A Sequence Element Necessary for Self-Cleavage of the Antigenomic Hepatitis Delta RNA in 20 M Formamide

Janet B. Smith,[‡] Philip A. Gottlieb,[§] and Gail Dinter-Gottlieb*,[‡]

Department of Bioscience and Biotechnology, Drexel University, Philadelphia, Pennsylvania 19104, and Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716

Received January 17, 1992; Revised Manuscript Received May 29, 1992

ABSTRACT: The genomic and antigenomic RNAs of hepatitis delta virus are capable of self-cleavage and show no significant sequence similarities to other known self-cleaving RNAs. We have derived an antigenomic delta RNA which cleaves to completion in 15 s in 9 mM magnesium at 37 °C and is capable of efficient self-cleavage in concentrations of formamide as high as 20 M. Cleavage in high concentrations of denaturant is dependent upon the presence of a polypurine sequence element, GGAGA, located between 81 and 85 nucleotides downstream of the cleavage site. Mutation of the initial $G_{81}G_{82}$ to $C_{81}C_{82}$, or removal of the sequence element, results in a loss of the ability to cleave in high formamide concentrations. Changing the final $U_{-2}C_{-1}$ of a pyrimidine-rich region, UCUUC, just upstream of the cleavage site, to $G_{-2}G_{-1}$ severely affects the self-cleavage, but introducing the two mutations, GG to CC and UC to GG, into the same molecule, restoring potential base pairing, partially restores the formamide stability. Relocating the GGAGA sequence upstream of the cleavage site also results in partial restoration of the formamide cleavage. Although the GGAGA sequence is important for self-cleavage under denaturing conditions, it does not appear to be necessary for HDV RNA cleavage in normal buffer conditions.

Hepatitis delta virus (HDV) is a unique human pathogen, whose 1679-nucleotide genome is a covalently closed circular RNA. Replication of hepatitis delta virus is carried out by an unknown polymerase, but appears to involve a rolling circle mechanism, similar to that seen with certain plant pathogens, the viroids, and the plant satellite virus RNAs (Branch & Robertson, 1984; Chen et al., 1986; Kos et al., 1986). Thus, the infectious genomic strand is copied into multimeric antigenomic RNAs, which are processed to monomeric viral units and circularized. The monomeric antigenomic RNAs are then copied into multimeric genomic strands, which are subsequently processed to monomeric circles. We and others earlier demonstrated that both the hepatitis delta virus genomic and antigenomic multimeric RNAs are capable of self-cleavage, in vitro (Sharmeen et al., 1988; Kuo et al., 1989; Wu et al., 1989). Truncated forms of the antigenomic RNA of hepatitis delta displayed the ability to cleave in high concentrations of denaturants such as formamide and urea, although such cleavage was dependent upon the presence of a 10-nucleotide region at the 3′ terminus of the molecule (Smith & Dinter-Gottlieb, 1991). These 10 nucleotides allowed the antigenomic delta RNAs to completely cleave in 18 M formamide and to cleave to 50% in 20 M formamide. Removal of these 10 nucleotides produced a molecule which cleaved poorly in concentrations of formamide above 10 M. In this work, we have further defined, through mutagenesis, a region of the hepatitis delta antigenomic RNA which confers the ability to self-cleave in high concentrations of formamide, in magnesium concentrations as low as 1 mM. Although the element GGAGA is required for stability in 20 M formamide, it is not required for cleavage of the antigenomic delta RNA in normal buffer conditions.

MATERIALS AND METHODS

Polymerase Chain Reaction. For mutations and a 3′ truncation reaction the polymerase chain reaction (PCR) was used, as previously described (Smith & Dinter-Gottlieb, 1991).

DNA Sequencing Reactions. The protocol contained in the Taq Sequencing Kit was followed with some revisions. When sequencing the 3′ end of hepatitis delta virus (HDV) cDNA, 1.5–2.0 μg of purified plasmid was used. A T7 primer was used in all 3′ sequence reactions. The amount of Taq polymerase was increased from the recommended 2.5 units/μL to 3.3 units/μL. Extension reactions were performed at 45 °C for 1 min, followed by a 5-min termination reaction at 75 °C. The samples were heat denatured at 90 °C for 5 min, quick cooled on ice, and loaded onto an 8% polyacrylamide/7 M urea sequencing gel. This gel was prerun at 45 mA for 30 min prior to loading the samples. Once the samples were loaded, the gel was run at 45 mA with constant power adjustment to a maximum of 60 W.

