JOURNAL OF VIROLOGY, Dec. 1991, p. 6438–6446
0022-538X/91/126438-09$02.00/0

Vol. 65, No. 12

# Regulation of Polyadenylation of Hepatitis Delta Virus Antigenomic RNA

SEN-YUNG HSIEH AND JOHN TAYLOR*

*Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111*

Received 29 May 1991/Accepted 24 August 1991

**Hepatitis delta virus (HDV) is a subviral agent with a small RNA genome that is replicated in the nucleus of an infected cell. During genome replication, there is the synthesis of a complementary RNA, known as the antigenome, and also of a smaller complementary species that is polyadenylated and acts in the cytoplasm as the mRNA for the only known HDV protein, the delta antigen. We have carried out an examination of the *cis*- and *trans*-acting elements that regulate the polyadenylation process involved in the synthesis of this mRNA for the delta antigen. Our experimental approach has been to study the processing of nascent antigenomic RNA as it occurs in transfected cells via DNA-directed RNA synthesis, in the absence of genome replication. Three conclusions have been made. (i) The polyadenylation process occurs independent of the functionality of a unique self-cleavage domain located just 3′ of the polyadenylation site. (ii) RNA transcripts that proceed beyond the polyadenylation site can be stabilized by the self-cleavage reaction. Thus, a single transcription initiation event can lead not only to the mRNA species but also to at least one more stable RNA species. (iii) If the nascent RNA species can fold on itself, into the so-called rodlike structure, then the presence of the delta antigen leads to a major suppression of polyadenylation. These results are incorporated into a more detailed model of the replication of the HDV genome.**

The genome of hepatitis delta virus (HDV) is a 1,700-base single-stranded RNA that is replicated in the nucleus of an infected cell via the RNA-directed RNA synthesis with a complementary RNA intermediate known as antigenomic RNA (5). Both the genome and antigenome actually have a covalently closed circular conformation and, because of an unusual self-complementarity, are each able to fold into an unbranched rodlike structure, with about 70% of the bases paired (14, 28). Moreover, each of these RNAs has a single site at which, at least in vitro, there can occur, in the absence of any protein, a specific self-cleavage event (15, 22). This process can be reversed, to give a self-ligation event (23). It has been speculated that the self-cleavage and self-ligation are part of a rolling-circle model of HDV genome replication (23, 27). Even though the process is very much like that for certain pathogenic RNA agents of plants (25, 27), there is one major difference. HDV encodes a protein, the delta antigen, that not only is packaged into progeny virions but also is essential for the replication of the genome (13).

The open reading frame for the delta antigen is not on the genomic RNA but rather on the antigenomic RNA (14, 28, 29). In a previous study, we detected and characterized an additional RNA species that is likely to be the mRNA for the delta antigen (11). This relatively minor species is about 800 bases in length, antigenomic, polyadenylated, and located in the cytoplasm. We mapped both the 5′ end of this RNA and also the site at the 3′ end where polyadenylation occurs. Figure 1 shows those sequences of antigenomic RNA downstream of the open reading frame for the delta antigen. The experimentally determined polyadenylation site, as indicated, is just 3′ of the sequence AAUAAA, which we showed to be essential for HDV genome replication (11). Thus, three features of HDV polyadenylation are reminiscent of the consensus picture for polyadenylation in animal cells of RNA transcribed off DNA by polymerase II (pol II): (i) the essential AAUAAA signal sequence and its location just 5′ of the polyadenylation site, (ii) the actual CA polyadenylation site, and (iii) the G+U-rich sequence just 3′ of the polyadenylation site (3, 18, 19). These similarities are probably more than just coincidence since the evidence is accumulating that HDV, even though it replicates by RNA-directed RNA synthesis, may actually do so by redirection of the host pol II (24). Certainly such redirection has already been claimed for viroids, which are a class of subviral agents of plants which are very much analogous to HDV (as cited in references 25 and 27). There is, however, a striking difference between HDV and the viroids. The viroids are entirely dependent on the host cell and encode no proteins. In contrast, HDV encodes one protein and, as mentioned above, requires that at least some of the antigenomic RNA transcripts become processed so as to form a cytoplasmic mRNA. However, there must be some transcripts that are not so processed and that become full-length antigenomic circular RNAs, which in turn act as the templates for the synthesis of new genomic RNA. In an attempt to understand how this might be achieved, we posed three questions. (i) What is the relevance for polyadenylation of the sequences in the vicinity of the polyadenylation site, especially the nearby self-cleavage domain? (ii) Can a nascent RNA transcript, subsequent to the polyadenylation, continue into the downstream sequences so as to produce a second stable RNA product? (iii) How is it that the polyadenylation process can be suppressed to allow the synthesis of full-length antigenomic RNA?

To answer these questions, we studied the processing of HDV antigenomic RNA as synthesized, in transfected cells, via DNA-directed RNA synthesis by RNA pol II. The rationale for this approach was that it enabled us to study RNA processing in the absence of genome replication. It was considered justified in that, as mentioned above, the available data implicated pol II in HDV RNA synthesis. Also, we already knew that when HDV cDNA spanning the region shown in Fig. 1 was expressed in a transfected cell via

* Corresponding author.

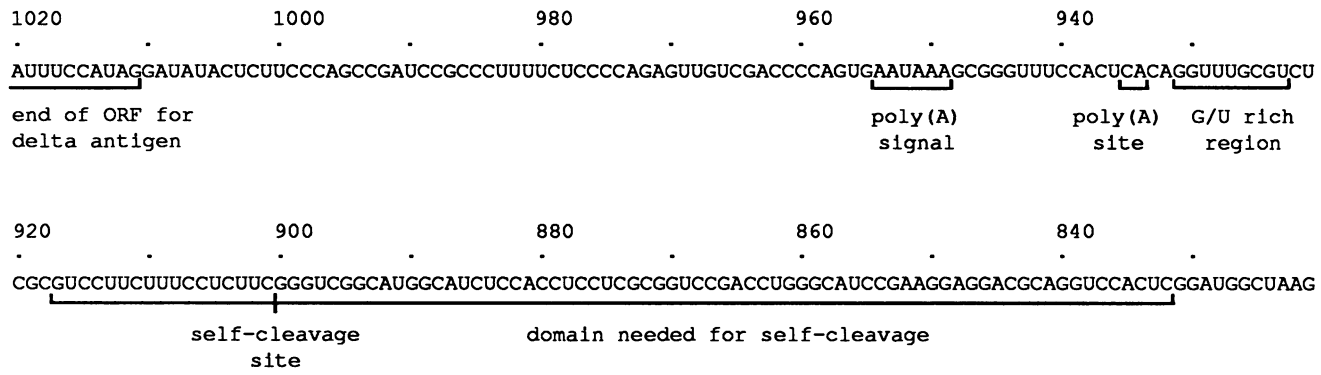


FIG. 1. Sequence features surrounding polyadenylation site on antigenomic HDV RNA. A total of 200 nucleotides of antigenomic HDV RNA, using the sequence of Kuo et al. (14), contains as indicated, the end of the open reading frame for the 195-amino-acid delta antigen, the poly(A) site and signal, as determined by Hsieh et al. (11), and the self-cleavage site and essential flanking domain, as determined by Kuo et al. (15).

