***In vivo* RNA-directed transcription, with template switching, by a mammalian RNA polymerase**

Jinhong Chang and John Taylor¹

Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111-2497, USA

¹Corresponding author
e-mail: jm_taylor@fccc.edu

These studies support the interpretation that a host polymerase, most likely RNA polymerase II, can not only carry out transcription that is RNA directed, but also achieve template switching on a discontinuous RNA template, and even perform non-templated nucleotide incorporation. As part of an *in vivo* analysis of the initiation of replication of the RNA genome of human hepatitis delta virus (HDV), a series of linear RNAs containing HDV sequences was tested in order to explain the ability of this host polymerase to initiate RNA-directed RNA synthesis *in vivo* and produce replicating circular HDV species. The data support the hypothesis that the input linear template RNAs were not converted to circles before transcription but rather that in the process of transcription, the polymerase was able to make an intra-molecular template switch. Furthermore, in certain cases this switch produced small deletions of template sequences, and in some cases even insertion of non-templated sequences. Thus, in an *in vivo* situation, polymerase II has several important capabilities in addition to what is considered typical DNA-directed transcription.

Keywords: aberrant RNA/hepatitis delta virus/RNA-directed transcription/RNA polymerase II/template switching

Introduction

RNA-directed RNA transcription is carried out by most RNA viruses using polymerases encoded in their own genomes. However, there are now several examples in plants and even animals of biologically relevant RNA-directed RNA transcription carried out by host-encoded polymerases. For instance, some of the plant agents known as viroids redirect the host RNA polymerase II (pol II) to replicate their genome (Schindler and Muhlbach, 1992) and others redirect a chloroplast polymerase (Navarro *et al.*, 2000). In addition, apparently all plants and some animals encode an RNA-directed RNA polymerase (Schiebel *et al.*, 1998), and there is solid genetic evidence that this gene is needed in some forms of post-transcriptional gene silencing (Cogoni and Macino, 1999). In mammals no such polymerase has yet been revealed, but there is just one example of RNA-directed transcription by a host enzyme, that for human hepatitis delta virus (HDV).

The RNA genome of HDV is very small (1679 nt), circular in conformation, essentially single-stranded, and

yet predicted to be able to fold into an unbranched rod-like structure with ~70% of the bases paired (Chen *et al.*, 1986; Kos *et al.*, 1986; Wang *et al.*, 1986). In these respects it is similar to that of the plant agents known as viroids (Taylor, 1999). In addition, both HDV and the viroid RNAs are able to redirect a host polymerase to carry out the RNA-directed RNA transcription needed to achieve the replication of their genomes. HDV replication expresses a small protein (195 amino acids) known as the delta antigen that is somehow essential for this replication (Kuo *et al.*, 1989). In contrast, the viroid RNAs, which are several fold smaller than HDV, encode no known proteins (Diener, 1991).

When multimeric RNA transcripts of genomic or antigenomic RNA are synthesized *in vivo*, they can be processed to unit length circular RNAs (Lazinski and Taylor, 1994). This processing involves site-specific cleavage by natural ribozymes. Both the genome and antigenome contain a ribozyme (Kuo *et al.*, 1988b; Sharmeen *et al.*, 1988). Cleavage produces unit length linear species. Then, what is considered to be a host-mediated ligation converts the linear species to circles (Reid and Lazinski, 2000).

The template requirements for the initiation of HDV and viroid RNA transcription are not yet clear. For HDV, there is evidence that in the synthesis of new antigenomic RNA initiation occurs at a specific site located very close to one end of the predicted rod-like folding of the genomic RNA template (Gudima *et al.*, 2000). Studies with mutated HDV sequences support but do not prove the requirement of the rod-like structure for this initiation (Macnaughton *et al.*, 1993; Beard *et al.*, 1996; Wu *et al.*, 1997; Gudima *et al.*, 2000). During viroid replication, linear RNAs with specific 5' ends have been found and there is evidence that these might be indicative of initiation sites (Navarro and Flores, 2000).

For both HDV and the viroids, the very observation of circular RNA species has led to the adoption of a variety of rolling circle models of replication (Branch and Robertson, 1984; Diener, 1989; Taylor, 1990; Symons, 1992; Daros *et al.*, 1994). The first question addressed in the present study was whether initiation requires the HDV RNA template to be in a circular conformation. The question was provoked by previous studies that showed that following transfection of cells with linear forms of *in vitro* transcribed HDV RNAs that were at least of unit length (1679 nt), initiation of genome replication occurred, leading to the accumulation of unit length RNA circles (Glenn *et al.*, 1990; Bichko *et al.*, 1994; Dingle *et al.*, 1998; Modahl and Lai, 1998). A similar study was reported with linear forms of a viroid RNA. It was concluded that if the 3' and 5' ends of the RNA represented a small inverted repeat, then that RNA was more likely to be infectious (Rakowski and Symons, 1994). In both the

