

---

# The self-cleaving domain from the genomic RNA of hepatitis delta virus: sequence requirements and the effects of denaturant

---

Anne T. Perrotta and Michael D. Been\*

Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

---

Received September 24, 1990; Revised and Accepted November 11, 1990

---

## ABSTRACT

The sequence requirements for self-cleavage of hepatitis delta virus genomic RNA were examined using precursor RNAs which were labeled at either the 5' or 3' ends and progressively deleted from the unlabeled end. In the presence of 50% formamide, which enhances self-cleavage in 2 mM MgCl<sub>2</sub> at 37°C, 84 nucleotides (nt) 3' of the break site were required. In the absence of formamide the minimum was reduced to 82 nt. Under both sets of conditions, precursors with 1 nt 5' to the break site cleaved. These results allowed two condition-dependent minimal domains for self-cleavage to be defined. However, in the absence of formamide, sequences flanking the minimal domain inhibited cleavage, possibly through involvement in the formation of non-cleaving structures. These data are consistent with the idea that cleavage *in vivo* could be regulated by alternative RNA structures.

## INTRODUCTION

Hepatitis delta virus (HDV) is a human pathogen with several unusual facets to its replication and structure (1). The HDV genome is a small (~1700 nt) circular single-stranded RNA (2–5) with extensive self-complementarity that allows the RNA to fold into an unbranched rod-like structure under non-denaturing conditions (6). Many of these features resemble those of some pathogenic RNAs of plants. An additional similarity to the plant pathogens is that genomic and antigenomic strands of HDV RNA contain sites that are capable of autocatalytic cleavage (self-cleavage) *in vitro* (7–9).

A number of RNAs can fold into structures capable of mediating either RNA cleavage (10–15) or concerted RNA cleavage and ligation (eg. self-splicing) *in vitro* (16–18). Self-cleavage reactions that are not associated with self-splicing can be mediated by relatively small RNA structures (19). Although many of these self-cleaving RNAs can be folded into a consensus secondary structure known as a hammerhead (15), other structures can also facilitate RNA self-cleavage reactions (20–23). Two recently described examples of non-hammerhead self-cleaving RNAs are those found in the genomic and the antigenomic sequences of HDV (7–9).

The self-cleavage of RNA from the HDV genomic sequence is enhanced by adding denaturants (5–7 M urea or 40–60% formamide) to the reaction (24). Moreover, in the absence of denaturants, the rate of cleavage at the genomic site is greatest at low Mg<sup>2+</sup> concentrations (<0.1 mM) or at elevated temperatures (>50°C). Together, these data were taken to indicate that conditions destabilizing for RNA structure may enhance the cleavage reaction. The mechanism for the enhancement is not known but, as suggested for other self-cleaving RNAs, it is possible that, upon folding, the RNA may form structures incapable of cleavage (15, 25). If such is the case for HDV RNA, the denaturants could promote cleavage by preferentially destabilizing the inactive forms or, if subsequent steps are fast when the RNA is in the cleavable conformation, by increasing the rate of interconversion of the various structures. Other mechanisms by which a denaturant could affect the rate of an RNA catalyzed reaction have been discussed by Zaug et al. (26).

Here we describe minimal sequences required for cleavage at the HDV genomic strand self-cleavage site and present evidence that sequences that fall outside of this domain can inhibit the cleavage reaction. The addition of denaturants appears to relax the inhibitory effect of flanking sequences. A short sequence at the 3' end of the minimal cleavage domain may confer extra stability to the self-cleaving structure. Together, these features could account for the enhanced rates of self-cleavage in the presence of denaturants.

## MATERIALS AND METHODS

### Enzymes, Nucleotides, Oligonucleotides and Reagents

T7 RNA polymerase was purified from an over-expressing clone provided by W. Studier (27). Restriction endonuclease *EcoRI* was a gift from P. Modrich (Duke University). Other enzymes, nucleotides and <sup>32</sup>P-labeled nucleotides were purchased from commercial sources. Oligonucleotides were synthesized by D. Wright (Duke University) on an Applied Biosystems DNA synthesizer. Formamide was recrystallized, and acrylamide solutions were deionized with mixed bed resin; all other chemicals were purchased from commercial sources and used as supplied.

---

\* To whom correspondence should be addressed

### Plasmid DNA

The plasmid pDM1X was constructed and kindly provided to us by M. Belinsky and G. Dinter-Gottlieb (Drexel University). It contains 135 nt of HDV sequence (5), position 651 to 786 on the map used by Wang et al. (2), inserted into the *PstI/XbaI* sites of pGEM4Z (Promega) such that transcription with T7 RNA polymerase yields a transcript with 19 nt of vector-derived sequence at the 5' end followed by the genomic HDV sequence and vector derived sequence at the 3' end, the length of which depends on the restriction enzyme site used. *Escherichia coli* (strain JM83) was transformed with pDM1X. Plasmid DNA was prepared from overnight cultures by boiling lysis and purified by CsCl equilibrium density centrifugation in the presence of ethidium bromide (28).

The other plasmid (pSD106) was generated by cloning a polymerase chain reaction (PCR) product (29) into the *SmaI* site of pUC19. The PCR primers were designed to generate a T7 promoter that overlapped the genomic HDV sequence 5' to the position of the break site such that the T7 RNA polymerase transcript began with nucleotide -3 (see Figure 1B). Miniprep DNA from several clones were sequenced as previously described (30). The construct used in this study contained the entire sequence of the intended PCR product and the orientation was such that runoffs were made after digesting the plasmid with *EcoRI*. Plasmid DNA was prepared as described above.