DNA-directed synthesis, not only did the polyadenylation signal function but also the polyadenylation site was the same (11).

Our answers to these above questions have made it possible to offer a more detailed model of the processing of nascent HDV RNAs that occurs during genome replication.

## MATERIALS AND METHODS

**Plasmids and transfections.** Cloned HDV cDNA sequences (14) were inserted into both the transcription vector pGEM4Z (Promega) and the eucaryotic expression vector pSVL (Pharmacia) as previously described (11, 13). The cDNA sequences were modified by standard procedures (20). All plasmid DNAs were purified by equilibrium centrifugation in cesium chloride (20). COS7 cells (9) were transfected by using DEAE-dextran (7).

**Oligonucleotides.** The four primers used for anchored polymerase chain reaction (PCR) (8) (see Fig. 2A) were as follows: A, 5′-GACTCGAGTCGACATCGA$(T)_{17}$-3′ (8); B, 5′-GACTCGAGTCGACATCGA-3′ (8); C, 5′-TGGGGGGT GTAACTTCGAAGGTTGGATCGAGGAG-3′ (positions 1152 to 1115); and D, 5′-CTCTTCCCAGCCGATCCC-3′ (positions 1014 to 987).

**RNA extraction.** In most cases, RNA was extracted from transfected cells by digestion with pronase in the presence of sodium dodecyl sulfate, followed by extractions twice with phenol and twice with ether. RNA to be used for RNase protection assays or for isolation of polyadenylated RNA on oligo(dT)-cellulose (1) was extracted via guanidine isothiocyanate followed by equilibrium centrifugation in cesium chloride (6). All RNA samples were given an additional treatment with RNase-free DNase (Promega) and reextracted prior to use in subsequent assays.

**PCR.** Anchored PCR assays were based on the procedure of Frohman et al. (8), as previously adapted by Hsieh et al. (11). To make the procedure quantitative as well as sensitive, we applied the strategy of Zack et al. (31), which makes use of an end-labeled primer, experimentally determined limiting amounts of input cDNA, and a reduced number of cycles of PCR amplification.

**Northern (RNA) analyses.** RNA samples were glyoxalated prior to electrophoresis into gels of 1.5% agarose. The RNA was then transferred electrophoretically to a nylon membrane and then hybridized one or more times as previously described (11) with specific RNA probes. The filter was subjected to autoradiography in the presence of two intensifying screens and then quantitated directly with an AMBIS radioanalytic imaging system to determine the amount of radioactivity in specific regions.

**RNase protection assays.** A labeled RNA probe was synthesized as described in the legend to Fig. 5A and used along with unlabeled RNAs from transfected cells, following the manufacturer's instructions for an RNase protection assay (Ambion).

## RESULTS

**Features on HDV antigenomic RNA that facilitate polyadenylation.** The aim of our first experiments was to investigate further the consequences of sequence alterations in the vicinity of the normal polyadenylation site, along with attempting to see whether the presence or absence of the delta antigen, which is known to be somehow needed for HDV genome replication (13), might have an effect on the polyadenylation process.

To do this, we inserted a cDNA representing part of the HDV sequence into the eucaryotic expression vector pSVL (Fig. 2A, lines 1 and 2). When this plasmid was transfected into animal cells, we expected to find polyadenylation either at the HDV site or at the downstream simian virus 40 (SV40) site of the vector. To assay these two polyadenylated species, we used a modified PCR procedure, referred to as anchored PCR (16) or 3′-RACE (8). The poly(A)-containing RNA from the transfected cells was reverse transcribed with primer A (Fig. 2A, line 3). The cDNA products containing HDV sequences were then specifically amplified by PCR using primer B, which is HDV specific, and radioactive primer C. The products so obtained were analyzed on a denaturing polyacrylamide gel (Fig. 2B). As expected, no specific product was amplified with use of RNA from untransfected cells (lane U). For the transfected cells, we observed a band at about 260 bases (lane 1). This band was heterogeneous in length, presumably because of different cDNA initiation sites within the poly(A) tail of the mRNA. As part of this experiment, we analyzed the RNA from three other transfections. One of these was to use a construct with the AAUAAA motif of the HDV cDNA changed to UUUAAA. As previously reported, such a mutation renders an HDV cDNA clone incapable of initiating HDV genome replication (11). In this assay of polyadenylation, the mutation caused a decrease in polyadenylation at the HDV site