Fig. 1. Representation of linear antigenomic RNAs of HDV that were tested for their ability to initiate transcription and accumulation of processed genomic RNA species. (A) Shows a circular form of unit length 1679 nt antigenomic HDV RNA that has been opened at a unique site p, and then extra sequences x and y added at the new 5' and 3' ends. These extra sequences may be of different lengths and sequence composition (HDV and/or non-HDV). In this example, indicated as species 1 $\times$, p = 434, x = 32 and y = 6. (For more details, see representation at top of Figure 4A.) The location of the antigenomic ribozyme, Rz, is indicated (shaded rectangle), as is the location of the open reading frame for the small delta protein. The position designated 1679/1 refers to the numbering for the genomic HDV RNA from a published HDV sequence (Kuo *et al.*, 1988a). (B) A specific case of a linear RNA, designated 1.2 $\times$, that is not only greater than unit length but has a direct repeat of 181 nt of HDV sequence (from the region nt 962–781), and as indicated contains a second copy of the ribozyme, Rz. The open reading frame for the small delta protein, which is essential for genome replication is present on but cannot be translated from such linear RNAs such as 1 $\times$ and 1.2 $\times$. However, as indicated in (C), during replication a small capped and polyadenylated RNA is produced for such translation. To initiate the transfection experiments described here, such a mRNA was transcribed *in vitro*, and co-transfected along with the larger linear RNA species.

HDV and viroid studies, it was left unsolved as to how such linear RNA templates were transcribed. For example, was the linear RNA somehow converted to a circle before being transcribed? Alternatively, did the transcription somehow pass through the discontinuity site on the linear template to produce RNA species that might go on to be processed into circles? Our studies addressed these questions in the case of HDV replication as initiated by linear RNA species. First, we found that such replication did not require a conversion or 'healing' of the linear antigenomic template RNA to a unit length circular form before transcription. This led us to test the alternative hypothesis that during the process of copying linear antigenomic templates, nascent genomic RNAs were transcribed across the discontinuity on the predicted rod-like folding. This led to the production of RNA multimers that were subsequently processed to form healed unit length circular genomic RNAs with the wild-type HDV sequence. We found that the length and sequence composition at the 5' and 3' ends of the linear antigenomic RNA template, adjacent to the predicted discontinuity, determined whether successful replication could occur. Moreover, in some situations, although there was genome replication and circle formation, small deletions and even insertions (of non-templated sequences) occurred at the site of the original discontinuity. Separate studies have provided evidence that the transcription of HDV antigenomic RNAs involves the re-direction of the host pol II

(Modahl *et al.*, 2000; Moraleda and Taylor, 2001). Thus, the present studies further show that pol II is capable of transcription across discontinuities in the RNA template, sometimes with precision, but sometimes with deletions and even non-templated insertions.

Results

Design of linear RNA templates for RNA transcription *in vivo*

During HDV replication, most of the genomic HDV RNA and its exact complement, the antigenome, are in the form of circular RNAs, 1679 nt in length, that can fold into an unbranched rod-like structure. For subsequent *in vivo* transcription experiments, we designed antigenomic RNA templates that were linear rather than circular. As represented in Figure 1A, these linear RNAs were equivalent to unit length 1679 nt circular species that had been opened up at a unique site, indicated as p, at which extra sequences were added at the new 5' and 3' ends, with lengths indicated as x and y. These extra sequences could be of HDV and/or not of HDV origin. The antigenomic RNA contains a natural ribozyme domain, which can create a cleavage at position 901/900 (Sharmeen *et al.*, 1988), as indicated by Rz in Figure 1A. The nucleotide positions on HDV use the notation of a published sequence (Kuo *et al.*, 1988a).

The species indicated as 1 $\times$ in Figure 1A is an example of one of the linear RNA templates tested the size of which was very close to unit length. For some greater than unit length linear RNAs tested, the additional sequences are no more than a direct terminal repeat of HDV sequences. The length of this repeat was varied from several nucleotides up to as much as 181 nt. A representation of the latter template is shown in Figure 1B as 1.2 $\times$. Note that this terminally redundant species is so long that it has two copies of the HDV ribozyme.

As indicated in Figure 1A and B, these antigenomic RNAs of at least unit length contain the open reading frame for the small delta protein, a 195 amino acid species that is essential for genome replication (Kuo *et al.*, 1989). However, such RNAs are not translated. It has been found previously that once HDV replication is under way, not only are the genomic and antigenomic RNAs transcribed, but also a mRNA for the translation of the small delta protein (Hsieh *et al.*, 1990). As indicated in Figure 1C, this mRNA has a 5' cap (Gudima *et al.*, 2000), and a 3' poly(A) tail (Hsieh *et al.*, 1990). It has also been observed that the small delta protein, or the mRNA for this protein, must be present at the time of RNA transfection (Glenn *et al.*, 1990; Modahl and Lai, 1998). Therefore, in the experiments described here (unless specifically stated otherwise), a mRNA species just as in Figure 1C was transcribed *in vitro*, and co-transfected along with the linear species, as indicated in Figure 1A or B. All linear templates and mRNA were produced by *in vitro* transcription using phage T7 RNA polymerase and DNA templates that were either recombinant plasmids or PCR products with the T7 promoter sequence appropriately placed.

Fig. 2. Northern assays demonstrated the ability of linear antigenomic RNAs to initiate genome replication, producing unit length circular genomic RNAs. Replication was initiated either by transfection of cultured HuH7 cells, at day 4, or of mouse hepatocytes, at day 5. The RNA species used for initiation were 1× and 1.2×, as described in Figure 1A and B, respectively. In each case a mRNA, as indicated in Figure 1C, was co-transfected as an initial source of the essential small delta protein. Following such co-transfection total RNA was extracted and subjected to northern blot analysis to detect HDV genomic RNA. As indicated at the right side, the electrophoresis conditions were such as to be able to separate circular and linear forms of 1679 nt unit length RNA.