### Preparation of Precursor RNA

Plasmid DNA was linearized with a restriction endonuclease (*XbaI* or *EcoRI*) and the reaction was stopped by adding EDTA to 25 mM. The products were extracted once with buffered phenol and once with chloroform, and the DNA was recovered by ethanol precipitation. The conditions used for transcription were: 40 mM Tris-HCl (pH 7.5), 15 mM MgCl<sub>2</sub>, 2 mM spermidine, 5 mM dithiothreitol (DTT), ribonucleoside triphosphates at 1 mM each, 50 µg/ml linear plasmid DNA, and 250 units of T7 RNA polymerase/µg of DNA. To make labeled transcript from pSD106 DNA, [ $\alpha$ -<sup>32</sup>P]CTP was added to 0.5 mCi/ml. After 60 min at 37°C, EDTA was added to 25 mM, formamide to 50%, and the RNA was fractionated by electrophoresis on a 6% polyacrylamide gel containing 7 M urea. RNA was located by UV shadowing or autoradiography and the precursor was excised and the RNA was eluted overnight at 4 °C in an elution mix containing either 0.75 M ammonium acetate, 1 mM EDTA, and 0.1% sodium dodecyl sulfate, or 10 mM EDTA and 0.1% sodium dodecyl sulfate. The RNA was recovered from the eluate by ethanol precipitation.

### Preparation of End-labeled Transcripts

To make 5' end-labeled precursor RNA, 2 µg of gel purified RNA was dephosphorylated with 1 unit of calf intestinal phosphatase at 37°C for 30 min in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. The reaction was extracted twice with buffered phenol and once with chloroform. The RNA was recovered by ethanol precipitation and resuspended in 7.5 µl of H<sub>2</sub>O and labeled in a 10 µl reaction containing 22 pmoles of [ $\gamma$ -<sup>32</sup>P]ATP (7000 Ci/mmole), 50 mM Tris-HCl (pH 8.9), 10 mM MgCl<sub>2</sub>, 5 mM DTT, and 15 units of T4 polynucleotide kinase. Incubation was for 2 hours on ice (0–4°C) and the reaction was terminated by addition of EDTA to 25 mM and formamide to 66%. The RNA was gel purified as described above.

RNA was labeled at the 3' end with [5'-<sup>32</sup>P]pCp (31). Labeled pCp was made by incubating 1 nmole of 3' CMP with 100 pmoles

of [ $\gamma$ -<sup>32</sup>P]ATP and 30 units of T4 polynucleotide kinase in 50 mM Tris-HCl (pH 8.0), 10 mM MgCl<sub>2</sub>, 15 mM DTT (final vol, 25 µl) at 37°C for 30 min followed by heating at 70°C for 5 min. The RNA was labeled in a 20 µl reaction containing 2 µg of RNA, 0.1 mM ATP, 10 units of T4 RNA ligase, 50 mM Hepes (pH 7.5), 3.3 mM DTT, 15 mM MgCl<sub>2</sub>, 10% dimethylsulfoxide, and 5 µl of the above pCp reaction. It was incubated on ice for 16 hours and stopped by adding EDTA to 25 mM and formamide to 66%. The labeled RNA was gel purified as described above.

### RNA Sequencing Reactions

Sequencing markers were generated by partial digestion of the end-labeled RNAs with T1 or U2 ribonuclease as described (32). The alkaline cleavage ladder was generated by heating RNA in 25 mM sodium bicarbonate, 1 mM EDTA (pH 9.0) to 100°C for 5 min.

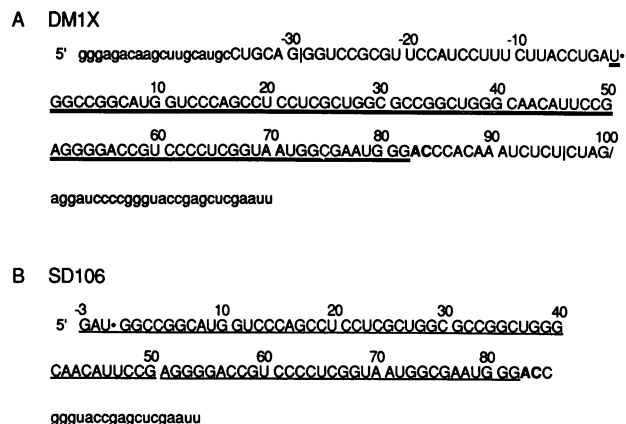
### Gel Electrophoresis

Sequencing gels contained 8% polyacrylamide (29:1 acrylamide:bis-acrylamide), 0.1 M Tris-borate (pH 8.3), 1 mM EDTA and either 7 M urea or 7 M urea and 40% formamide. The latter conditions were required to resolve fragments of 3' end-labeled RNA that migrated in a region near the cleavage product. Cleavage of internally labeled SD106 precursor was assayed on 6% polyacrylamide gels containing 7 M urea, 50 mM Tris-borate (pH 8.3), and 0.5 mM EDTA. Following electrophoresis, the gel was transferred to filter paper, dried, and an autoradiogram prepared using X-ray film.