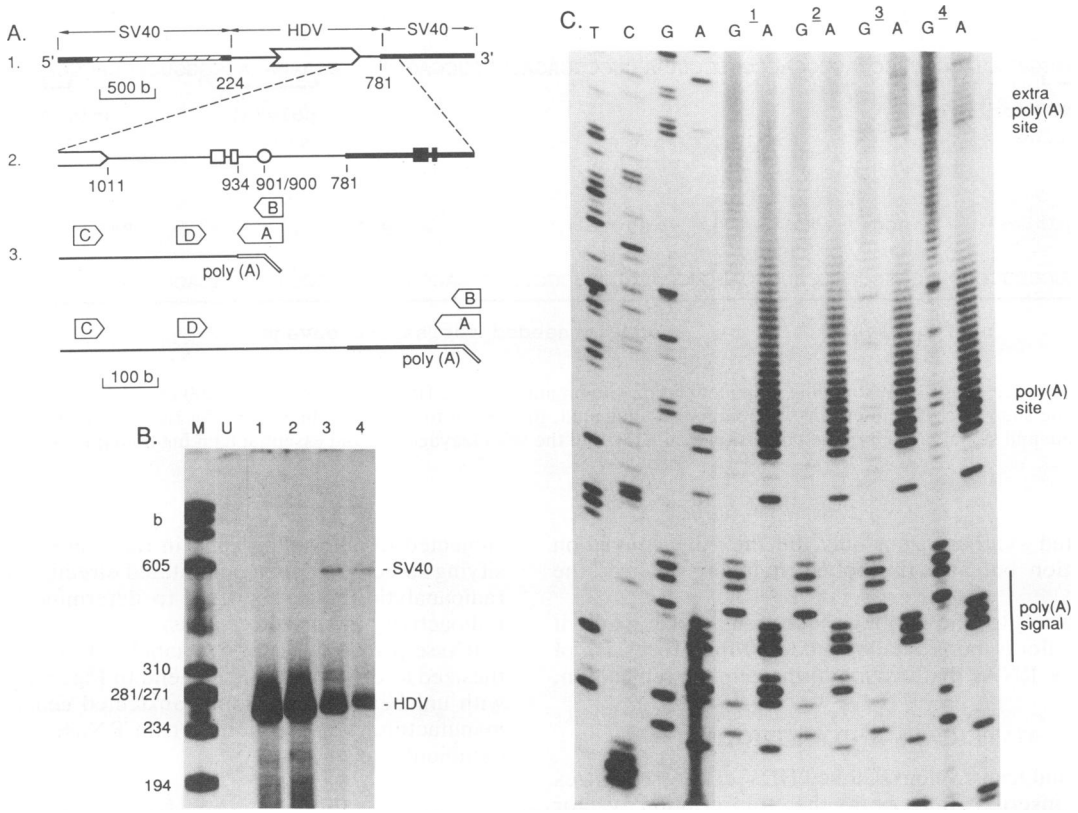


FIG. 2. Dependence of HDV polyadenylation on the poly(A) signal and effect of the delta antigen. (A) Region of HDV cDNA as inserted into the eucaryotic expression vector pSVL. The flanking SV40 sequences are indicated by the solid line, with one intronic region indicated by cross-hatching. On the HDV sequences, the open arrow indicates the open reading frame for the delta antigen. Line 2 amplifies the region around the HDV polyadenylation signal (open square), polyadenylation site (open rectangle), and self-cleavage site (open circle), including the corresponding signal for SV40 (solid symbols). The numbers refer to positions on the HDV sequence of Kuo et al. (14). Line 3 indicates the two predicted polyadenylated mRNAs along with the four oligonucleotide primers, A to D, defined in Materials and Methods. COS7 cells were transfected with the construct shown in line 1, and after 3 days the RNA was extracted and assayed for HDV-related polyadenylated RNA by using anchored PCR, with primer A for reverse transcription followed by primer B and end-labeled primer C for PCR, as previously described (11); the resultant gel analysis of the PCR products is shown in panel B. Location of the bands expected for polyadenylation at the HDV and SV40 sites are indicated at the right. End-labeled *Hae*III fragments of phage $\phi$X DNA provided size markers (lane M), with the sizes (in bases [b]) shown at the left. Anchored PCR was carried out on the RNA from untransfected cells (lane U) and four transfected cell samples (lanes 1 to 4). These correspond to HDV (lanes 1 and 2) and HDV modified to change the poly(A) signal from AAUAAA to UUUAAA (lanes 3 and 4). For lanes 2 and 4, the plasmids were deleted at bases 1433 to 1434 to disrupt the open reading frame of the delta antigen. (C) Results of dideoxynucleotide sequencing (10, 21) of the four anchored PCR products shown in lanes 1 to 4 of panel B. For this procedure, the products were asymmetrically amplified by using an excess of primer B relative to primer C and then sequenced, by using end-labeled primer D, alongside the ladder from normal HDV cDNA, as shown by the four lanes at the right. In this way it is possible, as indicated at the right, to identify on the four products the poly(A) signal sequence (AAUAAA for lanes 1 and 2) and the mutated sequence, (UUUAAA for lanes 3 and 4). Also indicated is the site of poly(A) addition for lanes 1 to 4 and, for lane 3 and 4, an extra poly(A) site, as discussed in the text.

and the appearance of a new band, consistent with polyadenylation at the SV40 site (lane 3).

The two HDV cDNA constructs used in these transfections each contained the open reading frame for the delta antigen. To determine whether expression of this protein had an effect on RNA processing, we made corresponding cDNA constructs in which the open reading frame was disrupted. This disruption did not bring about any change in RNA processing that was detectable by the PCR assay (Fig. 2B, lanes 2 and 4).

There are definite advantages and disadvantages to the PCR assay as used here. A disadvantage is that the efficiency of PCR amplification decreases with the length of the product (8, 16). Thus, in lanes 3 and 4 of Fig. 2B, the relative amount of polyadenylation at the SV40 site was probably underestimated relative to that of the shorter HDV product. (As revealed subsequently for Fig. 3 and 4, an additional cause for underestimation of RNAs processed at the SV40 polyadenylation site can be the additional ability of such RNAs to undergo self-cleavage.) At the same time, the major advantage of the method was the high sensitivity. As reported previously (11), the method allowed us to routinely detect the mRNA for the delta antigen in samples of HDV-infected cells and liver tissue under conditions in which a Northern analysis usually gave negative results. Another advantage was that the PCR product could be used for direct sequencing (10). Thus, as shown in Fig. 2C, we were able to confirm that for the RNA from the four transfections, the major polyadenylation site was the same as we have reported for infected cells (Fig. 1) (11). Also, for the transfec-

tions with the mutated polyadenylation signal (Fig. 2C, lanes 3 and 4), the sequence of the PCR product confirmed that this mutated site was actually associated with the polyadenylation. The method was so sensitive that it even detected for these two samples an unexpected additional polyadenylation site. By means of the adjacent sequencing ladder, this extra site was shown to correspond to position 901/900, which as indicated in Fig. 1 is the site at which we have shown self-cleavage to occur in vitro (22). While we do not consider such polyadenylation to be relevant to HDV genome replication, it does argue that in the transfected cells, some of the HDV antigenomic RNA did self-cleave at this site and also that the presence or absence of the delta antigen had no effect on such self-cleavage. These results are in fact the first evidence to show that the self-cleavage as first detected and studied in vitro (15, 22) also occurs in vivo.

The studies described above extend the analogy between the polyadenylation of HDV RNA and that of a typical mRNA precursor in terms of the requirements for *cis*-acting elements (3, 18, 19). However, one potentially major difference for HDV, as shown in Fig. 1, is the presence just 3′ of the G+U-rich region of a sequence that directs self-cleavage. The following experiments were undertaken to determine whether polyadenylation at the HDV site was affected by such self-cleavage or was dependent on any of the sequences essential for such self-cleavage.

The strategy for these experiments is summarized in Fig. 3A. We made a series of deletions in the HDV cDNA sequence. After transfection, we used anchored PCR, as for Fig. 2, to determine the relative amounts of polyadenylation at the HDV relative to the downstream SV40 polyadenylation site. As the extent of the deletion of HDV sequences increased, the relative amount of polyadenylation at the HDV site decreased and the amount at the SV40 site increased (Fig. 3B). Figure 3A shows the quantitation of these data as well as the sizes of the deletions in relationship to the domain that is known to be necessary for self-cleavage.

The first construct (Fig. 3A, line 1) contains 150 nucleotides of HDV sequence 3′ of the polyadenylation site and should be able to self-cleave. The results show that when this amount was reduced to 54 nucleotides (line 2), an amount that would remove self-cleavage ability, the efficiency of HDV polyadenylation was affected by no more than 1%. Thus, polyadenylation at the HDV site was not affected by the self-cleavage process, at least in this context. Also, as the amount of HDV sequences was further decreased (lines 3 to 5), there was a progressive decrease in the amount of HDV polyadenylation. The construct in line 4 includes the 34 nucleotides of HDV sequence from the polyadenylation site to the precise location at which self-cleavage otherwise occurs. This construct was thus interesting because it might give an indication of whether an RNA transcript that has self-cleaved could still be polyadenylated at the HDV site. Apparently it could, although there was some effect, and because of the limitations of the PCR assay (as discussed earlier), the efficiency was maybe less than the observed value of 87%. The construct with only 17 nucleotides of HDV sequence had none of the sequences needed for self-cleavage (see Fig. 1) and gave virtually no detectable polyadenylation at the HDV site.