In vivo transcription of linear RNA templates can lead to accumulation of circular RNAs

A series of such linear HDV RNA templates were transfected into cultured cells and thus tested for their ability to initiate genome replication. Four days after transfection the total RNA was extracted and tested by northern analyses for the accumulation of unit length genomic RNAs with circular or linear conformations.

Consider just the two examples of transfection with the species 1× and 1.2×, as described above and in Figure 1A and B, respectively. As shown in lanes 1 and 2 of Figure 2, following the transfection of cultured HuH7 cells with 1× and 1.2× RNAs, we detected accumulation of unit length genomic HDV RNA. The amount of accumulation for 1× was 19% relative to 1.2×. Furthermore, for both, the accumulated RNA was 80% circular.

As an alternative to this transfection of cultured cells, we also determined whether similar results would be obtained following hydrodynamic transfection of mouse hepatocytes *in vivo* (Chang *et al.*, 2001). As shown in lanes 3 and 4 of Figure 2, this was the case. The accumulation for 1× was 54% relative to 1.2×, and in both cases, 95% of the RNA was circular.

For each of the transfections using 1× RNA we then applied RT-PCR, cloning and nucleotide sequencing to examine the site p on the unit length genomic RNA species that had accumulated as a consequence of RNA-directed transcription. We found that the nucleotide sequence was exactly as for the unmodified HDV genome (data not shown); we suggest that the discontinuity on the linear template was perfectly 'healed', with precise removal of all the extraneous nucleotides. Such perfect healing was first observed some time ago following a somewhat similar transfection of cultured cells with a greater than unit length linear genomic RNA (Glenn *et al.*, 1990). That study did not continue to determine how the healing was achieved; this question became the aim of our subsequent experiments.

The 1.2× RNA, because it has a large terminal redundancy of HDV sequences, has two copies of the antigenomic ribozyme domain (Figure 1B). Thus, it is

Fig. 3. RT-PCR assays support the interpretation that linear RNAs were not converted or 'healed', to circles before RNA-directed transcription. HuH7 cells were transfected with the antigenomic linear species 1× (Figure 1A) either (i) alone (lane 2), (ii) co-transfected with mRNA to produce the small delta protein (lane 3) or (iii) co-transfected with mRNA to produce the large delta protein, which inhibits HDV genome replication (lane 4). At day 4 after transfection, the total RNA was extracted. (A) A representation of 1×, the linear antigenomic RNA, along with the three primers, L–N, used in the RT-PCR analyses. (B) and (C) show the products obtained after reverse transcription with primer L, followed by PCR using the primers L–M and L–N, respectively. Lane 5 was a positive control (greater than unit length HDV cRNA added to RNA from untransfected HuH7 cells) and lane 6, a negative control (untransfected HuH7 RNA). Relative to the DNA markers, lane 1, the band of 170 bp was as expected for 'healed' RNA.

plausible that following transfection some of this template RNA might undergo ribozyme cleavage followed by ligation to produce a circular RNA. In contrast, the 1× RNA has only one ribozyme domain (Figure 1A), and so if it were to be circularized following transfection and before transcription, another mechanism would have to be invoked. However, before considering such a possibility, we tested whether or not such circularization (i.e. what could be called 'prior healing') did occur.

Linear RNA templates are not circularized before transcription

To test for circularization we used the 1× template RNA and three forms of transfection into HuH7 cells. In the first, we used 1× alone, which we knew would not allow subsequent replication. In the second, we used 1× together with mRNA for the small delta protein, which does support replication. And, in the third we used 1× together with a different mRNA, that for a larger form of the delta protein, one that is known to have some properties in common with the small form, but is unable to support replication (Kuo *et al.*, 1989; Chao *et al.*, 1990). In each case, we extracted total RNA 4 days after transfection. As represented in Figure 3A, we used a RT-PCR assay to detect a predicted 170 bp product, L–N, across the site of discontinuity at site p = 434, on the transfected antigenomic linear RNA. As a positive control, we also assayed for the 115 bp product, L–M, adjacent to the discontinuity. Figure 3B shows that this control was

Fig. 4. Summary of ability of 12 linear RNAs to act as templates for RNA-directed transcription. These antigenomic linear RNAs are each approximately unit length. In each case, the template sequence shown is that at and around the 5' and 3' ends, relative to the rod-like folding of the HDV sequences. In (A) the value of the opening site $p = 434$, while in (B) it is 317. The HDV sequences are shown in upper-case letters with the numbering according to a published sequence (Kuo *et al.*, 1988a). Bullets indicate deleted HDV nucleotides and lower-case letters refer to non-HDV sequence. At the right side, for both panels, is an indication as to whether accumulation of genomic RNA was detected by northern analysis following transfection of each RNA.

actually positive for all samples. Figure 3C, however, which shows the assays for the product spanning the discontinuity, reveals that for $1\times$ alone (lane 2) or in combination with the mRNA for the large delta protein (lane 4) there was no RT-PCR product. In contrast, for $1\times$ in combination with mRNA for the small delta protein there was a product (lane 3), just as for a positive control (lane 5) which, relative to size markers (lane 1), was the expected size of 170 bp.