Sequencing of the 5′ end of HDV cDNA required less template, so 1 μg of plasmid was used in each reaction. The plasmid DNAs were extended at 37 °C for 1 min, followed by a 5-min termination reaction at 75 °C. The samples were heat denatured as above, but were loaded onto a 10% polyacrylamide/7 M urea sequencing gel. The gel running conditions were as noted above.

Transcription of the RNA in Vitro and Self-Cleavage Analysis. Transcription of the delta RNAs and self-cleavage analysis, in the presence and absence of formamide, were carried out as previously described (Smith & Dinter-Gottlieb, 1991). All experiments were repeated at least three times, and the values derived from densitometry readings of the autoradiographs were averaged. The averaged values differed by less than 10%.

* Author to whom correspondence should be addressed. Telephone: (215) 895-2631; FAX: (215) 895-1273; email: gottlieb@duvm.ocs.drexel.edu.

[‡] Drexel University.

[§] University of Delaware.

0006-2960/92/0431-9629$03.00/0

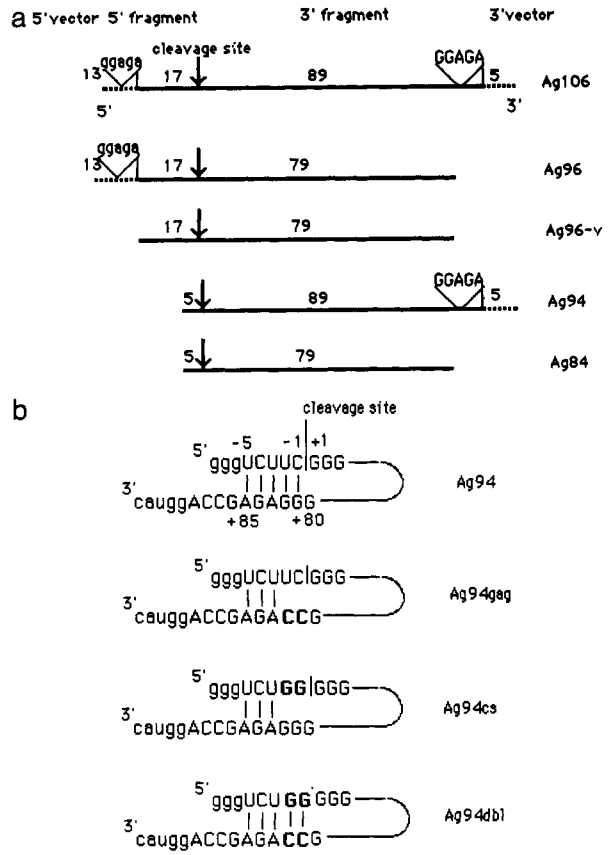

FIGURE 1: (a) Antigenomic RNAs from hepatitis delta virus which were tested for self-cleavage activity and stability in high concentrations of formamide. The cleavage site is marked by an arrow. Some RNAs contained 5′ or 3′ vector sequences, designated by dotted lines. Vector sequences are written in lower case letters. (b) Proposed pairing between the GGAGA region at the 3′ end of the molecule and the 5 nucleotides upstream of the cleavage site (Ag94). The effects of the various mutations on the pairing are shown below.

RESULTS

Five HDV Nucleotides Upstream of the Cleavage Site Are Sufficient for Cleavage in High Concentrations of Formamide. We proposed earlier that the ability of the 106-nucleotide (nt) antigenomic HDV RNA Ag106 (Figure 1) to self-cleave in 20 M formamide depended on interactions between the pentanucleotide GGAGA, located near the 3′ terminus of the RNA, and portions of the polypyrimidine stretch immediately 5′ to the cleavage site (Smith & Dinter-Gottlieb, 1991). Mutagenesis of nucleotides involved in proposed interaction has now been carried out and the effect on self-cleavage in formamide evaluated. The mutations were as follows (Figure 1): first, truncation of the 5′ region of the molecule to produce Ag94; second, altering the $G_{81}G_{82}$ in the terminal region to $C_{81}C_{82}$ (Ag94gag); third, changing the $U_{-2}C_{-1}$ at the cleavage site to $G_{-2}G_{-1}$ (Ag94cs); and finally, introducing both the Ag94gag and Ag94cs mutations into the same molecule (Ag95dbl).