In summary, these experiments offered two major conclusions. First, the potential of the downstream sequences to undergo a self-cleavage was neither an inhibitor nor an enhancer of the polyadenylation at the HDV site. Second, the HDV sequence necessary for in vitro self-cleavage, independent of such cleavage, contained elements (both 5′ and 3′ of the cleavage site) that were necessary for optimal efficiency of polyadenylation at the upstream HDV site. This second conclusion regarding the need for downstream sequences, indicated in Fig. 3A as the deduced domain for polyadenylation, was thus very similar to the consensus picture for the polyadenylation of cellular mRNAs (3, 18, 19).

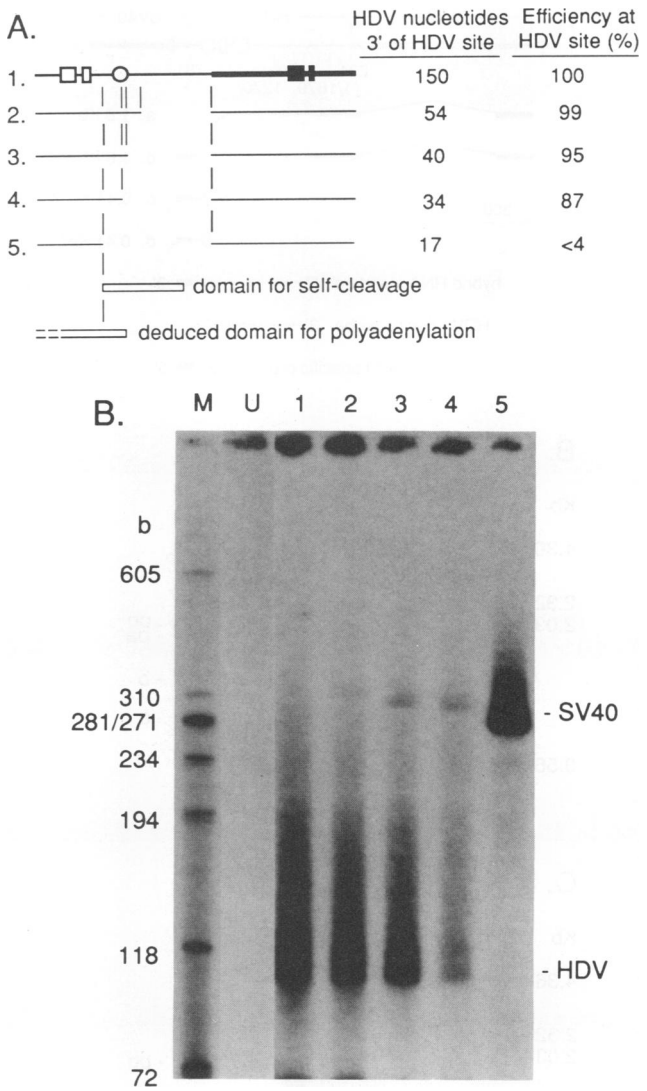


FIG. 3. Role of adjacent *cis*-acting sequences in polyadenylation. HDV cDNAs that were inserted into pSVL and tested by transfection and subsequent anchored PCR, using primer B and end-labeled primer D (see Fig. 2A) to assay the RNA species polyadenylated at either the HDV or SV40 site. Line 1 and the associated symbols are as described in the legend to Fig. 2. A set of five overlapping HDV cDNAs was tested (B). Products corresponding to polyadenylation at the HDV and SV40 sites are indicated at the right. For each sample, densitometry was used to determine the relative amount of the two polyadenylated species, allowing for the fact that in such an assay the two species have a broad range of electrophoretic mobility. The results of this quantitation are summarized at the right of panel A. Also shown in panel A are the relationship of the HDV cDNA constructs to the known essential domain for self-cleavage (15) and the deduced essential domain for polyadenylation. b, bases.

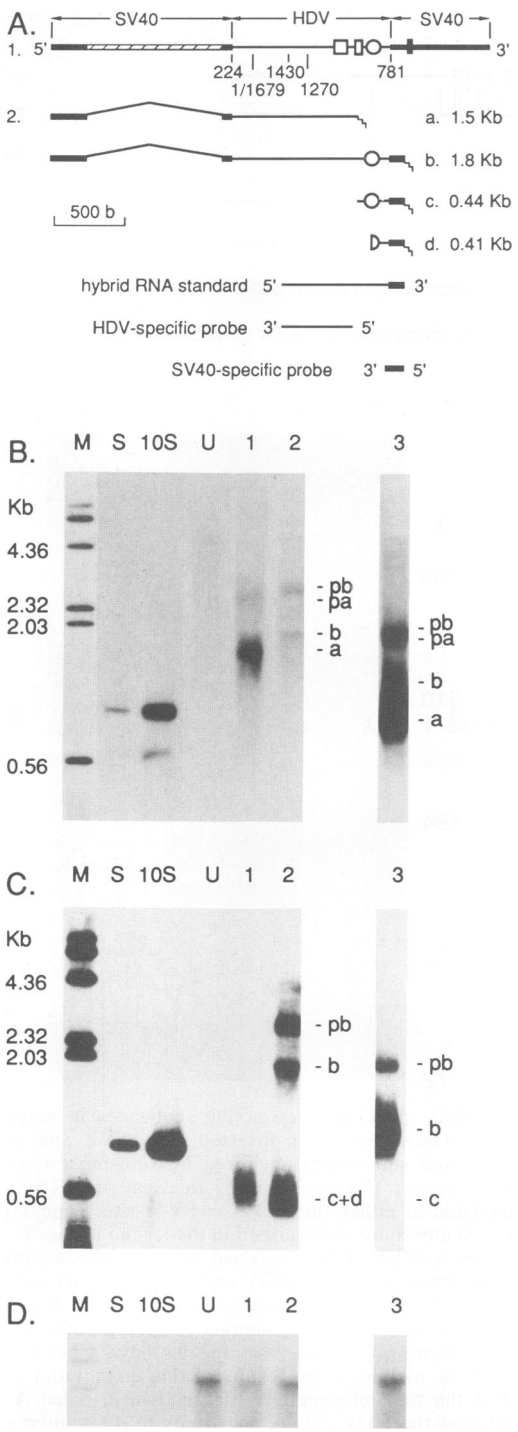


FIG. 4. Elongation of RNA transcripts subsequent to polyadenylation. As indicated in panel A, an HDV cDNA sequence was inserted into the vector pSVL and transfected into COS7 cells. The sequence features and symbols are as for Fig. 2A. After 3 days, the RNA was extracted and then the polyadenylated RNA was isolated (1). Because of the presence of the SV40 intron, the HDV and SV40 polyadenylation sites, and the HDV self-cleavage site, we expected various polyadenylated species to be synthesized. These species are as indicated in line 2 as a to d along with the predicted sizes, assuming the poly(A) to be 0.15 kb (30). Such species could be separated on a Northern analysis using both size and the region-specific RNA probes: HDV (B) and SV40 (C). The cells were actually cotransfected with pRSV(CAT) (generously provided by John Burch of this institute) so that the transfection efficiency could be determined with a probe to the chloramphenicol acetyltransferase sequence, with results as shown in panel D. Four cell samples were tested: lane U, no HDV sequences; lane 1, wild-type HDV; lane 2, HDV modified at the poly(A) signal; lane 3, HDV modified by a small deletion in the domain essential for self-cleavage. The latter construct involved a second and larger deletion, of about 600 bp, from the region *Bgl*III (224) to *Xho*I (1270). The bands are identified at the right. pa and pb refer to the immature forms of species a and b that still contain the SV40 intron. Lane S and the 10-fold-higher amount, lane 10S, refer to the hybrid RNA standard indicated in panel A and described in the text. End-labeled *Hin*dIII fragments of phage lambda DNA provided size markers (lane M), with the sizes shown at the left.