These studies show that following transfection but in the absence of genome replication, we were unable to detect healing of the input $1\times$ antigenomic RNA at the site p . It

remains possible that such healing occurred but at levels not detected in our assays. With this in mind, similar studies were also carried out for the $1.2\times$ template using an appropriate pair of primers. As described earlier (Figure 1B), this template contains not one but two copies of the HDV ribozyme, and so might be at an advantage for putative prior circularization. Nevertheless, in the absence of replication (that is, with or without the large delta protein, but in the absence of the small delta protein), we again found no evidence for the prior circularization of the transfected template RNA (data not shown).

In summary, the negative RT-PCR results with $1\times$ and $1.2\times$ RNAs support but do not prove the interpretation that prior healing was not occurring. It was only after separate experiments, as described below, that we could exclude the possibility that small amounts of prior healing occurred but failed to be detected by the RT-PCR.

Sequences at the 5' and 3' ends of linear RNA templates can affect the ability to initiate RNA transcription and the 'healing' at the site of discontinuity

Since the above studies did not support the hypothesis that the linear RNA was healed before transcription, we considered the alternative hypothesis, that healing was only achieved on the nascent transcript, by some kind of template-switching mechanism. With this new hypothesis in mind, we tested a series of different linear template RNAs, some of which are summarized in Figure 4.

We considered two values of p , the site of discontinuity. For the RNAs in Figure 4A, $p = 434$. This is the same site as that used in an earlier study (Glenn *et al.*, 1990) in which the unit length HDV DNA template was permuted at a unique *NheI* site. RNAs with $p = 317$ are shown in Figure 4B. This site was chosen because (as explained in Materials and methods) it allowed us to transcribe *in vitro* template RNAs with $x = 0$. We also considered variations in the size and/or composition of the sequences x and y . In Figure 4A, the first five RNAs shown contained a short terminal palindrome, arising because of the *NheI* site. For the next three sequences the repeat size was progressively reduced, until the last RNA, which was 2 nt shorter than unit length. In Figure 4B, the first two sequences were to distinguish between a direct repeat and an inverted repeat, respectively. The remaining two sequences tested exactly unit length and 2 nt short of unit length, respectively.

In each case, 4 days after transfection of HuH7 cells we used northern analysis to detect the accumulation of unit length genomic RNA. The predicted structure and replication ability of 12 of the RNAs tested are summarized in Figure 4.

As indicated, transfection of one of these RNAs did not lead to detectable accumulation of processed genomic RNA (Figure 4A) (i.e. $<0.5\%$ relative to the ability of the $1.2\times$ RNA species discussed earlier). This RNA was 2 nt less than the complete unit length sequence, although another construct with a 2 nt deletion at another location did give accumulation (Figure 4B). Five other RNAs also failed to produce HDV RNA accumulation; these were those with larger values of x and y , together with an ability of their extensive non-HDV sequences to base pair (data not shown).

Fig. 5. Nucleotide sequencing of imperfectly healed genomic RNAs that accumulated following initiation of replication with linear RNA templates. The four templates shown in (A–D) are exactly as in Figure 4B. In each panel, the diagram is of the linear antigenomic RNA used as template, with focus on the site of discontinuity. Beneath this are examples of the sequence of genomic RNA that contained imperfect healing through the site of discontinuity. As in Figure 4, the HDV sequences are shown in upper-case letters with the numbering according to a published sequence (Kuo *et al.*, 1988a). Bullets indicate deleted HDV nucleotides and lower-case letters refer to non-HDV sequence. Letters in lower-case italics represent sequences at the junction on the replicating RNA that are not of HDV origin.

Next, we used RT–PCR, cloning and sequencing to examine the site of healing on those genomic RNAs that accumulated during replication. For the first three cases in Figure 4A, where $p = 434$, we detected only perfect healing (data not shown). In addition, for the second last RNA shown in Figure 4A, we found perfect healing for seven of eight sequenced clones. The exception was a clone with what could be a 1 nt deletion coupled with a 2 nt insertion (data not shown). This result gave us a clue that there could be what we will refer to as ‘imperfect’ healing at the discontinuity site. As will now be explained, many more examples were found when we considered replication initiated with four different linear RNAs, each with the discontinuity at $p = 317$ (Figure 4B).

Figure 5 summarizes the altered sequences obtained for replication of the four RNAs shown in Figure 4B. Each panel shows the sequence of the linear antigenomic RNA template around the site of its discontinuity. Beneath this are the deduced sequences of the corresponding region of the genomic RNAs that were able to replicate and accumulate, and yet were different from the wild-type sequence. Consider the four sets of results.

In Figure 5A, the template RNA had an 8 nt direct repeat at the discontinuity. After replication, seven of nine cloned sequences were perfectly healed to reconstitute the natural genomic HDV sequence. However, as shown, two of nine sequences were imperfectly healed, with a deletion of either 1 or 2 nt.

In Figure 5B, the length of the template was maintained but the 8 nt direct repeat sequence at the 3' end was replaced with an exact inverted repeat of the 5' end. In nine of nine, the healings were imperfect. As can be seen, there were assorted deletions, and combinations of deletions and replacements (replacements are sequences not predicted from the copying of either the HDV or non-HDV sequences of the input RNA).

In Figure 5C, the template was exactly unit length, with $x = y = 0$. Here four of 13 of the replicated sequences were perfectly healed. As shown, nine of 13 had deletions, deletions and replacements, or simply an insertion.