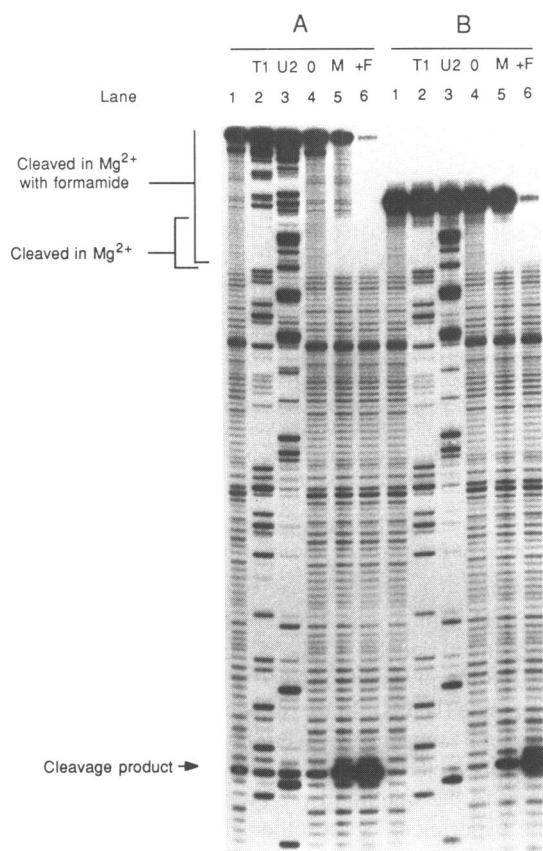
## RESULTS

### The Minimal Sequence Requirement 3' of the Break Site Varies with Reaction Conditions

Initial studies were carried out with precursor RNA transcribed from pDM1X. The effects of denaturants, Mg<sup>2+</sup> concentration, and reaction temperature on the rate and extent of cleavage of this transcript have been described (24). To identify minimal HDV sequence requirements for self-cleavage, we used the experimental approach described by Forster and Symons (19) in their study on the self-cleaving RNA from the virusoid of lucerne transient streak virus (vLTSV). Precursor RNAs (Figure 1A) were generated by transcribing pDM1X DNA digested with different restriction endonucleases. These precursors were 5' end-labeled and gel purified. A portion of this RNA was then partially cleaved at random positions by alkaline treatment in the absence of Mg<sup>2+</sup> to generate a set of labeled fragments which would share a common 5' end. On a sequencing gel, this randomly cleaved RNA appeared as a 'ladder' of fragments, each differing by 1 nt at the 3' end (Figure 2, lanes 1). Randomly cleaved RNA was then incubated at pH 7.5 in Mg<sup>2+</sup> (2 mM MgCl<sub>2</sub> and 0.25 mM EDTA) to test for self-cleavage. Specific RNA fragments capable of self-cleaving were identified by their decrease in or their absence from the experimental lane. The smallest fragment that cleaved in the absence of denaturant had 82 nt 3' of the break site (lanes 5). In the presence of 50% formamide this minimum size increased to 84 nt (lanes 6) (this difference will be difficult to see in the reproduced photograph, it is more apparent in Figures 5 and 6, below). After cleavage, the <sup>32</sup>P-labeled RNA accumulated at the position expected for 5' cleavage product (arrow).



**Figure 1.** Sequences of the Transcripts Made from Plasmids pDM1X and pSD106. Upper case letters are HDV sequence; lower case are vector sequences; the heavy dot marks the break site and numbering is from this position. The double line. The sequences shown are for transcripts made from *EcoRI* cut plasmid DNA, the diagonal slash indicates the end of the transcript made from *XbaI* cut pDM1X DNA.



**Figure 2.** Self-Cleavage of 5' End-labeled DM1X Precursor: Identification of Required and Inhibitory Sequences 3' of the Break Site. Runoff precursor DM1X RNAs made from pDM1X DNA digested with *EcoRI* (set A) or *XbaI* (set B) were 5' end-labeled and gel purified. Portions were partially digested with RNase T1 (G ladder, lanes 2) or RNase U2 (A ladder, lanes 3), or randomly cleaved by alkaline treatment (lanes 1). The alkaline treated RNA was diluted and the pH adjusted to 7.5 by the addition of Tris-HCl (pH 7.0). The randomly cleaved RNA was incubated for 15 min at 37°C in either 50 mM Tris-HCl (pH 7.5) and 0.25 mM EDTA (lanes 4), or in the same buffer containing in addition either 2 mM MgCl<sub>2</sub> (lanes 5), or 2 mM MgCl<sub>2</sub> and 50% formamide (lanes 6). In each case the reactions were terminated by adding an equal volume of 95% formamide containing 25 mM EDTA. Samples were analyzed by electrophoresis in an 8% polyacrylamide gel containing 7M urea.

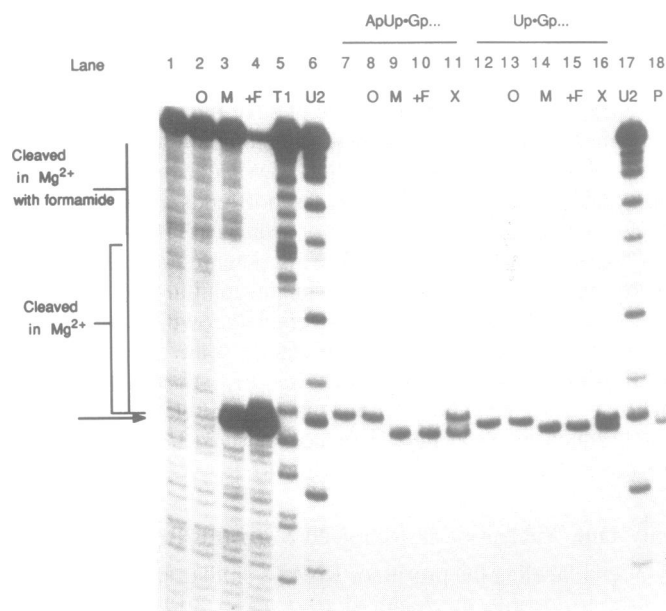
### HDV Sequences can Inhibit Cleavage in the Absence of Denaturants

With the *EcoRI* runoff transcript, cleavage was inefficient for those fragments with more than 96 nt 3' of the break site unless formamide was added to the reaction (compare tops of lanes 5 and 6, Figure 2A). Note that the HDV sequence extends to nucleotide position 100. The *XbaI* runoff ends at position 100 (Figure 1A) so there are no 3' vector sequences. However, there was little or no cleavage of the full length *XbaI* transcript. Therefore, in the absence of formamide, inhibition of cleavage was seen with fragments that terminated within the HDV sequence. It is not known to what extent, if any, the inhibition by flanking regions is dependent on the specific sequence; questions concerning potential for inhibition of cleavage by more distal sequences, especially sequences derived from the vector, are not addressed with this experiment.

