at 37 °C. Aliquots of each sample were analyzed for radioactively labeled P protein by reducing SDS-polyacrylamide gel electrophoresis on 7.5% polyacrylamide gels. Signal intensities were quantified using a phosphoimager system (BAS1500; Fuji).

RESULTS

Design of Artificial De RNA/DNA Chimeras—The authentic De stem-loop encompasses 57 nucleotides in length (Fig. 1B), a size beyond the range currently accessible with most conventional RNA synthesis equipment. Although a test RNA of 35 nucleotides could be isolated using phosphoramidite chemistry and t-butyldimethylsilyl protection for the 2′-OH groups, attempts to obtain a completely synthetic De RNA did not yield detectable amounts of full-length material (data not shown). Therefore, to minimize length for more efficient synthesis, some nonessential nucleotides were deleted in the further De constructs: in the upper stem and lower stem, U-2596 and the single unpaired U-2610 were removed (indicated in Fig. 1B by encircled nucleotides). Both deletions have been reported to be compatible with the packaging function of De in transfected cells (20). To directly monitor its effect on priming, a corresponding T7 transcript (De RNAΔU, Fig. 3B) was analyzed, at 50 and 500 nM concentrations, side-by-side with wild-type De RNA in the in vitro priming reaction. At these concentrations, wild-type De RNA leads to a dose-dependent labeling of P protein, whereas the signal saturates above 1 μM De RNA (24, 25). At both concentrations, the De RNA ΔU displayed an only 2-fold reduction in priming efficiency (Fig. 3A). To further shorten the sequence, the bottom base pair (positions 2560 and 2616) was deleted, and the penultimate base pair was changed from G-U to G-C (see Fig. 1B). These modifications are known not to affect the De/P protein interaction in vitro (26). The first chimeric molecule, variant DRb1, was designed to contain ribonucleotides in the loop, plus in the apical bulge region (Fig. 3B). Synthesis of this molecule, with 22 ribonucleotides in a total of 53 nucleotide positions, yielded clearly detectable amounts of full-length material that was functionally active in the priming assay (see below). Accordingly, all additional synthetic derivatives were based on this sequence.

![Fig. 3. Ribonucleotides in the De bulge region are necessary and sufficient for priming.](image-url)
All-DNA Analogues of De Are Functionally Inactive but Ribonucleotides in the Bulge and Loop Regions Restore Activity—As a first step toward analyzing the functional importance of the ribose moieties within De we analyzed the performance in trans-priming assays of an unmodified DNA analogue of De, De DNAAT. It corresponded to DHBV nucleotides 2557 to 2617 and hence contained the complete, authentic stem-loop sequence. Neither at 100 nM nor at 1000 nM could any priming signal be detected (data not shown). The same negative result was obtained with the homologous oligonucleotide De DNAAT, which contained deoxyuridine instead of deoxythymidine (Fig. 3A). Hence, complete substitution of all ribose moieties is incompatible with a productive P protein interaction, and the 5-methyl groups of deoxythymidine are not responsible for this inactivity.

Next we asked whether partial restoration of the RNA backbone would reestablish functional activity. Because RNA-specific non-Watson-Crick interactions are frequently found in single-stranded regions, we first focussed on the bulge and its vicinity, and on the apical loop region. The corresponding chimeras, DRbl1 and DRbl1 (Fig. 3B), when tested at 50 and 500 nM concentrations, both stimulated covalent labeling of P protein to essentially the same levels as the reference RNA De RNAAT. This indicated that the principal elements for a productive interaction with P protein reside mainly in the RNA bulge region and its vicinity, whereas in the apical region, the 2'-hydroxyls are of minor, if any, importance.

The Tip of the Lower Stem and the Bulge Harbor the Minimal Ribose-dependent Element for Productive Interaction with P Protein—In DRbl1, the bulge region, the opposite unpaired U and the two adjacent base pairs in the lower and the upper stem consisted of RNA (Fig. 3B). To further narrow down the essential ribonucleotide positions, we used the next set of variants, DRb2 through DRb5 (Fig. 4B), all of which contain a complete ribonucleotide bulge. In DRb2, only the flanking residues on the left side are composed of ribonucleotides. In DRb3 and DRb4, two and one bulge-proximal ribo base pair(s), respectively, in the lower stem are maintained. In DRb5, the two bulge-proximal ribo base pairs in the upper stem consist of ribonucleotides.