Since the original Ag106 RNA molecule contained 13 vector nucleotides and 17 pyrimidine-rich delta nucleotides upstream of the cleavage site, offering a number of sites for GGAGA to pair (Smith & Dinter-Gottlieb, 1991), we truncated the RNA to leave only five delta nucleotides, UCUUC, upstream of the cleavage site. This RNA transcript, Ag94, retained the ability to cleave to 65% in 20 M formamide in 5 min (Figure 2a,c). Furthermore, this Ag94 RNA cleaved to completion in 15 s at 37 °C in the absence of formamide (Table I). This was in comparison to other delta RNAs, which cleaved rapidly, but never to completion, or cleaved to completion, but only over an extended period (Belinsky & Dinter-Gottlieb, 1991; Perrotta & Been, 1991; Smith & Dinter-Gottlieb, 1991; Smith, Gottlieb, Prasad, and Dinter-Gottlieb, in preparation).

Mutagenesis of a 3′ GG Eliminates the Ability To Cleave in High Formamide Concentrations. To test the importance of the 3′ sequence in the interaction, mutations were introduced in the terminal 10-nucleotide stretch which is necessary for the cleavage in 20 M formamide (Smith & Dinter-Gottlieb, 1991). Removal of the five terminal nucleotides at the 3′ end of the molecule had no effect on the cleavage reaction (data not shown), so attention was focused on the GGAGA sequence. Changing the GG at positions 81 and 82 to a CC (Figure 1b, Ag94gag) abolished the cleavage in high formamide concentrations (Figure 2b,c). In fact, the cleavage was inhibited above 6.8 M formamide, in contrast to Ag94, which was unaffected by concentrations of formamide up to 18 M. This mutation also had an effect on cleavage in the absence of formamide, reducing the rate of cleavage at 37 °C by 50-fold (Table I).

Mutagenesis at the Cleavage Site Inhibits Self-Cleavage, but Increasing Concentrations of Formamide Restore the Activity. Changing the $U_{-2}C_{-1}$ nucleotides just upstream of the cleavage site in Ag94 to $G_{-2}G_{-1}$ (Figure 1b, Ag94cs) had a striking effect on the self-cleavage of the antigenomic RNA. In the absence of formamide, virtually no cleavage was seen in 10 min at 37 °C, while the wild type cleaved to almost 100% (Figure 2a,c). With addition of formamide up to 11.1 M, the cleavage was enhanced, peaking at 55% in 5 min at 37 °C. Concentrations higher than 11.1 M resulted in diminished cleavage, which we attributed to formamide destabilizing the structure of the RNA. In the absence of the denaturant, less than 30% of the molecules cleaved, even after 60-min incubation (data not shown).

Because the mutation at the cleavage site severely affected the self-cleavage, the proposed model for the GGAGA interaction near the cleavage site (Smith & Dinter-Gottlieb, 1991) was difficult to assess fully. Nevertheless, the compensatory mutation was introduced. The self-cleavage activity in the absence of denaturant was restored in the Ag95dbl mutant, but not the cleavage in the highest concentrations of formamide.

Reestablishing the potential base-pairing between the GGAGA region and the cleavage site in Ag95dbl (Figure 1b) restored cleavage activity comparable to that of the Ag94gag mutant in the absence of formamide (Figure 2c and Table I). In the absence of denaturant, cleavage of 50% of the molecules was seen in 5 min, and cleavage continued until 90% cleavage was achieved in 60 min at 37 °C, approaching wild-type levels (Smith, Gottlieb, Prasad, and Dinter-Gottlieb, in preparation).

Formamide affected the Ag95dbl mutant in a fashion similar to the Ag94gag mutant. Cleavage to 50% in 5 min was seen in concentrations of formamide up to 6.8 M (Figure 2c), after which the cleavage reaction rapidly diminished, indicating that denaturation of the active site had occurred.