In Figure 5D, the template was 2 nt shorter than unit length. In six of eight sequences, the genomic RNA maintained this 2 nt deletion. The other two, as shown here, had either an additional 4 nt deletion or a 2 nt replacement.

Overall, these studies show that the ability of linear RNA species to achieve initiation and accumulation of processed unit length RNA transcripts was very much determined by the site of discontinuity, p , and by the length and sequence composition of the flanking sequences, x and y (Figure 4). Moreover, even when replication was initiated and circular RNAs formed, the site of healing could be perfect, as in the wild-type sequence, or imperfect, with a variety of small deletions and/or insertions of non-templated sequences (Figure 5). As considered further in the Discussion, these data strongly disfavor a model in which the linear RNA templates are somehow healed before transcription. Instead, they favor a model where the linear RNAs are transcribed and during this transcription, template switching could occur to get across the discontinuity.

Modifications of the 5' and 3' ends of the linear RNA templates did not block healing

For each of the linear RNA templates described so far, the 5' end was a triphosphate and the 3' end a hydroxyl. We were interested in determining whether modification of the structure at these ends might alter the ability of the RNAs to initiate genome accumulation. We considered modification of the 5' end with a cap structure and conversion of the 3' end via pretreatment with sodium meta-periodate.

As one example, consider the case when $p = 317$, with $x = y = 0$, as indicated at the third species in Figure 4B. We transcribed such linear RNAs and tested them with and without a cap structure, and with and without a blocked 3' end. There was no major difference in the ability of these four RNAs to achieve accumulation of genomic

RNA (data not shown). A similar experiment was carried out with the template RNA with $p = 434$, indicated at the top of Figure 4A. Again, there was no effect of 3'- and 5'-end changes on the ability to initiate accumulation of genomic RNA (data not shown). The finding that blocking of the 3' end of the template RNAs did not compromise their ability to initiate RNA synthesis argues against the biological relevance of *in vitro* studies that suggest that pol II can carry out 3'-end addition of HDV RNA templates (Filipovska and Konarska, 2000; Gudima *et al.*, 2000). We did not detect any increase in the frequency of imperfect healing when the 3' and/or 5' ends were blocked (data not shown).

These findings with end-blocked RNAs, together with the earlier RT-PCR studies (Figure 3) argue against a model in which the linear RNA template is healed before transcription. However, as considered in the Discussion, the strongest argument comes from the observations of imperfect healing (Figure 5).

Discussion

We have described here experiments in which linear RNAs containing antigenomic HDV RNA sequences were transfected into mammalian cells and, by a process involving *in vivo* RNA-directed RNA transcription by a host polymerase that is considered to be pol II, this led to the accumulation of processed genomic RNA transcripts that were circular. Our results confirm and significantly extend an earlier observation (Glenn *et al.*, 1990). We report that the length and composition of the sequences at the 5' and 3' ends of the input linear RNAs could have an effect on the ability of such RNAs to initiate replication. What was truly surprising was that for some of the accumulated circular RNA products we could detect sequence changes at the site corresponding to the discontinuity on the input RNA templates. For four such templates, we observed small and specific deletions, with or without the insertion of similar sized non-template sequences (Figure 5). We call this 'imperfect healing'. The deletion of HDV genomic sequences never corresponded to 5' of the discontinuity site on the HDV-specific sequences of the template RNA; it was always immediately 3' of this site. Each of the four templates directed its own characteristic set of insertions/deletions/replacements. For example, in Figure 5B the insertions/deletions/replacements were over a larger range than for Figure 5A and C. Also, in Figure 5C there were mainly small insertions, while in Figure 5A we only detected small deletions, or in most cases, perfect healing.

Our results allowed us to distinguish between two models of how healing occurred. One model was that following transfection some of the linear RNAs were somehow healed to produce circular RNA templates for RNA-directed transcription. Contrary to such a model, by using RT-PCR we were unable to detect such prior healing of template RNAs (Figure 3). However, the strongest evidence against this model was the observations of imperfect healing (Figure 5). This model could not explain the generation of healed transcripts that contained both deletions of template sequences and insertions of non-template sequences. We therefore propose a second model in which imperfect healing is achieved during

transcription through the discontinuity of a linear RNA template. In such a model not only does template switching occur to produce the observed deletions, but there must also be some prior addition of non-templated nucleotides to explain the observed replacements and insertions. Maybe the simplest interpretation of the origin of the extra sequences is that they represent non-templated synthesis added to nascent RNA chains by a stalled polymerase. Interestingly, from an examination of what we interpret as replacement and inserted nucleotides, 50% were adenosine in the genomic transcript. This is like other polymerases that can demonstrate non-templated additions with a preference for adenosine (Magnuson *et al.*, 1996).

The assay we have used involves biological selection of transcriptional 'winners' that are competent for further transcription and ultimate accumulation as processed circular HDV RNA species. However, the frequency at which such winners arose and accumulated genomic RNA transcripts was surprisingly high. It was usually >4% relative to that initiated by transfection of a linear RNA template with a much more extensive direct terminal redundancy of HDV RNA sequences (Figure 2B and data not shown). Therefore, it is tempting to suggest that the polymerase can rescue a large fraction of what might otherwise be failed attempts at template switching. After all, it is known that when pol II in transcribing from a DNA template comes to a transcriptional arrest, the accessory factor TFIIS stimulates hydrolytic removal of the 3' single-stranded RNA, to generate a new 3' end (Palangat and Landick, 2001).