Primase assays with these variants at 50 and 500 nM concentration (Fig. 4A) revealed that three of the four chimeras (DRb2, -4, and -5) were severely impaired in their activity. However, DRb3 produced signals with an intensity comparable to that of DRbl1. This indicated the presence of an important ribose-dependent recognition element at the tip of the lower but not the base of the upper stem. Because variant DRb4 with a single ribo base pair at the tip of the lower stem was only marginally active, this element involves the two tip base pairs of the lower stem but not the unpaired U (compare DRb4 with -3).

Combined with Two Ribo Base Pairs at the Tip of the Lower Stem, a Single Ribonucleotide at The Initiation Site Is Necessary and Sufficient for Productive Interaction with P Protein—In all variants analyzed thus far, the bulge consisted completely of RNA. To test whether all of these ribonucleotides are essential, we used the next series of variants in which the two ribo base pairs at the tip of the lower stem were maintained as RNA, and various parts of the bulge itself were replaced by DNA (Fig. 5B).

Variant DRb35 with a complete DNA bulge produced only a very marginal priming signal (about 5% of DRb3), emphasizing the need for the presence of one or more ribonucleotides in the bulge. Restoration of ribonucleotides in the 5'-proximal half of the bulge did not rescue priming activity (DRb32). By contrast, reintroduction of ribonucleotides in the 3' part of the bulge, i.e. the actual template region, restored activity. Both DRb31 and DRb33 led to a similar labeling of P protein as the all-ribo bulge chimera DRb3, and even the chimera DRb34 with a single ribonucleotide at the 3'-terminal position of the bulge, where DNA synthesis most likely initiates, produced a signal of about one-third the intensity seen with De RNAAT. This positive effect on priming efficiency of the single additional 2'-hydroxyl clearly establishes the essential role of 2'-hydroxyls in the template region. Notably, because the priming assays were performed using [α-32P]dATP, incorporation of the label into the protein must have occurred in a U-templated fashion. This suggests that once initiated at a ribonucleotide, P protein is able to proceed on a DNA template during primer synthesis, although the efficiency may be somewhat reduced (compare the signal intensities for variants DRb33 and DRb34). Together, these data demonstrate that as few as 5 ribonucleotides in the entire De RNA element are sufficient for productive interaction with P protein.

Functional De / P Interaction Also Depends on the Base Iden-
tities in the Ribonucleotide Pairs at the Tip of the Lower Stem—Next, we asked whether the nature of the bases at the above defined ribose-dependent positions was also important. The experimental strategy involved the generation of a series of variant De in vitro transcripts containing mutated lower stems. To maintain base pairing while at the same time changing the sequence we first replaced the entire lower De stem with the corresponding part from HBV e (variant H/D0; see Fig. 6B). In accord with a previous report (24), this complete domain swap abolished the template function of De (Fig. 6A). We then progressively reintroduced De-specific base pairs and monitored the priming activities of the corresponding variants. Based on the above described results and the previous finding that the lower stem can be shortened from its base to between 5 and 7 base pairs without losing P protein binding competence (26), we concentrated on the base pairs at the tip of the lower stem.

The corresponding variants H/D1, H/D2, and H/D3 contain one, two, and three authentic De base pairs, respectively, underneath the bulge; in variant H/D2.5 the left-hand C residue in the third pair from the top was changed back from C-G to U-G (see Fig. 6B). As a final control, we used the variant synD in which the 8 bottom base pairs of De are substituted for by 5 G-C pairs (Fig. 6B).

The corresponding in vitro transcripts were used in trans-priming assays at a 500 nM concentration as described above (Fig. 6A). Variant H/D1, with a single De-specific base pair, was as inactive in priming as H/D0 RNA. However, a slight but reproducible increase in signal intensity (to about 5% of the wild-type signal) was observed with variant H/D2 containing 2 authentic De pairs. The activity was significantly increased (to about 20% of wild-type level) when the left-hand C residue in the third pair was changed back from C to U (H/D2.5), and to about 50% of wild-type activity by replacement with the original U-A pair from De (H/D3). Hence, two De base pairs are essential for at least a basal priming activity, and three De base...
pairs are essential for a substantial priming activity. The crucial role of the base identities in these three pairs but not in the base of the lower stem was further corroborated by variant synD which displayed about 80% of the wild-type priming activity. That it performed even better than H/D3 might be due to the facts that the two G-C pairs underneath the three tip base pairs correspond to authentic Dε RNA and differ only by the identity of the bases at the top base pair of the lower stem, which is U-G in Dε and A-U in HBV ε.