An Upstream GGAGA Can Confer Some Stability in Formamide. Subcloning the delta RNA molecules occasionally caused vector sequences, which may alter the self-cleavage activity, to be incorporated into the resulting RNA transcripts (Belinsky & Dinter-Gottlieb, 1991). Analysis of the Ag106 and Ag96 transcripts (Figure 1a) revealed that a portion of the upstream vector sequence, GGAGA, precisely duplicated the downstream 3′ GGAGA sequence that allows cleavage of Ag106 in high concentrations of formamide. As previously described (Smith & Dinter-Gottlieb, 1991), Ag96,

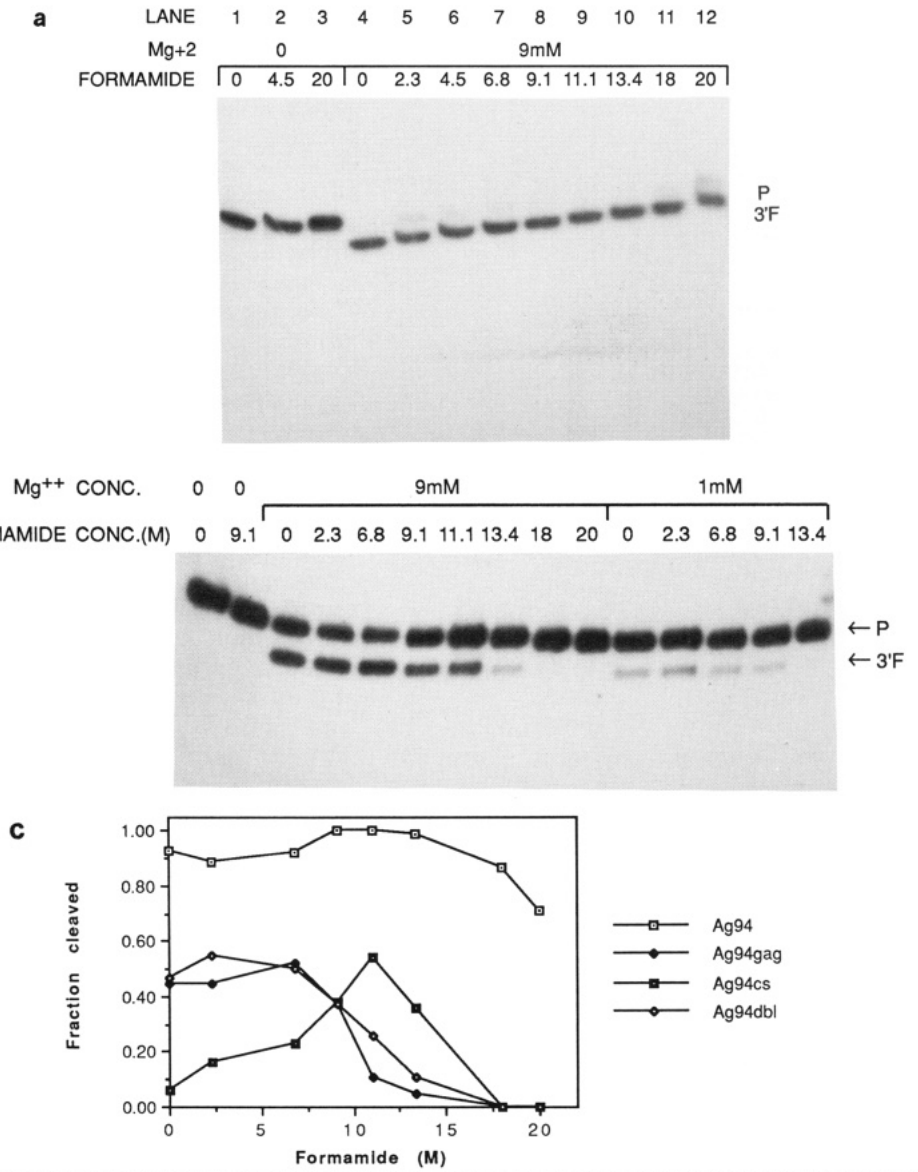

FIGURE 2: (a) Polyacrylamide gel electrophoresis of the cleavage products of the wild-type Ag94 with increasing formamide concentrations in 9 mM magnesium. Lanes 1–3 were incubated in the absence of magnesium and show no cleavage. P designates the presursor RNA, and 3′F the 3′ product. The 5′ product was run off the gel. (b) Analysis of Ag94gag in the presence of increasing formamide. Lanes 1 and 2 were incubated in the absence of magnesium. Reactions shown in lanes 3–10 were incubated in the presence of 9 mM magnesium, while those in lanes 11–15 contained 1 mM magnesium. (c) Effect of increasing formamide concentrations on the self-cleavage of the Ag94 RNA transcript (□) and mutants Ag94gag (♦), Ag94cs (■), and Ag95dbl (◇).