The PCR assay as used above was not designed to detect RNA that had self-cleaved. Thus, to detect both cleaved and uncleaved species and to examine additional questions regarding the processing of the RNA transcripts, we used the assay as described below.

**RNA transcripts extending beyond the HDV polyadenylation site are stabilized by self-cleavage.** It is known that the transcription of mRNA precursors can proceed up to 4,000 nucleotides beyond the polyadenylation site (19), but such amputated or continuing 3′ transcripts are highly unstable (2, 19). However, we reasoned that if in the case of HDV such transcripts could somehow be rendered stable, this stabilization may contribute to the ultimate synthesis of 1,700-nucleotide, full-length, antigenomic RNA. The experiment described below supports the interpretation that such stabilization can be achieved as a consequence of self-cleavage.

The experimental strategy is summarized in Fig. 4A. Cells were transfected with a construct containing HDV cDNA sequences (line 1). After transfection, the polyadenylated RNA was extracted and examined by a Northern analysis. DNA-directed RNA transcription should initiate at a single site, as directed by the SV40 promoter. However, because of the different possibilities of processing, we expected to detect as many as the four different size classes of polyadenylated RNA, as indicated in line 2 by species a to d. Species a arises simply via polyadenylation at the HDV site. Species b represents read-through of this HDV site followed by polyadenylation at the downstream SV40 site. (Actually species a and b are each derived by splicing precursor RNAs, pa and pb, respectively, that contain the SV40 intron. We detected each of these precursors.) Since species b contains the self-cleavage domain, it could self-cleave and produce the fourth species, d. Another possibility is that after polyadenylation occurs at the HDV site and produces species a, the 3′ end of the RNA transcript continues on and ultimately is polyadenylated at the SV40 site, to create species c. An additional possibility is that species c, if it contains the self-cleavage domain, may process itself, to become another source of species d.

To quantitate these four polyadenylated RNAs as studied by Northern analysis, we made use not only of the size differences but also of region-specific probes, as indicated in Fig. 4A. We hybridized first with the HDV probe (Fig. 4B) and then rehybridized with the SV40 probe (Fig. 4C). To normalize the results of these two hybridizations, we included in the Northern analysis a special hybrid RNA standard (Fig. 4A) that was able to hybridize to both probes (Fig. 4B and C, lane S and 10S).

In this study, we examined the RNAs from four different transfections. Results for the construct indicated in line 1 of

TABLE 1. Quantitation of polyadenylated RNA species

| Polyadenylated RNA species[a] | Relative amt of RNA species with given cDNA[b] | | |
|---|---|---|---|
| | Lane 1, wild type | Lane 2, poly(A) signal mutant | Lane 3, self-cleavage mutant |
| a | 75.2 | 11.2 | 76.7 |
| b | <1.5 | 43.0 | 20.8 |
| c+d | 23.3 | 45.7 | <2.5 |

[a] These four RNA species are defined in the legend to Fig. 4A. Also, species a includes the unspliced form, species pa, and likewise species b includes pb.
[b] The results are for the three transfections shown in lanes 1 to 3 of Fig. 4. The individual RNA species were quantitated from the Northern analyses of Fig. 4B and C, using an AMBIS radioanalytic image analyzer. The results for each transfection are normalized to 100% total. The quantitation refers not to the mass distribution but to the relative number of molecules, because the probes used were small and region specific.

Fig. 4A are shown in lane 1 of Fig. 4B and C. Similarly, lane 2 shows the results for a construct mutated in the HDV polyadenylation signal. Lane 3 shows the results for a construct deleted so as to remove the self-cleavage ability. Lane U was not transfected with an HDV construct. All four transfections included as a control for transfection efficiency plasmid pRSV(CAT). Expression from this plasmid was monitored by hybridization with an appropriate specific probe (Fig. 4D).

The radioactivity in each of the labeled RNA bands was quantitated directly from the filter. These data are presented in Table 1; species c and d have been grouped together, since in the absence of additional data, the Northern analysis is not sufficient to resolve them. The following observations can be made.

For the wild-type HDV sequence as template (i.e., lane 1 in Fig. 4B and C and Table 1), species b, polyadenylation at the SV40 site, was apparently less than 1.5% relative to polyadenylation at the HDV site. The situation was, however, more complicated because species c and/or d existed and accounted for more than 23% of the total transcripts. These are both polyadenylated at the SV40 site. How did they arise? One possibility was simply read-through of the HDV polyadenylation site, followed by self-cleavage. Another possibility was that HDV polyadenylation occurred but that the 3′ transcripts continued on, with or without a subsequent self-cleavage.

To help distinguish between these possibilities, we considered the consequences of transfecting with a construct mutated at the HDV polyadenylation signal, so as to increase those species that arise by read-through of the HDV site (i.e., lane 2 in Fig. 4B and C and Table 1). As expected, this change caused a dramatic increase in species b relative to species a. Also, and of most relevance, species c+d was also increased, to about 46%. Our interpretation is that this latter increase was due to species d (and not species c) that arose via self-cleavage of some of the read-through species b. According to this interpretation, maybe only about half of the total read-through species actually self-cleaved (leaving a residual amount of 43%). Thus, if self-cleavage of read-through transcripts is of this efficiency, we can now begin to sort out the possible interpretations for transfection with the wild-type construct. Specifically, the majority of the 23% of transcripts pooled together as c+d could not have arisen via self-cleavage of the read-through species b, of which less than 2% was left. This leaves us with the other possibility, that is, that we can have polyadenylation at the HDV site, to generate species a, followed by continuation of the 3′ transcript so as to generate a second species polyadenylated at the SV40 site. This second species might be d or c, depending on whether or not there is a self-cleavage event. From the work of others on the processing of mRNA precursors, there is ample precedent for the concept of polyadenylation being accompanied by a continuing 3′ transcript (19). Such continuing transcripts are relatively unstable, although the reason for such instability is not yet clear. We therefore speculated that the self-cleavage reaction might somehow act to confer stability on the continuing transcript. As explained below, we used yet another HDV construct to test and confirm this possibility.