**An Essential Role for the G Residue in the U-G Base Pair at the Tip of the Lower De Stem**—One prediction from the above described results was that, in a reverse experiment, replacement of the authentic base pairs at the tip of the lower De stem should interfere with the priming function. Indeed, a corresponding variant, -1A/U (Fig. 7B), containing just the tip A-U base pair of HBV ε in an otherwise completely De-derived context, showed a very low priming activity (about 5% of wild-type level), even at a 500 nM concentration (Fig. 7A). The same was true for variant -1C/U. By contrast, approximately 50% of the wild-type activity was seen with variant -1C/G having the right-hand G residue from De. To confirm the special role of this G residue suspected from this result, we finally combined it with an A residue on the left. The corresponding variant -1A/G exhibited the same activity as the wild-type RNA. Hence, the presence of either U (wild-type Dε), C, or A at the left-hand position of the tip base pair is compatible with significant priming activity, regardless of Watson-Crick pairing, whereas replacement of the authentic right-hand G by U strongly decreases priming. This establishes a dominant discriminating role of this single G residue in a productive interaction with HBV P protein.

**DISCUSSION**

The unique mechanism of replication initiation in hepatitis B viruses requires that the information for exact start site selection by P protein be defined by the specific interaction between the ε RNA template and the reverse transcriptase. Because of the size and multicomponent nature of the initiation complex, direct biophysical analyses of its structure are currently not possible. Using biochemical methods instead, we have here defined a small 2'-OH- and base-dependent determinant at the tip of the lower De stem that, in concert with a single ribose residue in the bulge, is necessary and sufficient for efficient hepatitis B virus replication initiation (Fig. 8). Unusually, this specific element is located downstream of the initiation site.

**Structural Implications for the P Protein/De Interaction**—A comparison with the limited number of other protein-RNA complexes that have been investigated by similar biochemical or, in a few cases, crystallographic methods indicates that the interaction of P protein with De bears some striking similarities to several proteins that recognize RNA hairpins. One example is the coat protein of bacteriophage MS2, which specifically binds to a 19-nucleotide stem-loop, the operator RNA (29). Deoxy substitution experiments defined a ribose requirement in the upper two base pairs of the stem, and at a single position in the four-member loop (30). As proposed and later directly confirmed by x-ray crystallography, the 2'-hydroxyl of this loop residue is involved in hydrogen bonding contacts to the protein (31, 32). The single ribose residue in the template region of the De bulge required for efficient priming may also be in direct contact with P protein but a more subtle effect on its template quality rather than its physical binding ability cannot be excluded at present. It should also be noted that, in contrast to the MS2 operator and the spliceosomal RNAs discussed below, there are probably few base-specific contacts in the single-stranded De bulge. In human HBV, for instance, a substantial variety of bulge sequences is compatible with the template and the encapsidation functions of ε (11, 33).
The ribose moieties at the tip of the operator stem have been suggested to maintain the helix in the A-form conformation observed in the crystal structure (30). For some model compounds, it has been shown that a single ribo-base pair at the end of a double helix can indeed induce a A-like conformation in the entire molecule (34, 35). Hence, the essential ribose residues at the tip of the lower De stem could serve a similar, general function. The simultaneous importance of base identities is, however, compatible with a more specific role. First, the ends of RNA helices provide a rich source for specific recognition because the distortions caused by the adjacent single-stranded regions allow for contacts that are not possible in an ideal A-form RNA with its deep and narrow major groove (36).