While it is true that the above model might be the first claim for template switching by a mammalian RNA polymerase, presumably pol II, it is true that such switching by other polymerases is well documented. In some cases, it is dependent upon some level of nucleotide homology. For example, reverse transcriptase is considered to need at least 12 nt of direct repeat to achieve an accurate template switch on a retroviral genome (Dang and Hu, 2001; Pfeiffer and Telesnitsky, 2001). The hepadnaviral reverse transcriptase can even use as little as 4 nt (Wang and Seeger, 1992). Along these lines, some of the RNA templates tested in our experiments possessed a direct terminal repeat. The presence and extent of this repeat did enhance correct healing and RNA accumulation. Such results therefore support a homology-dependent form of template-switching model (Nagy *et al.*, 1998). However, for other templates there was little or no such homology and the healing, if observed, was more likely to be imperfect. Such results are more similar to *in vivo* and *in vitro* studies with phage T7 and RNA virus polymerases, where template switching has been observed in the absence of obvious homology, and is sometimes associated with deletion of template sequences and/or insertion of non-templated sequences (Nagy *et al.*, 1998; Rong *et al.*, 1998). Nevertheless, it has been considered that both homology-dependent and -independent forms of template-switching mechanisms may have basic similarities in the role played by the polymerase (Nagy *et al.*, 1998).

All the available evidence suggests that *in vivo* transcription of HDV antigenomic RNA is via pol II (Modahl *et al.*, 2000; Moraleda and Taylor, 2001). Our new data provide unexpected findings regarding what the

polymerase can do when it reaches a discontinuity on the template strand. Similar findings might also be observed for RNA transcription by pol II on DNA templates that have an obstruction such as a discontinuity or nucleotide modification of the template strand. We might speculate that somatic hypermutation in B cells might arise in part via RNA polymerase errors when trying to transcribe through discontinuities in the template strand (Muramatsu and Honjo, 2001). Nevertheless, how pol II functions when acting as an RNA-directed polymerase and its normal functions as a DNA-directed polymerase may have very little in common.

Finally, our studies of RNA-directed RNA transcription, and especially the detection of what we call imperfect healing, might have some implications for the phenomenon of post-transcriptional gene silencing (PTGS) and its possible provocation by what have been vaguely described as 'aberrant RNAs' (Matzke *et al.*, 2001; Tuschl, 2001). Certainly, in plants and some animals, PTGS can involve RNA-directed RNA synthesis by a host polymerase acting on single-stranded RNA templates. However, as mentioned earlier, such a polymerase has not been found in cells of mammalian origin; thus, the RNA-directed transcription of HDV-related sequences remains the only example of such transcription by a mammalian host polymerase.

Materials and methods

In vitro RNA transcription

RNAs needed for transfection were transcribed *in vitro* using phage T7 RNA polymerase using a RiboMax RNA production system (Promega). In one case (Figure 4A, first sequence) the RNA was transcribed from a pGem4Z (Promega) construct. Specifically, a unit length *NheI* fragment of HDV cDNA was inserted into the *XbaI* site of the vector. Before transcription *in vitro* with T7 polymerase, this construct was opened at the *BamHI* site, downstream of the inserted HDV sequences. In all other situations T7 transcription was by expression-PCR. That is, DNA primers were chosen to amplify the regions of HDV sequence described, in all cases placing a T7 promoter sequence 5' of the sequences to be transcribed. Because of the requirements of the T7 promoter, all *in vitro* transcripts begin with GGG; however, in Figures 4B and 5, this also corresponded to a GGG sequence of the HDV RNA species at around position 317. 5'-capped RNAs were generated using a mMessage mMachine kit (Ambion). In some experiments the purified RNA transcripts were then blocked at the 3' end by treatment with a 1000-fold molar excess of sodium metaperiodate (Hunt, 1968). In each RNA transfection, unless stated otherwise, we included a 5'-capped and 3'-polyadenylated mRNA species capable of being translated *in vivo* to produce the small delta protein. This mRNA was obtained by expression-PCR using a primer that placed a 3'-tail of (A)₂₅ on the T7 transcripts.

RNA transfection

Sub-confluent monolayer cultures of a human hepatoblastoma cell line, HuH7 (Nakabayashi *et al.*, 1982), were transfected by RNA combinations, as indicated in the text, using LipofectAMINE Plus (Invitrogen). Hydrodynamics-based transfection of Balb/c mice was carried out exactly as described recently (Chang *et al.*, 2001).

RNA extraction and northern analyses

Total RNA from transfected HuH7 cells or mouse liver was extracted by Tri Reagent (Molecular Research Center). Samples of 4 µg were glyoxalated and submitted to agarose gel electrophoresis and northern assay. In order to separate linear from circular forms of unit length HDV RNA, the gel concentration was raised to 3% (Chen *et al.*, 1986). After electro-transfer the HDV-specific sequences were detected using a 5'-end-labeled synthetic oligonucleotide. Following hybridization, radioactivity was detected using a bio-imager (Fuji).

RT-PCR, cloning and sequencing

RT-PCR was used to detect whether or not HDV species were closed at the site corresponding to the initial discontinuity on the linear HDV antigenomic RNA. It was also used to characterize the 'healing' that occurred on genomic RNAs that accumulated following transfection. In such cases the PCR products so obtained were subjected to TOPO-II cloning (Invitrogen) after which relevant clones were grown up and characterized by automated nucleotide sequencing.