### Only One Nucleotide is Required 5' to the Break Site

By 3' end-labeling the precursor RNA, a similar approach could be used to identify the required sequence 5' to the break site. As before, a ladder of fragments was generated and these were tested for the ability to self-cleave at pH 7.5 after the addition of Mg<sup>2+</sup> (Figure 3). There is a set of fragments that cleaved in the absence of denaturants; the 5' end of these fragments extended from a position near the break site to about 29 nt 5' to the break site (lane 3). Following cleavage, the <sup>32</sup>P-labeled product migrated with the 3' fragment (arrow). Longer fragments which did not cleave in the absence of denaturant, cleaved when 50% formamide was included in the reaction (lane 4). The shortest fragment that inhibited cleavage in the absence of denaturants contained only HDV sequences at the 5' end. Again, the potential role of non-HDV sequences in inhibition of cleavage cannot be determined from this experiment.

From data shown in lanes 2–4 of Figure 3 it was apparent that, at most, only a few nucleotides were required 5' to the break site. The exact number was difficult to determine because the product obscured the region of interest. To determine minimal 5' requirements, it was essential to test uncleaved RNA that contained only 1 or 2 nt 5' to the break site. To make these precursors, a 3' end-labeled DM1X transcript (*XbaI* runoff) was partially digested with RNase T1 (cuts 3' of G) and RNase U2 (cuts 3' of A). The sequence preceding the break site is ...GpApUp\* (Figure 1) so that truncated precursor RNAs, containing just the dinucleotide (ApUp) or the mononucleotide (Up) 5' to the break site, could be generated in these reactions. Both of these RNAs, when gel purified and incubated with Mg<sup>2+</sup>, produced a fragment that co-migrated with purified 3' cleavage product (Figure 3, lanes 9, 14, and 18). This demonstrated that both RNAs were capable of cleavage and that one nucleotide 5' to the break site was sufficient for the self-cleavage reaction. These results also make it unlikely that cleavage of the shorter fragments in the hydrolysis mixture was catalyzed by larger versions of the RNA acting enzymatically. The addition of formamide to the reaction did not alter the result (lanes 10 and 15). Thus, even under conditions that affected the 3' sequence requirements, a single nucleotide 5' to the break site was sufficient for cleavage. However, [5'-<sup>32</sup>P]phosphorylated 3' cleavage product showed no evidence for auto-dephosphorylation in 2 and 10 mM MgCl<sub>2</sub> at pH 5, 6, 7, or 8 (data not shown), so a 5'-terminal phosphate group at the break site was not labile. The results of the above experiments on the 5' and 3' end-labeled RNA are summarized for the DM1X sequence in Figure 1A.

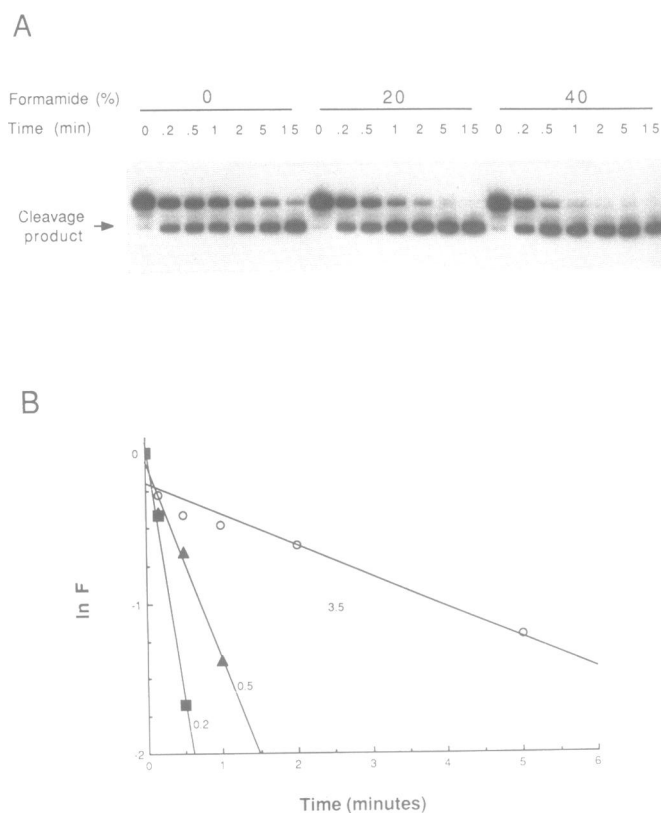


**Figure 3.** Self-cleavage of 3' End-labeled DM1X Precursor: Identification of Required and Inhibitory Sequences 5' to the Break Site. Lanes 1–6: *Xba*I runoff DM1X RNA was 3' end-labeled and gel purified. Portions were partially digested with RNase T1 (lane 5) or RNase U2 (lane 6) or randomly cleaved by alkaline treatment and the pH adjusted as described in the legend to Figure 2. The randomly cleaved RNA (lane 1) was incubated for 15 min at 37°C in either 50 mM Tris-HCl (pH 7.5) and 0.25 mM EDTA (lane 2), or in the same buffer together with either 2 mM MgCl<sub>2</sub> (lane 3) or 2 mM MgCl<sub>2</sub> and 50% formamide (lane 4). In each case, the reactions were terminated as described for Figure 2. The arrow indicates the position of the labeled cleavage product. Lanes 7–18: RNA fragments containing either 2 nt (lane 7) or 1 nt (lane 12) 5' to the break site were gel purified from T1 and U2 partial digests, respectively. The RNA fragments were incubated for 15 min at 37°C in either 50 mM Tris-HCl (pH 7.5) and 0.25 mM EDTA (lanes 8 and 13), or in the same buffer together with 2 mM MgCl<sub>2</sub> (lanes 9 and 14) or 2 mM MgCl<sub>2</sub> and 50% formamide (lanes 10 and 15). After stopping the cleavage reactions as described above, portions of the reactions shown in lanes 8 and 9 were mixed and loaded on to lane 11, and portions of 13 and 14 were mixed and loaded onto lane 16. The T1 fragment was analyzed in lanes 8–11 and the U2 fragment was analyzed in lanes 13–16. Lane 18 contained gel purified 3' end-labeled cleavage product. Samples were analyzed on an 8% polyacrylamide sequencing gel containing 7 M urea and 10 M formamide; a similar gel was used to purify the specific T1 and U2 fragments.