Table I: First-Order Rate Constants for Cleavage of Antigenomic RNAs in 10 mM Magnesium at 37 °C[a]

Ag RNA	k (min^{-1})	k_{rel}[c]	Ag RNA	k (min^{-1})	k_{rel}[c]
Ag94	10	1.0	Ag94dbl	0.37	0.04
Ag94gag	0.19	0.02	Ag96	3.2	0.32
Ag94cs	*b*	*b*	Ag96-v	0.77	0.07

[a] With the exception of Ag94 and Ag96, which cleaved to completion in less than 15 s, the RNAs showed a biphasic curve of cleavage, with a rapid initial cleavage followed by slower phase. We interpret this to reflect the proportion of molecules in an active form which cleave rapidly, followed by those which must undergo a transition to the active form. These values reflect only the first phase. On the other hand, Ag96cs showed cleavage of less than 3% for the first 5 min, then gradually began to cleave, and reached only 20% cleavage after 60 min of incubation. The mutant RNAs Ag94gag and Ag94dbl, as well as Ag96-v, show a $t_{1/2}$ of cleavage of 5 min, but the remainder of the RNA takes 60 min to complete cleavage. Thus a $t_{1/2}$ of cleavage is also an unsatisfactory measure for comparison of activities. [b] Less than 0.001 min^{-1}. [c] $k_{rel} = k(\text{test})/k(\text{Ag94})$; k_{rel} is unitless.

which lacks the 3′ GGAGA, showed inhibition of cleavage above 11.1 M formamide, with total inhibition of cleavage by 20 M formamide. Although the downstream GGAGA was absent in Ag96, the presence of the upstream GGAGA might have stabilized the reaction in formamide. We now show that, with removal of the upstream GGAGA (Ag96-v, Figure 1a), addition of formamide caused a reduction in self-cleavage activity (Figure 3). The presence of the GGAGA region, either upstream or downstream of the cleavage site, appeared to stabilize the RNA structure in formamide, but the optimal stabilization was only achieved when the GGAGA was positioned at the 3′ terminus of the molecule.

In the absence of formamide, Ag96-v showed cleavage equivalent to the Ag94gag or Ag95dbl mutants (Table I). While the cleavage in high formamide clearly required all five nucleotides intact, the cleavage reaction in the absence of formamide could still occur, even when the downstream GGAGA was removed.

Cleavage in High Formamide Is Possible in 1 mM Magnesium, If an Intact GGAGA Is Present. Changing two nucleotides in the Ag94gag mutant destroyed the ability of the RNA to self-cleave in formamide concentrations above

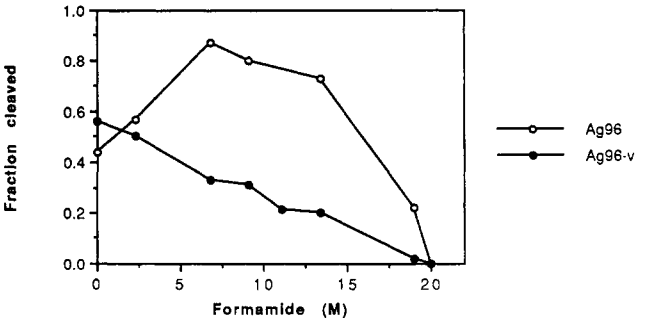

FIGURE 3: The position of the GGAGA region affected stability in formamide. In Ag96 (○) the GGAGA sequence is at the 5′ end of the molecule, and in Ag96-v (●), it has been removed. The resulting RNA shows decreased cleavage in formamide.

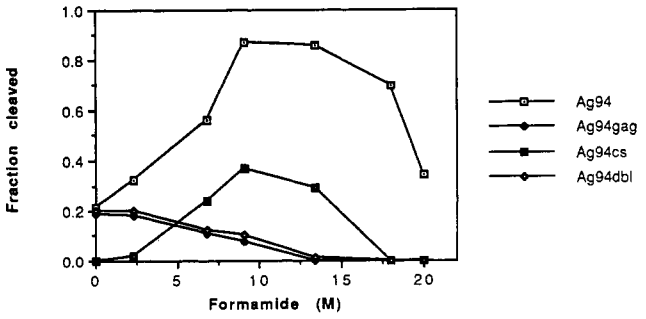

FIGURE 4: Stability in formamide of Ag94 and mutants in 1 mM magnesium. Cleavage profiles in increasing formamide of Ag94 (□), Ag94gag (♦), Ag94cs (■), and Ag95dbl (◇) RNAs are shown.