Acknowledgements

We thank Anita Cywinski and the Fox Chase DNA Sequencing Facility. In addition, we thank Tony Lerro and Luis Sigal for the hydrodynamics-based transfections. Constructive comments on the manuscript were given by Robert Perry, Christoph Seeger, William Mason and Eric Moss. This work was supported by grants AI-26522 and CA-06927, from the NIH, and by an appropriation from the Commonwealth of Pennsylvania.

References

- Beard, M.R., Macnaughton, T.B. and Gowans, E.J. (1996) Identification and characterization of a hepatitis delta virus RNA transcriptional promoter. *J. Virol.*, **70**, 4986–4995.
- Bichko, V., Netter, H.J. and Taylor, J. (1994) Introduction of hepatitis delta virus into animal cell lines via cationic liposomes. *J. Virol.*, **68**, 5247–5252.
- Branch, A.D. and Robertson, H.D. (1984) A replication cycle for viroids and small infectious RNAs. *Science*, **223**, 450–455.
- Chang, J., Sigal, L.J., Lerro, A. and Taylor, J. (2001) Replication of the human hepatitis delta virus genome is initiated in mouse hepatocytes following intravenous injection of naked DNA or RNA sequences. *J. Virol.*, **75**, 3469–3473.
- Chao, M., Hsieh, S.-Y. and Taylor, J. (1990) Role of two forms of the hepatitis delta virus antigen: Evidence for a mechanism of self-limiting genome replication. *J. Virol.*, **64**, 5066–5069.
- Chen, P.-J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J. and Taylor, J. (1986) Structure and replication of the genome of hepatitis δ virus. *Proc. Natl Acad. Sci. USA*, **83**, 8774–8778.
- Cogoni, C. and Macino, G. (1999) Gene silencing in *Neurospora crassa* requires a protein homologous to RNA-dependent RNA polymerase. *Nature*, **399**, 166–169.
- Dang, Q. and Hu, W.S. (2001) Effects of homology length in the repeat region on minus-strand DNA transfer and retroviral replication. *J. Virol.*, **75**, 809–820.
- Daros, J., Marcos, J., Hernandez, C. and Flores, R. (1994) Replication of avocado sunblotch viroid: Evidence for a symmetric pathway with two rolling circles and hammerhead ribozyme processing. *Proc. Natl Acad. Sci. USA*, **91**, 12813–12817.
- Diener, T.O. (1989) Circular RNAs: Relics of precellular evolution? *Proc. Natl Acad. Sci. USA*, **86**, 9370–9374.
- Diener, T.O. (1991) Subviral pathogens of plants: viroids and viroidlike satellite RNAs. *FASEB J.*, **5**, 2808–2813.
- Dingle, K., Bichko, V., Zuccola, H., Hogle, J. and Taylor, J. (1998) Initiation of hepatitis delta virus genome replication. *J. Virol.*, **72**, 4783–4788.
- Filipovska, J. and Konarska, M.M. (2000) Specific HDV RNA-templated transcription by pol II *in vitro*. *RNA*, **6**, 41–54.
- Glenn, J.S., Taylor, J.M. and White, J.M. (1990) *In vitro*-synthesized hepatitis delta virus RNA initiates genome replication in cultured cells. *J. Virol.*, **64**, 3104–3107.
- Gudima, S., Wu, S.-Y., Chiang, C.-M., Moraleda, G. and Taylor, J. (2000) Origin of the hepatitis delta virus mRNA. *J. Virol.*, **74**, 7204–7210.
- Hsieh, S.-Y., Chao, M., Coates, L. and Taylor, J. (1990) Hepatitis delta virus genome replication: a polyadenylated mRNA for delta antigen. *J. Virol.*, **64**, 3192–3198.
- Hunt, J.A. (1968) The use of isoniazid for the detection of terminal sequences in periodate-oxidized RNA. *Methods Enzymol.*, **12**, 240–243.
- Kos, A., Dijkema, R., Arnberg, A.C., van der Meide, P.H. and Schellekens, H. (1986) The hepatitis delta (δ) virus possesses a circular RNA. *Nature*, **323**, 558–560.
- Kuo, M.Y.-P., Goldberg, J., Coates, L., Mason, W., Gerin, J. and Taylor, J. (1988a) Molecular cloning of hepatitis delta virus RNA from an infected woodchuck liver: sequence, structure and applications. *J. Virol.*, **62**, 1855–1861.