### Characterization of the Domain 3' of the Break Site

The above experiments independently defined minimal sequences required either 5' to or 3' of the site of cleavage. However, it is possible that portions of both the 5' and the 3' domain could act autonomously to provide sequence or structure for cleavage. If so, then the non-essential 5' and 3' sequences, as defined above, could not be deleted in the same precursor without loss of cleavage activity. To test this, a T7 promoter sequence was incorporated just 5' to the position of the site of cleavage such that transcription with T7 RNA polymerase generated a precursor with only the sequence pppGpApUp preceding the break site. This precursor RNA (SD106, Figure 1B) also had 85 nt of HDV sequence 3' of the break site followed by cloning site vector sequence. SD106 precursor RNA cleaved efficiently in the presence of Mg<sup>2+</sup> (Figure 4A). The release of a 5' fragment of the correct size was confirmed by analyzing the cleavage products of 5' end-labeled precursor on a 20% polyacrylamide sequencing gel (data not shown).

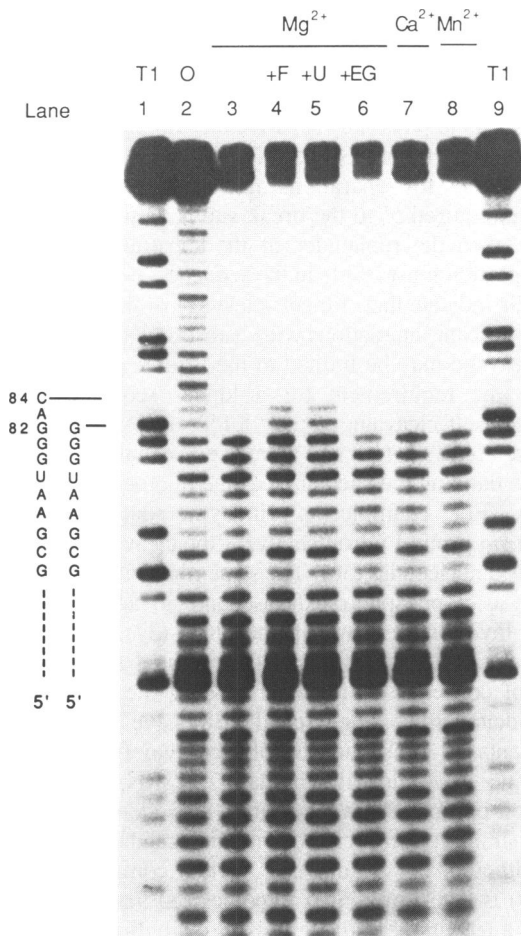
To test if deleting all but 3 nt of the 5' portion would affect the size of the required 3' domain, the 3' deletion experiment



**Figure 4.** Self-cleavage of SD106 Precursor RNA. **A.** Internally-labeled SD106 RNA was synthesized using [ $\alpha$ -<sup>32</sup>P] CTP so that no label is incorporated 5' to the cleavage site (see Figure 1). The gel-purified precursor was incubated for 5 min at 37°C in 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, and either 0%, 20% or 40% formamide. The self-cleavage reaction was initiated by adding MgCl<sub>2</sub> to 3 mM and aliquots were removed at the indicated times and stopped with 3 volumes of 95% formamide containing 25 mM EDTA and tracking dyes. Samples were run on a 6% polyacrylamide gel containing 7 M urea, the gel was dried and exposed to X-ray film. **B.** RNA was located using the autoradiogram, excised and the results quantified by counting Cherenkov scintillation. Results are expressed as the natural log of the fraction of uncleaved precursor (ln F) as a function of time. Open circles, no formamide; closed triangles, 20% formamide; closed squares, 40% formamide. The numerical value adjacent to each line is the half time ( $t_{1/2}$ , min) of the precursor calculated from the the slope ( $-k$ ).