13 M, while the Ag94 RNA cleaved to almost 80% in 20 M formamide (Figure 2). The protocol for cleavage in formamide requires a 5-min incubation at 37 °C in the formamide, then addition of 9 mM magnesium, and a further 5-min incubation (Rosenstein & Been, 1990; Smith & Dinter-Gottlieb, 1991). It was suggested that the 9 mM magnesium was providing structural stability, compensating for the increasing concentrations of formamide. To test the effect of the magnesium, only 1 mM magnesium was added after the 5-min preincubation in formamide. The Ag94 RNA retained the ability to cleave to 40% in 20 M formamide. Increasing the concentration of formamide enhanced the cleaved reaction, with maximal cleavage only seen in 10–13 M formamide, and loss of activity above 13 M formamide (Figures 2b and 4). The mutation of the $U_{-2}C_{-1}$ at the cleavage site (Ag94cs) diminished the self-cleavage, but cleavage was restored by increasing the formamide concentration, as had been seen with the higher concentration of magnesium (Figure 4). Both the mutation in the GGAGA region (Ag94gag) and the compensatory Ag95dbl mutations showed cleavage in the absence of formamide equivalent to the Ag94 RNA, at 20% (Figure 4). But while the Ag94 and Ag94cs mutant cleavage was enhanced by the addition of formamide, the cleavage of the gag and dbl mutants was abolished by increasing the formamide concentration in the presence of low magnesium ion concentration.

DISCUSSION

The Truncated HDV Is an Efficient Self-Cleaving RNA. We have derived an RNA molecule (Ag94) from the antigenomic RNA of deltahepatitis virus which is able to cleave rapidly and completely, and whose self-cleaving activity is relatively unaffected by formamide. The self-cleaving activity of catalytic RNAs derived from HDV has been shown to be sensitive to changes in concentrations of formamide (Rosenstein & Been, 1990; Belinsky & Dinter-Gottlieb, 1991; Smith & Dinter-Gottlieb, 1991). For example, the antigenomic HDV RNA containing 62 nucleotides 5′ and 119 nucleotides 3′ of the cleavage site (Ag180; Smith & Dinter-Gottlieb, 1991) showed cleavage of an increasing proportion of the RNA as the concentration of formamide was increased. This phenomenon, in which catalytic capacity was enhanced by increasing the concentration of a denaturant, can be explained if the self-cleaving RNA is assumed to adopt alternative RNA structures. Catalytically inactive HDV RNAs can arise from the formation of alternative secondary and tertiary interactions. These inactive structures are kinetically trapped, and formamide presumably lowers the barrier for the interconversion between the alternative structures and the active form or forms. The presence of such alternative structures has been demonstrated for the hammerhead RNA (Behlen & Uhlenbeck, 1991) and may play a role in other RNA molecules (Tang & Draper, 1990).

Mutation of Two Nucleotides 5′ of the Cleavage Site Reintroduced Competing, Inactive RNA Foldings. The change of $U_{-2}C_{-1}$ at the cleavage site to GG (Ag94cs), in the absence of any other mutations, severely inhibited the self-cleavage reaction (Figures 1b and 2c), reducing cleavage in the absence of formamide to less than 10%. Addition of 11.1 M formamide allowed cleavage of 50% of the molecules, implying that inactive, competing structures are forming at the lower concentration of formamide. Above 11.1 M, the activity is inhibited, with no cleavage observed at formamide concentrations above 18 M. This mimics the results obtained by Perrotta and Been (1991), who changed the $U_{-5}C_{-4}$ upstream of the cleavage site to AA, altering the polypyrimidine stretch and resulting in a population of molecules incapable of cleaving beyond 50% in 10 min at 37 °C, even in the presence of 10 M formamide.

The measured inactivation effect of the Ag94cs mutant RNA is clearly related to the choice of conditions used in the assay (Figure 2c). With no formamide present or in low concentrations of formamide the Ag94cs RNA cleaves poorly. At higher concentrations, increased formation of an active structure is evidenced by the increase in the amount of cleaved RNA. This makes interpretation of the effect of nucleotide substitutions on catalytic activity difficult to assess using only one set of conditions, since inactivation may be attributed to either a trapped inactive structure or an RNA molecule that is incapable of forming a catalytically competent structure under any conditions.