To clarify the role of self-cleavage in RNA processing, we transfected cells with another cDNA mutant, one with a small deletion in the region essential for self-cleavage (lane 3 in Fig. 4B and C and Table 1). (This construct had a second deletion, much larger but much 5′ of the polyadenylation site. This caused a reduction in the sizes of species a and b and the corresponding unspliced versions, pa and pb, as indicated in Fig. 4.) Since self-cleavage was no longer possible, the total for c+d was really only c. Moreover, such species could only have arisen via processing of what we have referred to above as a continuing 3′ transcript. Since the relative amount of species c was now less than 2.5%, we made two conclusions. First, such a low value was consistent with what others have pointed out as the instability of continuing 3′ transcripts (2, 19). Second, such a low amount was not true for lane 1, in which self-cleavage could occur. In other words, the data support our interpretation that self-cleavage can lead to significant stabilization of the continuing 3′ transcript.

Along the lines of this interpretation, we would expect that species c+d as seen in lane 1 for the wild-type construct would actually, because of self-cleavage, be predominantly of species d. To test this possibility, we used anchored PCR with an HDV-specific primer either just 5′ or just 3′ of the self-cleavage site. The results indicated that self-cleavage had occurred (12). However, an unavoidable limitation of such an experiment was that the self-cleavage could equally well have occurred during the assay.

If our interpretation that a single transcription initiation can lead to both an mRNA processed at the HDV site and a stabilized 3′ continuing transcript is correct, then we can deduce from Table 1 that the efficiency of such stabilization is not 23.3% but about $23.3 \times 100/75.2 = 31\%$. (We speculate that if the downstream RNA had been able to fold into the rodlike structure, this may have increased the stabilization even further.)

In summary, the main conclusion from this experiment was that stabilization of the 3′ continuing transcript was achieved because of the self-cleavage. From this conclusion, we speculate that in HDV genome replication, both antigenomic mRNA and one or even more copies of full-length antigenomic RNA can be generated following a single transcription initiation. However, in this or any model of HDV genome replication, there remains, as considered below, yet another question: How are the full-length transcripts able to escape polyadenylation?

**Suppression of a functional HDV polyadenylation site.** We have shown that transcripts of antigenomic RNA can be readily processed by polyadenylation. However, any acceptable model of HDV genome replication must also explain how there can be antigenomic RNA species that are not so processed and that go on to become the full length and act as templates for the synthesis of new genomic RNA. We

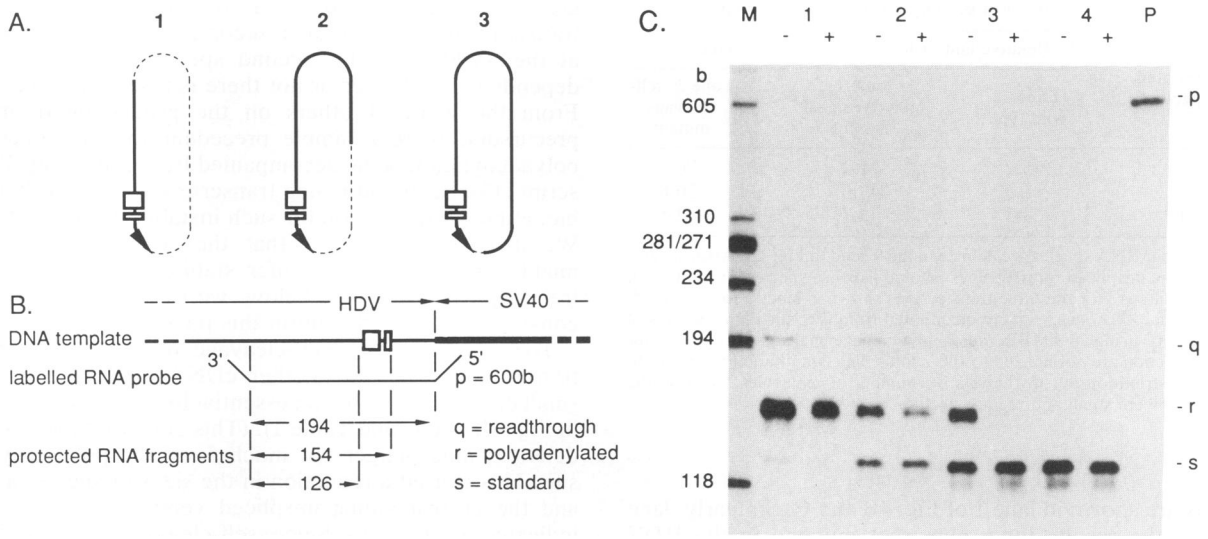


FIG. 5. Ability of RNA folding, with and without the delta antigen, to suppress polyadenylation. (A) Representation of three different HDV cDNAs that were inserted into pSVL. The inserts are represented in terms of their relationship to the rodlike structure. The solid lines represent the inserts, with the arrowhead indicating the 3′ end of the antigenomic sequence. The broken lines represent the remainder of the rodlike structure. The other symbols are as in Fig. 2A. The precise inserts were as follows: 1, *Xho*I (1270) to 894; 2, *Nhe*I (430) to 894; and 3, *Xba*I (781) to 894. The latter two contained the open reading frame of the delta antigen but, to occlude such expression, were deliberately mutated by a deletion at 1433 to 1434. (B) Enlargement of the region of the DNA that contains the HDV polyadenylation signals. COS7 cells were transfected not only with the constructs mentioned above but also with a second construct that expresses antigenomic HDV sequences from *Bgl*I (224) to *Sal*I (962). This construct provided two functions. First, it provided an internal standard for the RNA assays; the associated mRNA involved did not use the HDV site but rather used an SV40 site. Second, it contained sufficient HDV sequences to allow the synthesis of the delta antigen; an alternative version of the construct contained a deletion of two bases (at positions 1569 and 1570) so as to disrupt the delta antigen. RNA was extracted from the transfected cells, and then an RNase protection assay, using the labeled RNA probe indicated in panel B as p, was employed to detect HDV-specific RNA transcripts. The expected protected RNA fragments are indicated as species q for read-through, r for polyadenylation at the HDV site, and s for RNAs transcribed from the internal standard mentioned above. These fragments were separated by gel electrophoresis and then subjected to autoradiography (C) and direct quantitation. Transfections with constructs 1 to 3 of panel A are indicated as 1 to 3, respectively. All transfections contained the standard construct, as described above, that could or could not express the delta antigen, as indicated by + or −. Lane P represents a 1% aliquot of the intact probe (p). End-labeled *Hae*III fragments of phage ϕX DNA provided size markers (lane M), with the sizes (bases [b]) shown at the left.

propose that one possible solution might be to use a single initiation event and allow polyadenylation to occur, but then the 3′ transcript continues and is stabilized by self-cleavage, as supported by the data presented above, and then by some mechanism, when the transcript passes through the polyadenylation signal region, for the second time, suppress polyadenylation. As described below, we have data to support this possibility. The mechanism of suppression seems to be provided by the ability of the antigenomic RNA to fold into the rodlike structure, which in turn can act as a binding site for the delta antigen.