Second, there are several instances in which exactly the ends of a helix harbor important specificity determinants enabling highly similar RNAs to be accurately discriminated by very similar proteins. The cognate RNA elements of the spliceosomal proteins U1A and the U2B'-U2A complex (U1 snRNA hairpin II and U2 snRNA hairpin IV) share not only a common secondary structure but also almost identical sequences in the single-stranded loops. A key role for selective recognition is played by the loop-closing base pairs, C-G in the U1 and the noncanonical U-U pair in the U2 snRNA (37, 38). The underlying stems contribute to the specific structure adopted by the closing base pairs; in addition, the rigidly structured 3'-part of the loop in the U2 snRNA allows for direct interactions between the RNA stem and the second protein component in the complex. Hence, we propose that the tip of the lower De stem, and probably at least part of the bulge, harbor a similarly important specificity determinant. This view is supported by the crucial importance of the right-hand G residue at the tip De base pair, the counterpart of which in U1 snRNA makes specific contacts to the protein. Such a discriminating role of G-2605 in De could plausibly explain why exchange of this single nucleotide for the HBV-specific U residue abolishes functional interaction with DHBV P protein. The base dependence at the next base pairs further down in the lower stem implies that also this part of the structure is involved in protein contacts.

Functional Implications for Replication Initiation in Hepatitis B Viruses—In addition to the specific ribose- and base-dependent recognition element at the tip of the lower De stem defined in this study, the upper stem and the loop are also important for a productive interaction as established by previous mutational analyses (27). We have not found any ribose requirement in this apical region that, overall, is also relatively tolerant toward base alterations, given they do not grossly alter the conformation of the free RNA and are compatible with formation of the largely unpaired structure induced by P protein binding. There are, however, some positions at which a single base exchange can completely abrogate the interaction with P protein, for instance a U to C mutation at position 2590 in the loop (27). This demonstrates that also the apical De region contains specific recognition elements. Hence, the actual replication initiation site is bracketed between specific interaction sites for the protein.

Two distinct features of hepadnaviral replication are de novo initiation and a programmed synthesis arrest. Only after the covalent complex of P protein with the self-made primer is translocated to DR1* does DNA synthesis resume. All other reverse transcriptases, like DNA polymerases in general, require a nucleic acid primer. Therefore, the known structures from retroviral systems, for instance HIV-1 RT primer/template complexes (39, 40), are of limited comparative value. The only known other exception is the unusual Mauriceville retroplasmid reverse transcriptase, which can initiate DNA synthesis from the penultimate residue at the 3'-end of the RNA template. Its positioning over the initiation site is mediated by a tRNA-like structure (41), a mechanism closely resembling that employed by viral RNA-dependent RNA polymerases. Unfortunately, little structural information is available about the corresponding initiation complexes. The other large class of polymerases capable of de novo initiation comprises the DNA-dependent RNA polymerases. The recent high resolution structure of an initiation complex formed by their simplest representative, the single subunit enzyme from bacteriophage T7 (42), has revealed how this protein achieves specific initiation site selection without a primer. Three features for proper positioning are (i) specific tight interactions with distinct upstream sequences within the double-stranded T7 promoter, (ii) melting of the template by site-specific contacts with the template strand immediately before the initiation site to create a transcription bubble, and (iii) generation of a distinct structure in the template strand such that the first template

![Fig. 9. Comparison of primer-less synthesis initiation by hepadnaviral P proteins and T7 RNA polymerase.](image)
nucleotide, dC, can optimally base pair when the NTP binding pocket is occupied with GTP (Fig. 9).

Although the hepadnaviral template is RNA rather than DNA, there are several parallels and one major difference: like the phage enzyme, P protein is directed selectively to the template region by, in this case, RNA- instead of DNA-specific recognition elements. The bulged structure of the template region resembles a transcription bubble already in free DNA; however, most likely, the constraints imposed on the region resembles a transcription bubble already in free DNA, there are several parallels and one major difference: like the phage enzyme, P protein is directed selectively to the template region.

We thank A. Bosserhoff (Center for Molecular Biology, Heidelberg, Germany) for oligonucleotide synthesis and A. Wahl-Feuerstein for excellent technical assistance.