- Kuo,M.Y.-P., Sharmeen,L., Dinter-Gottlieb,G. and Taylor,J. (1988b) Characterization of self-cleaving RNA sequences on the genome and antigenome of human hepatitis delta virus. *J. Virol.*, **62**, 4439–4444.
- Kuo,M.Y.-P., Chao,M. and Taylor,J. (1989) Initiation of replication of the human hepatitis delta virus genome from cloned DNA: role of delta antigen. *J. Virol.*, **63**, 1945–1950.
- Lazinski,D.W. and Taylor,J.M. (1994) Expression of hepatitis delta virus RNA deletions: *cis* and *trans* requirements for self-cleavage, ligation and RNA packaging. *J. Virol.*, **68**, 2879–2888.
- MacNaughton,T.B., Wang,Y.-J. and Lai,M.M.C. (1993) Replication of hepatitis delta virus RNA: effect of mutations of the autocatalytic cleavage sites. *J. Virol.*, **67**, 2228–2234.
- Magnuson,V.L., Ally,D.S., Nylund,S.J., Karanjawala,Z.E., Rayman,J.B., Knapp,J.I., Lowe,A.L., Ghosh,S. and Collins,F.S. (1996) Substrate nucleotide-determined non-templated addition of adenine by *Taq* DNA polymerase: implications for PCR-based genotyping and cloning. *Biotechniques*, **21**, 700–709.
- Matzke,M., Matzke,A.J. and Kooter,J.M. (2001) RNA: guiding gene silencing. *Science*, **293**, 1080–1083.
- Modahl,L.E. and Lai,M.M.C. (1998) Transcription of hepatitis delta antigen mRNA continues throughout hepatitis delta virus (HDV) replication: a new model of HDV RNA transcription and regulation. *J. Virol.*, **72**, 5449–5456.
- Modahl,L.E., MacNaughton,T.B., Zhu,N., Johnson,D.L. and Lai,M.M.C. (2000) RNA-dependent replication and transcription of hepatitis delta virus RNA involve distinct cellular RNA polymerases. *Mol. Cell. Biol.*, **20**, 6030–6039.
- Moraleda,G. and Taylor,J. (2001) Host RNA polymerase requirements for transcription of the human hepatitis delta virus genome. *J. Virol.*, **75**, 10161–10169.
- Muramatsu,M. and Honjo,T. (2001) Complex layers of genetic alteration in the generation of antibody diversity. *Trends Immunol.*, **22**, 66–68.
- Nagy,P.D., Zhang,C. and Simon,A.E. (1998) Dissecting RNA recombination *in vitro*: role of RNA sequences and the viral replicase. *EMBO J.*, **17**, 2392–2403.
- Nakabayashi,H., Taketa,K., Miyano,K., Yamane,T. and Sato,J. (1982) Growth of human hepatoma cell lines with differentiated functions in chemically defined medium. *Cancer Res.*, **42**, 3858–3863.
- Navarro,J.-A. and Flores,R. (2000) Characterization of the initiation sites of both polarity strands of a viroid RNA reveals a motif conserved in sequence and structure. *EMBO J.*, **19**, 2662–2670.
- Navarro,J.-A., Vera,A. and Flores,R. (2000) A chloroplastic RNA polymerase resistant to targetitoxin is involved in replication of avocado sunblotch viroid. *Virology*, **268**, 218–225.
- Palangat,M. and Landick,R. (2001) Roles of RNA:DNA hybrid stability, RNA structure and active site conformation in pausing by human RNA polymerase II. *J. Mol. Biol.*, **311**, 265–282.
- Pfeiffer,J.K. and Telesnitsky,A. (2001) Effects of limiting homology at the site of intermolecular recombinogenic template switching during Moloney murine leukemia virus replication. *J. Virol.*, **75**, 11263–11274.
- Rakowski,A.G. and Symons,R.H. (1994) Infectivity of linear monomeric transcripts of citrus exoritis viroid: terminal sequence requirements for processing. *Virology*, **203**, 328–335.
- Reid,C.E. and Lazinski,D.W. (2000) A host-specific function is required for ligation of a wide variety of ribozyme-processed RNAs. *Proc. Natl Acad. Sci. USA*, **97**, 424–429.
- Rong,M., Durbin,R.K. and McAllister,W.T. (1998) Template strand switching by T7 RNA polymerase. *J. Biol. Chem.*, **273**, 10253–10260.
- Schiebel,W., Pelissier,T., Riedel,L., Thalmeir,S., Schiebel,R., Kempe,D., Lottspeich,F., Sanger,H.L. and Wasseneger,M. (1998) Isolation of an RNA-directed RNA polymerase-specific cDNA from tomato. *Plant Cell*, **10**, 2087–2101.
- Schindler,I.-M. and Muhlbach,H.-P. (1992) Involvement of nuclear DNA-dependent RNA polymerases in potato spindle tuber viroid replication: a reevaluation. *Plant Sci.*, **84**, 221–229.
- Sharmeen,L., Kuo,M.Y., Dinter-Gottlieb,G. and Taylor,J. (1988) The antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *J. Virol.*, **62**, 2674–2679.
- Symons,R.H. (1992) Small catalytic RNAs. *Annu. Rev. Biochem.*, **61**, 641–671.
- Taylor,J. (1990) Hepatitis delta virus: *cis* and *trans* functions needed for replication. *Cell*, **61**, 371–373.
- Taylor,J.M. (1999) Replication of human hepatitis delta virus: influence of studies on subviral plant pathogens. *Adv. Virus Res.*, **54**, 45–60.
- Tuschl,T. (2001) RNA interference and small interfering RNAs. *Chem. Biochem.*, **2**, 239–245.
- Wang,G.H. and Seeger,C. (1992) The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. *Cell*, **71**, 663–670.
- Wang,K.-S. *et al.* (1986) Structure, sequence and expression of the hepatitis delta viral genome. *Nature*, **323**, 508–513.
- Wu,T.-T., Netter,H.J., Lazinski,D.W. and Taylor,J.M. (1997) Effects of nucleotide changes on the ability of hepatitis delta virus to transcribe, process and accumulate unit length, circular RNA. *J. Virol.*, **71**, 5408–5414.

Received September 26, 2001; revised November 9, 2001;
accepted November 19, 2001