The strategy can be explained in terms of Fig. 5A. Three HDV cDNA inserts, indicated as 1 to 3, were tested for their ability to lead to polyadenylation at the HDV site. These inserts all contained the HDV polyadenylation region, and all lacked a self-cleavage domain. Two of the inserts were much longer than those used in the earlier experiments, such as for Fig. 2, and long enough so that at least some of the antigenomic RNA transcribed was expected to fold into the previously described rodlike structure, characteristic of HDV genomic and antigenomic RNA (14, 28). Species 3, the largest, was such that even the region needed for polyadenylation was able to be included in the rodlike structure. In contrast to these, species 1 contained no rodlike structure. It was chosen both because of this and because, like the known precursor to HDV mRNA, the initiation of the HDV antigenomic RNA sequences was close to the top of the rod (11).

These three HDV cDNAs were thus inserted into the expression vector pSVL and, as in the previous experiments, used to transfect COS7 cells. Equal amounts of one of two additional constructs were used in each transfection. One of these additional constructs was able to provide the delta antigen (the 195-amino-acid species [14, 17]), and the other, because of a two-base deletion in the open reading frame, was not. These constructs used an SV40 rather than an HDV polyadenylation site.

To sort out and quantitate the various HDV RNA species in the transfected cells, we used a 600-base labeled RNA probe, designated p in Fig. 5B, and an RNase protection assay. Our expectation was that unprocessed transcripts would yield a 194-base protected fragment (q), while an RNA polyadenylated at the HDV site would yield a 154-base fragment (r). Processed RNA from the cotransfected plasmid would yield a 126-base fragment (s) and act as an internal standard. These labeled RNAs were resolved on a polyacrylamide gel (Fig. 5C). Lanes 1 to 3 refer to cotransfections using the corresponding three constructs, while lane 4 refers solely to the internal control. The ability of the internal control to express the delta antigen is indicated above each lane.

Consider first those transfections in which the delta antigen was not expressed. As seen in lane 4, the internal control alone yielded, as expected, only species s. The cotransfections involving constructs 1 to 3 (lanes 1 to 3) yielded, as

expected, species r, corresponding to polyadenylation at the HDV site, and relatively little of the read-through transcript (species q). One striking difference between lanes 1 to 3 was that lane 1 had much more of the species r. One possible interpretation of this finding is that the absence of potential rodlike structure correlates with an increased efficiency of polyadenylation. Another possibility could be that the rod structure interferes with RNA transcription efficiency. Further experiments are needed to understand this phenomenon.

When we compared these results with those of the transfection in which the delta antigen was present, a major effect was seen. The amount of species r, corresponding to polyadenylation at the HDV site, was affected. From direct quantitation of lanes 1 to 3, as normalized to the standard and relative to the absence of antigen as 100%, the amounts of polyadenylation at the HDV site when delta antigen was present corresponded to 107, 72, and 14%, respectively. Thus, if the nascent RNA can fold into the rodlike structure, then the presence of the delta antigen can cause a major suppression of polyadenylation.

We already know from other studies that the delta antigen has HDV-specific RNA binding ability, with a specificity for binding to the rodlike structure (4). Thus, our interpretation of the present data was that the delta antigen, as a consequence of such binding, was able to act in *trans* and suppress the functionality of the polyadenylation site on antigenomic HDV RNA.

It is important to note that consistent with the above interpretation regarding the absolute amounts of polyadenylation for the internal standard, which actually had a very small amount of rodlike structure (about half of that for construct 2), there was in each of constructs 1 to 4 no significant effect correlating with the expression of functional delta antigen.

One final observation that deserves comment was that the read-through transcript, species q, was virtually unchanged for lanes 1 to 3. That is, it was not correspondingly increased when the polyadenylation was dramatically decreased. A possible explanation, based on the results presented above, is that the read-through transcripts in these experiments were unstable because they lacked the self-cleavage domain.

## DISCUSSION

In this report, we have addressed three major questions concerning the processing of HDV RNA transcripts. We have used DNA-directed transcription of HDV antigenomic RNA, via cDNA constructs, to examine these questions inside cells, but in the absence of complications arising from the full cycle of HDV genome replication. We have shown that (i) the polyadenylation of antigenomic HDV RNA can be an efficient process and is dependent on certain sequence features and motifs, very much like the consensus picture for cellular mRNA precursors, and yet independent of the self-cleavage reaction, previously characterized by in vitro studies (15, 23) and known to occur just 34 nucleotides downstream of the actual polyadenylation site; (ii) a single nascent antigenomic RNA can give rise to not only a 5′ species polyadenylated at the HDV site but also a continuing 3′ transcript that can be rendered stable by the process of self-cleavage; and (iii) the folding of a nascent transcript into the rodlike structure can give major suppression of polyadenylation at the HDV site if the delta antigen is also present. This in vivo effect of the delta antigen in the presence of the rodlike structure is consistent with our recent in vitro studies which show that the delta antigen will specifically bind to HDV RNAs that can fold into the rodlike structure (4).

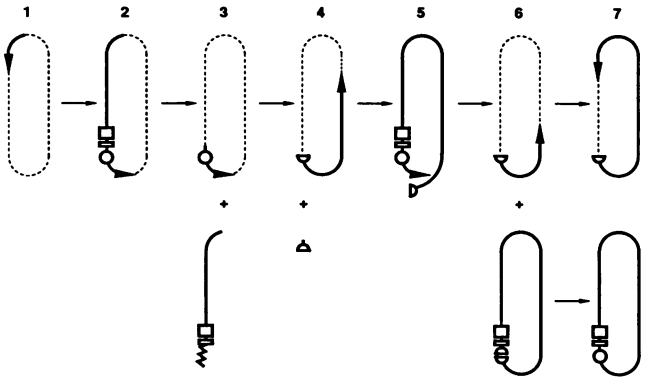


FIG. 6. Model for the regulation of antigenomic synthesis during HDV genome replication. The symbols are as used in the previous figures, and the seven steps shown are described in the text.

We have incorporated these results along with other findings to propose a model of HDV genome replication that is more detailed than the rolling-circle model previously presented (26). The model shown in Fig. 6 focuses on the transcription of genomic RNA into antigenomic RNA. In step 1, transcription is considered to start at the same location as that which has been experimentally shown to be the 5′ end of the mRNA, that is, only five bases from the top of the rod structure (11). In step 2, when the nascent transcript extends beyond the signals for polyadenylation, it is processed, as indicated in step 3, to generate both the mRNA for the delta antigen and a 3′ continuing transcript. On the basis of the results of this study, we propose that the 3′ end of the nascent transcript can continue and that because of self-cleavage, as indicated in step 4, this continuing transcript is stabilized. [We do not yet know how self-cleavage causes such stabilization. The self-cleavage would leave the HDV RNA with a 5′-OH terminus (15), which is unlike the 5′-phosphate of the typical amputated fragment generated by the poly(A) processing and known for its high instability (2, 19).] This stabilized transcript can now continue, as indicated in step 5, until it reaches, for the second time, the signals for polyadenylation. This time, however, because of the ability of the nascent RNA to fold into the rodlike structure and the ability of the delta antigen, if present in sufficient amounts to bind to this rodlike structure (4), the polyadenylation process is suppressed. Thus, the transcript continues further and reaches the self-cleavage domain for the second time. The RNA now self-cleaves, as indicated in step 6, releasing a fragment of exactly unit length antigenomic RNA; because of the folding of this fragment into the rodlike structure, its two ends are brought into precise apposition, an alignment which, as we have previously shown, is a potent facilitator of self-ligation (23), as indicated in step 7. Meanwhile, the 3′ end of the nascent RNA may continue, as indicated in step 7, through one or more additional rolling-circle events.