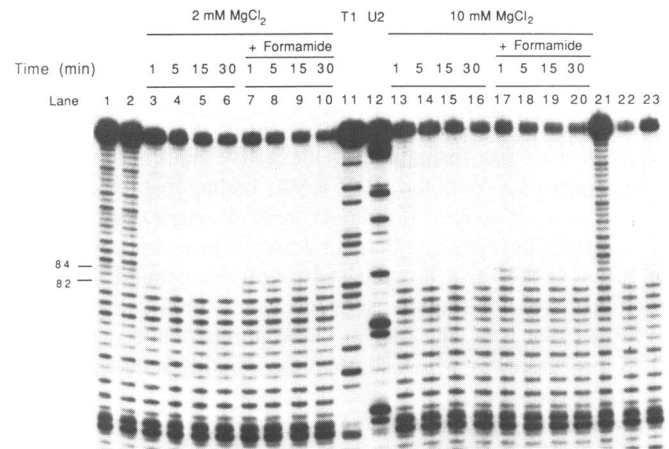
was repeated with 5' end-labeled SD106 precursor RNA. An alkali generated ladder of the 5' end-labeled precursor was incubated in the presence of Mg<sup>2+</sup> (Figure 5, lane 3) or Mg<sup>2+</sup> and formamide (lane 4). In the absence of denaturant, a fragment with 82 nt of 3' sequence cleaved, but removal of one additional nucleotide prevented cleavage. Adding formamide to 50%, again added 2 nt to the minimum required for efficient cleavage. In the same experiment, some other conditions for self-cleavage were surveyed. Adding urea to 5 M (lane 5) had an effect similar to the formamide, whereas adding ethylene glycol to 50% did not (lane 6). [Ethylene glycol at 50% will stimulate the rate of cleavage of the DM1X precursor but it does so to a lesser extent than either 5 M urea or 50% formamide (unpublished results).] Either Mn<sup>2+</sup> or Ca<sup>2+</sup> could substitute for Mg<sup>2+</sup> in the cleavage reaction and, in the absence of denaturants, the RNA fragments capable of cleavage were the same as with Mg<sup>2+</sup> (compare lane 3 with lanes 7 and 8). These results are summarized for the sequence of SD106 in Figure 1B.

All of the above reactions using alkali generated fragments of 5' end-labeled SD106 precursor RNA were incubated for 15 min in buffers containing 2 mM MgCl<sub>2</sub>. However, cleavage is time



**Figure 5.** Self-cleavage of 5' End-labeled SD106 Precursor RNA; Identification of Required Sequences. Sequence requirements for self-cleavage of SD106 precursor RNA was examined. An RNase T1 partial digest (G ladder) is shown in lanes 1 and 9. Randomly cleaved RNA was incubated at 37°C for 15 min in 50 mM Tris-HCl (pH 7.5), 0.25 mM EDTA (lane 2). Other reactions, in addition, contained: 2 mM MgCl<sub>2</sub> (lane 3); 2 mM MgCl<sub>2</sub>, 50% formamide (lane 4); 2 mM MgCl<sub>2</sub>, 5 M urea (lane 5); 2 mM MgCl<sub>2</sub>, 50% ethylene glycol (lane 6); 2 mM CaCl<sub>2</sub> (lane 7); or 2 mM MnCl<sub>2</sub> (lane 8). The sequence at the 3' end of the shortest fragments to cleave in the absence or presence of 50% formamide is shown. The 5' cleavage product was run off the gel.

dependent and can be relatively slow under some conditions (24). The cleavage of a 5' end-labeled SD106 ladder was followed in a 30 min time course (Figure 6). The patterns indicated that in 2 mM MgCl<sub>2</sub> without formamide, most fragments ending at position 83 had cleaved in 5 min (lane 4), and most fragments ending at position 82 had cleaved by 15 min (lane 5). Less obvious perhaps, after 30 min in 10 mM MgCl<sub>2</sub> without formamide, there was a small decrease in the amount of the the fragment containing only 81 nt of 3' sequence (lane 16) suggesting that it was slowly being cleaved. However, raising the MgCl<sub>2</sub> concentration to 20 mM (lane 22) only slightly enhanced cleavage of this particular fragment. For the most part, to effect rapid cleavage, 82 nt appear to be a lower limit for the 3' domain. In 2 mM MgCl<sub>2</sub> and formamide there was some cleavage of the fragment containing 84 nt of 3' sequence after 30 min (lane 10) but it was incomplete. The effect of 50% formamide could be partially offset by raising the MgCl<sub>2</sub> concentration from 2 to 10 mM (lanes 7-10 and 17-20) in that some of the fragments containing 82 and 83 nt of 3' sequence were cleaved with the higher MgCl<sub>2</sub> concentration. Therefore, it appears that the



**Figure 6.** Time course for cleavage of 5' end-labeled SD106 ladder at low and high MgCl<sub>2</sub> concentrations. The 5' end-labeled SD106 ladder was prepared as described for Figure 5. The RNA in lane 1 received no further treatment. The RNA in lanes 2 and 21 were incubated for 30 min in 50 mM Tris-HCl (pH 7.5) 0.25 mM EDTA at 37°C. For the cleavage reaction, RNA was incubated in the same buffer with the addition of 2 mM MgCl<sub>2</sub> (lanes 3-10), or 10 mM MgCl<sub>2</sub> (lanes 13-20) and 50% formamide (lanes 7-10 and 17-20). Aliquots were removed at 1, 5, 15 and 30 min as indicated and the reaction terminated by the addition of an equal volume of 95% formamide containing 25 mM EDTA. Lane 22 contained RNA incubated for 30 min at 37°C in 20 mM MgCl<sub>2</sub>, and lane 23 contained RNA incubated at 30°C in 2 mM MgCl<sub>2</sub>. Lanes 11 and 12 contained sequencing ladders as indicated.

nucleotides at positions 83 and 84 are not essential for cleavage activity but may stabilize the structure required for cleavage.

### Deleting Distal Flanking Sequences Can Enhance the Rate of Cleavage

With the DM1X precursor, denaturants enhance the cleavage reaction (24). One explanation of the enhancement effect, consistent with results presented in this paper is that sequences outside of the cleavage domain inhibit the cleavage reaction only in the absence of denaturants. If so, deleting the inhibitory sequence should increase the rate of cleavage. The data for cleavage of SD106 precursor from the experiment shown in Figure 4A were quantified to determine the rates of cleavage (Figure 4B). At 37 °C in 2 mM Mg<sup>2+</sup> and in the absence of denaturant, the half time ( $t_{1/2}$ ) of the reaction was about 3.5 min. This is approximately a 20-30 fold enhancement over the rates of cleavage seen with the longer (DM1X) precursors under similar conditions ( $t_{1/2}$  = 65-120 min) (24, A.T.P. and M.D.B. unpublished data). However, cleavage of the SD106 precursor is still enhanced by the addition of formamide ( $t_{1/2}$  = 0.5 min in 20% formamide and  $t_{1/2}$  = 0.2 min in 40% formamide). The SD106 precursor contained 18 nt of 3' vector derived sequence and, if this vector sequence is inhibitory, it may account for the residual enhancement upon addition of formamide to the reaction. Alternatively, it could be that formamide also acts at another step in the reaction, a step which becomes rate limiting with the SD106 precursor.