REFERENCES

1. Nossal, M., and Schaller, H. (1993) Trends Microbiol. 1, 221–228
2. Nossal, M., and Schaller, H. (1996) J. Virol. Hepat. 3, 217–226
3. Seeger, C., and Hu, J. (1997) Trends Microbiol. 5, 447–450
4. Nossal, M. (1999) Interferon 42, 103–116
5. Bartenschlager, R., Junker-Niepmann, M., and Schaller, H. (1990) J. Virol. 64, 5324–5332
6. Wang, G. H., and Seeger, C. (1992) Cell 71, 663–670
7. Bartenschlager, R., and Schaller, H. (1992) EMBO J. 11, 3413–3420
8. Junker-Niepmann, M., Bartenschlager, R., and Schaller, H. (1990) EMBO J. 9, 3289–3306
9. Wang, G. H., and Seeger, C. (1993) J. Virol. 67, 6507–6512
10. Tavis, J. E., Perri, S., and Ganem, D. (1994) J. Virol. 68, 5336–5343
11. Nossal, M., and Rieger, A. (1996) J. Virol. 70, 2764–2773
12. Weber, M., Brunessen, V., Bartos, H., Bosserhoff, A., Bartenschlager, R., and Schaller, H. (1994) J. Virol. 68, 2994–2999
13. Zoulim, F., and Seeger, C. (1994) J. Virol. 68, 6–13
14. Lanford, R. E., Notvall, L., and Beames, B. (1995) J. Virol. 69, 4431–4439
15. Holt, S. E., Aisner, D. L., Baur, J., Tesmer, V. M., Dy, M., Ouellette, M., Trager, J. B., Morin, G. B., Toft, D. O., Shay, J. W., Wright, W. E., and White, M. A. (1999) Genes Dev. 13, 817–826
16. Bohan, S. P., Kralil, A., and Yamamoto, K. R. (1995) Science 268, 1303–1304
17. Pratt, W. B., and Toft, D. O. (1997) Endocur. Rev. 18, 306–360
18. Pollack, J. R., and Ganem, D. (1995) J. Virol. 69, 3254–3263
19. Knapp, T., and Nossal, M. (1993) Nucl. Acids Res. 21, 3967–3975
20. Pollack, J. R., and Ganem, D. (1994) J. Virol. 68, 5579–5587
21. Hu, J., and Seeger, C. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 1060–1064
22. Hu, J., Toft, D. O., and Seeger, C. (1997) EMBO J. 16, 58–68
23. Beck, J., and Nossal, M. (1998) Mol. Cell. Biol. 18, 6262–6272
24. Wang, G. H., Zoulim, F., Leber, E. H., Kitson, J., and Seeger, C. (1994) J. Virol. 68, 8437–8442
25. Beck, J., Bartos, H., and Nossal, M. (1997) Virology 270, 500–504
26. Beck, J., and Nossal, M. (1996) Nucl. Acids Res. 24, 4346–4346
27. Beck, J., and Nossal, M. (1997) J. Virol. 71, 4971–4980
28. Usga, N., Egl, M., and Rich, A. (1992) Nucl. Acids Res. 20, 6695–6699
29. Witherell, G. W., Gott, J. M., and Uhlenbeck, O. C. (1991) Prog. Nucl. Acids Res. Mol. Biol. 40, 185–220
30. Baidya, N., and Uhlenbeck, O. C. (1995) Biochemistry 34, 12363–12368
31. Valesgard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994) Nature 371, 623–626
32. van den Worm, S. H., Stonehouse, N. J., Valegard, K., Murray, J. B., Walton, C., Fridborg, K., Stockley, P. G., and Liljas, L. (1994) Nature 371, 623–626
33. Rieger, A., and Nossal, M. (1995) Nucl. Acids Res. 23, 3909–3915
34. Egl, M., Usman, N., and Rich, A. (1993) Biochemistry 32, 3221–3227
35. Ban, C., Ramakrishnan, B., and Sundaralingam, M. (1994) J. Mol. Biol. 236, 275–285
36. Steitz, T. A. (1993) in The RNA World (Gesteland, R. F., and Atkins, J. F., eds) pp. 219–228, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
37. Oubridge, C., Itou, N., Evans, P. R., Too, C. H., and Nagai, K. (1994) Nature 372, 432–438
38. Price, S. R., Evans, P. R., and Nagai, K. (1998) Nature 394, 645–650
39. Jacobs-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P. et al. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6210–6214
40. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Science 282, 1669–1675
41. Chen, B., and Lambowitz, A. M. (1997) J. Mol. Biol. 271, 311–332
42. Cheetham, G. M., Jeruzalmi, D., and Steitz, T. A. (1999) Nature 399, 80–83