We envision that the unit-length circular antigenomic RNA produced in step 7 can go on to be the RNA template for the synthesis of genomic RNA. The process is expected to be similar, although simpler, because there is no polyadenylation site that needs to be regulated.

Finally, it must be pointed out that the synthesis of a special mRNA for the delta antigen, and the role of the delta

antigen in this process, is unlikely to be a purely selfish process. That is, the delta antigen is not present just to regulate its own synthesis. There is no doubt that the delta antigen is also needed for the packaging of progeny genomic RNA into new virions. However, the situation is still not entirely clear; we know that the delta antigen, independent of such packaging, is also needed for genome replication (13). Moreover, even if the polyadenylation signal is inactivated, so that regulation might not be expected to be needed, there is still at least one more essential role, as yet not clarified, for the delta antigen in genome replication (11).

## ACKNOWLEDGMENTS

J.T. was supported by grants CA-06927, RR-05539, and AI-26522 from the National Institutes of Health, by grant MV-7M from the American Cancer Society, and by an appropriation from the Commonwealth of Pennsylvania.

Tony Yeung provided the chemical synthesis of the oligonucleotides used in the study. Karen Trush was responsible for the computer graphics. Rich Katz, Bob Perry, Wang-Shick Ryu, and Christoph Seeger provided valuable critical readings of the manuscript.

## REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. Wiley Interscience, New York.
2. **Bernstein, P., and J. Ross.** 1989. Poly(A), poly(A) binding protein and the regulation of mRNA stability. Trends Biochem. Sci. **14**:373–377.
3. **Birnstiel, M. L., and F. J. Schaufele.** 1989. Structure and function of minor snRNPs, p. 155–182. *In* M. L. Birnstiel (ed.), Small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin.
4. **Chao, M., S.-Y. Hsieh, and J. Taylor.** 1991. The antigen of hepatitis delta virus: examination of in vitro RNA-binding specificity. J. Virol. **65**:4057–4062.
5. **Chen, P.-J., G. Kalpana, J. Goldberg, W. Mason, B. Werner, J. Gerin, and J. Taylor.** 1986. The structure and replication of the genome of the hepatitis delta virus. Proc. Natl. Acad. Sci. USA **83**:8774–8778.
6. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
7. **Cullen, B. R.** 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. Methods Enzymol. **152**:684–704.
8. **Frohman, M. A., M. K. Dush, and G. R. Martin.** 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85**:8998–9002.
9. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
10. **Gyllensten, U.** 1989. Direct sequencing of *in vitro* amplified DNA, p. 45–60. *In* H. A. Erlich (ed.), PCR technology. Stockton, New York.
11. **Hsieh, S.-Y., M. Chao, L. Coates, and J. Taylor.** 1990. Hepatitis delta virus genome replication: a polyadenylated mRNA for the delta antigen. J. Virol. **64**:3192–3198.
12. **Hsieh, S.-Y., and J. Taylor.** Unpublished data.
13. **Kuo, M. Y.-P., M. Chao, and J. Taylor.** 1989. Initiation of replication of the human hepatitis delta virus genome from cloned DNA: role of delta antigen. J. Virol. **63**:1945–1950.
14. **Kuo, M. Y.-P., J. Goldberg, L. Coates, W. Mason, J. Gerin, and J. Taylor.** 1988. Molecular cloning of hepatitis delta virus RNA from an infected woodchuck liver: sequence, structure and applications. J. Virol. **62**:1855–1861.
15. **Kuo, M. Y.-P., L. Sharmeen, G. Dinter-Gottlieb, and J. Taylor.** 1988. Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. J. Virol. **62**:4439–4444.
16. **Loh, E. Y., J. F. Elliott, S. Cwirla, L. L. Lanier, and M. M. Davis.** 1989. Polymerase chain reaction with single-sided specificity: analysis of T-cell receptor δ chain. Science **243**:217–220.
17. **Luo, G., M. Chao, S.-Y. Hsieh, C. Sureau, K. Nishikura, and J. Taylor.** 1990. A specific base transition occurs on replicating hepatitis delta virus RNA. J. Virol. **64**:1021–1027.
18. **Proudfoot, N.** 1991. Poly(A) signals. Cell **64**:671–674.
19. **Proudfoot, N. J.** 1989. How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. **14**: 105–110.
20. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Sharmeen, L., M. Y.-P. Kuo, G. Dinter-Gottlieb, and J. Taylor.** 1988. The antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. J. Virol. **62**:2674–2679.
23. **Sharmeen, L., M. Y.-P. Kuo, and J. Taylor.** 1989. Self-ligating RNA sequences on the antigenome of human hepatitis delta virus. J. Virol. **63**:1428–1430.
24. **Taylor, J.** 1990. Hepatitis delta virus *cis* and *trans* functions needed for replication. Cell **61**:371–373.
25. **Taylor, J.** 1990. Structure and replication of hepatitis delta virus. Semin. Virol. **1**:135–141.
26. **Taylor, J., M. Kuo, P.-J. Chen, G. Kalpana, J. Goldberg, C. Aldrich, L. Coates, W. Mason, and J. Summers.** 1987. Replication of hepatitis delta virus, p. 541–548. *In* W. Robinson, K. Koike, and H. Will (ed.), Hepadna viruses. Alan R. Liss, Inc., New York.
27. **Taylor, J. M.** 1991. Human hepatitis delta virus. Curr. Top. Microbiol. Immunol. **168**:141–166.
28. **Wang, K.-S., Q.-L. Choo, A. J. Weiner, J.-H. Ou, R. C. Najarian, R. M. Thayer, G. T. Mullenbach, K. J. Denniston, J. L. Gerin, and M. Houghton.** 1986. Structure, sequence and expression of the hepatitis delta viral genome. Nature (London) **323**:508–513.
29. **Weiner, A. J., Q.-L. Choo, K.-S. Wang, S. Govindarajan, A. G. Redeker, J. L. Gerin, and M. Houghton.** 1988. A single antigenomic open reading frame of the hepatitis delta virus encodes the epitope(s) of both hepatitis delta antigen polypeptides $p24^{\delta}$ and $p27^{\delta}$. J. Virol. **62**:594–599.
30. **Wickens, M.** 1990. How the messenger got its tail: addition of poly(A) in the nucleus. Trends Biochem. Sci. **15**:277–281.
31. **Zack, J. A., S. L. Arrigo, S. R. Weitsman, A. S. Go, A. Haislip, and I. S. Y. Chen.** 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. Cell **61**:213–222.