### DISCUSSION

This work was prompted by the unexpected finding that denaturants dramatically enhance the rate of self-cleavage at the HDV genomic RNA self-cleavage site (24). It would now appear that at least part of this effect is the result of destabilizing inhibitory structures generated by flanking sequences. Inhibition

of cleavage by flanking sequences is not without precedent. The appearance of a 'window' of cleavable fragments in the alkali generated ladder was also seen in the sequence 5' to the break site of the vLTSV RNA (19). Those authors hypothesized that the formation of a stable hairpin structure involving 5' sequence interferes with the formation of the active hammerhead self-cleaving structure. With the HDV RNA, finding that the cleavage of many of the longer fragments with 3' ends distal to the boundaries of the cleavage domain could be promoted by adding denaturant would also be consistent with the general idea that structures involving flanking sequences may inhibit cleavage. In support of that idea, the rate of cleavage is enhanced with the SD106 precursor for which some of the flanking sequences have been removed.

This study has defined the boundaries of the minimal sequence required for efficient self-cleavage of the HDV genomic-strand RNA. The methods used do not address the question of what sequences within these boundaries are required for activity. The accurate determination of these boundaries, however, will be useful both in generating a model for the secondary structure of the RNA and in designing experiments aimed at further reducing the size of the cleavage domain through internal deletions. The minimum is dependent on the reaction conditions. In the absence of denaturant, the smallest fragment that efficiently cleaved extended from position -3 to 82, but fragments containing only one nucleotide 5' to the break site also cleaved. Potentially, the minimum extends from position -1 to 82 although, due to the difficulty in generating purified labeled substrate with that specific sequence, that fragment has not been tested. In the presence of 50% formamide or 5 M urea, and low concentrations of  $Mg^{2+}$ , two additional nucleotides are required for efficient self-cleavage. The smallest fragment that cleaved efficiently under these conditions extended from -3 (or potentially from -1) to 84. In the presence of denaturants, the additional nucleotides may be required to stabilize an interaction required for the formation of the cleavable structure. It is possible that the observed essential sequence could be reduced further by conditions that provide additional stability to the structure. However, in the absence of denaturants, higher  $Mg^{2+}$  did not dramatically reduce the size of the 3' sequence necessary for cleavage. It is also necessary to keep in mind that, if the self-cleavage reaction is physiologically important, it must occur in the context of precursors much larger than the fragments used to define the minimal cleavage domains. As such, it is possible that additional sequences, not required for the reactions *in vitro*, are essential to either generate the active structure or to suppress the formation of inhibitory structures *in vivo*.

It would appear that the sizes of the self-cleaving domains, as we have defined them, differ significantly from results published by Kuo et al. (7). In that paper it was reported that self-cleavage of the genomic strand RNA requires 30 nt 5' and 74 nt 3' of the break site. However, to identify sequences required 5' to the break site, those authors used primer extension on hydrolysis products after cleavage. That method would not have identified a minimal requirement of only a single nucleotide 5' to the break site. In addition, that method may be subject to other artifacts as a result of template fragments in the hydrolysis mixture. With regard to the 3' domain, careful inspection of patterns in the hydrolysis ladder (such as compressions and enhanced cleavages 5' to Gs) shown in Figure 4 of Kuo et al. (7) indicates that those authors got results similar to those of Figure 2 of this paper (82 nt required), but in the absence of markers they underestimated the size of the smallest fragment

that cleaved. We therefore consider it unlikely that there are actual inconsistencies with the results of Kuo et al. (7), but the correct numbers are as reported here.

It has often been possible with other self-cleaving and self-splicing RNAs to physically separate the 'ribozyme' into enzyme and substrate portions (33-35). For HDV self-cleaving RNA, similarly successful separations may be possible. If the single nucleotide required 5' to the break site is viewed as part of the substrate, then the remainder of the substrate and the entire catalytic portion must reside in the sequence 3' of the break site. Our results indicate that, for self-cleavage of the HDV genomic RNA, significant interactions with 'substrate' sequence 5' to the site of cleavage may be limited to the uridine at the break site. Although the requirement for a longer sequence might be anticipated, self-cleavage of the vLTSV RNA requires only 3 nt 5' to the break site (19) and there is precedent for even shorter sequences flanking the sites of cleavage in other RNA catalyzed reactions. For example, a ribozyme derived from the *Tetrahymena* group I intron catalyzes the cleavage of substrates as small as dinucleotides (36). The same ribozyme is even capable of acting as a phosphotransferase and an acid phosphatase, reactions involving a terminal phosphate (37).

The accurate determination of the minimal cleavage domain may be of consequence to understanding the pathogenesis of hepatitis delta virus. Negro et al. (38) have noted a striking complementarity of sequence between two regions of HDV antigenomic RNA and the human 7SL RNA of the signal recognition particle. This, of course, also identifies regions of sequence similarity between the genomic strand of HDV and 7SL RNA. Although the significance of this complementarity (or similarity) is not known, it is hypothesized that the viral RNA may play a role in the pathogenesis through direct action on the RNA component of the signal recognition particle. Two hypothesized mechanisms for this action involve either formation of a duplex with the antigenomic strand (38) or trans-cleavage of 7SL RNA by an HDV genomic RNA ribozyme (39). We would like to draw attention to the fact that the region of similarity between the 7SL RNA (nt 10-55) and the genomic strand of HDV (nt 683-724, numbered according to Wang et al., (2)) falls almost entirely within the 5' half of the minimal self-cleaving domain of HDV RNA and includes the break site (nucleotide positions -3 to 39 relative to the site of cleavage). While it is possible that these similarities or complementarities are the molecular basis of HDV pathogenesis, an additional interesting prospect is that the sequence relationships allow HDV RNA to subvert cellular factors for virus replication.

## ACKNOWLEDGEMENTS

We thank P. Pavco, S. Rosenstein, and T. Cech for many helpful comments on the manuscript, C. Fuller for suggesting the formamide containing gels and G. Vergara for making prints. This work was supported by a grant from the NIH (GM-40689). M.D.B. was supported by a Junior Faculty Research Award from the American Cancer Society (JFRA-233).

## REFERENCES

1. Taylor, J.M. (1990) *Cell* **61**, 371-373.
2. Wang, K.-S., Choo, Q.-L., Weiner, A.J., Ou, J.-H., Najarian, R.C., Thayer, R.M., Mullenbach, G.T., Denniston, K.J., Gerin, J.L., and Houghton, M. (1986) *Nature* **323**, 508-514.
3. Chen, P.-J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., and Taylor, J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8774-8778.



4. Makino, S., Chang, M.-F., Shieh, C.-K., Kamahora, T., Vannier, D.M., Govindarajan, S., and Lai, M.M.C. (1987) *Nature* **329**, 343–346.
5. Kuo, M.Y.P., Goldberg, J., Coates, L., Mason, W., Gerin, J., and Taylor, J. (1988) *J. Virol.* **62**, 1855–1861.
6. Kos, A., Dijkema, R., Arnberg, A.C., van der Meide, P.H., and Schellekens, H. (1986) *Nature* **323**, 558–560.
7. Kuo, M.Y.-P., Sharmeen, L., Dinter-Gottlieb, G., and Taylor, J. (1988) *J. Virol.* **62**, 4439–4444.
8. Sharmeen, L., Kuo, M.Y.P., Dinter-Gottlieb, G., and Taylor, J. (1988) *J. Virol.* **62**, 2674–2679.
9. Wu, H.-N., Lin, Y.-J., Lin, F.-P., Makino, S., Chang, M.-F., and Lai, M.M.C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1831–1835.
10. Epstein, L.M., and Gall, J.G. (1987) *Cell* **48**, 535–543.
11. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) *Cell* **35**, 849–857.
12. Zaug, A.J., Kent, J.R., and Cech, T.R. (1984) *Science* **224**, 574–578.
13. Prody, G.A., Bakos, J.T., Buzayan, J.M., Schneider, I.R., and Bruening, G. (1986) *Science* **231**, 1577–1580.
14. Hutchins, C.J., Rathjen, P.D., Forster, A.C., and Symons, R.H. (1986) *Nucleic Acids Res* **14**, 3627–3640.
15. Forster, A.C., and Symons, R.H. (1987) *Cell* **49**, 211–220.
16. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982) *Cell* **31**, 147–157.
17. Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A., and Cheng, H.-L. (1986) *Cell* **44**, 213–223.
18. Van der Veen, R., Arnberg, A.C., Van der Horst, G., Bonen, L., Tabak, H.F., and Grivell, L.A. (1986) *Cell* **44**, 225–234.
19. Forster, A.C., and Symons, R.H. (1987) *Cell* **50**, 9–16.
20. Brown, R.S., Hingerty, B.E., Dewan, J.C., and Klug, A. (1983) *Nature* **303**, 543–546.
21. Rubin, J.R., and Sundaralingam, M. (1983) *Journal of Biomolecular Structure and Dynamics* **1**, 639–646.
22. Haseloff, J., and Gerlach, W.L. (1989) *Gene* **82**, 43–52.
23. Hampel, A., Tritz, R., Hicks, M., and Cruz, P. (1990) *Nucleic Acids Res.* **18**, 299–304.
24. Rosenstein, S.P., and Been, M.D. (1990) *Biochemistry* **28**, 8011–8016.
25. Ruffner, D.E., Dahm, S.C., and Uhlenbeck, O.C. (1989) *Gene* **82**, 31–41.
26. Zaug, A.J., Grosshans, C.A. and Cech, T.R. (1988) *Biochemistry* **27**, 8924–8931.
27. Davanloo, P., Rosenberg, A.H., Dunn, J.J., and Studier, F.W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2035–2039.
28. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor University Press, Cold Spring Harbor).
29. Innis, M.A., Gelfand, D.H., Sninsky, J.T., and White, T.J. (1990) *PCR protocols: a guide to methods and applications*. Academic Press, Inc.
30. Been, M.D., and Cech, T.R. (1985) *Nucleic Acids Res.* **13**, 8389–8408.
31. Bruce, A.G., and Uhlenbeck, O.C. (1978) *Nucleic Acids Res.* **5**, 3665–3677.
32. Donis-Keller, H., Maxam, A.M., and Gilbert, W. (1977) *Nucleic Acids Res.* **4**, 2527–2538.
33. Zaug, A.J., and Cech, T.R. (1986) *Science* **231**, 470–475.
34. Zaug, A.J., Been, M.D., and Cech, T.R. (1986) *Nature* **324**, 429–433.
35. Uhlenbeck, O.C. (1987) *Nature* **328**, 596–600.
36. Kay, P.S., and Inoue, T. (1987) *Nature* **327**, 343–346.
37. Zaug, A.J., and Cech, T.R. (1986) *Biochemistry* **25**, 4478–4482.
38. Negro, F., Gerin, J.L., Purcell, R.H., and Miller, R.H. (1989) *Nature* **341**, 111.
39. Young, B., and Hicke, B. (1990) *Nature* **343**, 28.