Alternative Pairings Are Possible between the GGAGA and Three Regions within the Molecule. The sequence of Ag94 RNA has three possible regions which can interact with GGAGA (Figure 5). These are located just upstream of the cleavage site (region 1), 15 nucleotides downstream of the cleavage site (region 2), and 22 nucleotides downstream of the cleavage site (region 3). The exact pairing needed for folding the RNA into a correct structure is currently not known, although a pseudoknot has been postulated for both the genomic and the antigenomic HDV RNA ribozymes (Perrotta & Been, 1991).

The two base changes producing Ag94gag (Figures 1b and 5b) allowed cleavage, though less efficiently than Ag94 (as indicated by the $t_{1/2}$ of 5 min instead of less than 15 s), while not providing the formamide stability. Formamide addition did not enhance the cleavage of these RNA molecules (Figures 2b,c). From 0 to 6.8 M formamide Ag94 and Ag94gag showed no change in their percent cleavage. The absence of any formamide effect on Ag94 or Ag94gag indicated that no competing structures capable of interconversion in formamide

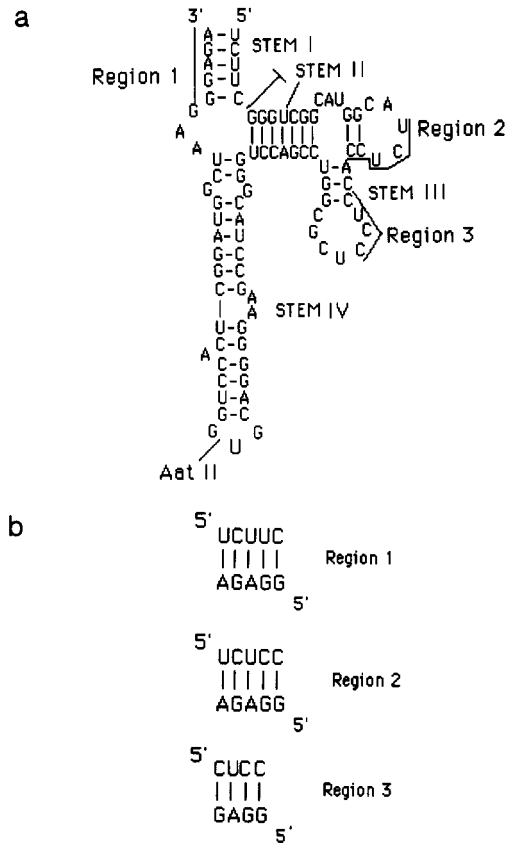

FIGURE 5: (a) A proposed folding of the Ag94 wild-type RNA, with the GGAGA/region 1 pairing and alternate regions 2 and 3 marked. This folding was derived by locating conserved base-paired regions which are also found in the genomic RNA. (b) Potential GGAGA pairing with regions 1, 2, and 3 in the Ag94 RNA molecule.

were formed. Above 6.8 M formamide, inhibition of the cleavage activity of Ag94gag began, with complete inhibition by 15 M formamide. These results indicated that an interaction important for self-cleavage as well as for formamide stability was disrupted by the $G_{81}G_{82}$ to CC change in the GGAGA sequence.

The alteration of the two nucleotides 5′ to the cleavage site in Ag94cs reduced the cleavage activity to less than 20% of wild type in the absence of formamide (Figure 2c). This mutation leaves the GGAGA sequence intact while mutating region 1. The addition of formamide up to 11.1 M enhanced the cleavage reaction, possibly by melting out alternate structures. However, only 50% of the molecules cleaved, similar to the results by Perrotta and Been (1991) when the delta $U_{-5}C_{-4}$ was changed to a vector AA upstream of the cleavage site. These five nucleotides upstream of the cleavage site thus appear to be essential for the most active conformation at the cleavage site. Above 11.1 M, with the Ag94cs mutant, inhibition of cleavage is seen.

Creation of the pseudorevertant, Ag95dbl, in an attempt to reestablish base-pairing in the region of the GGAGA and region 1, and consequently, formamide stability, proved to be complex. The introduction of the two-base change to restore pairing in the dbl mutant (Figure 1b) did not restore the "wild-type" cleavage kinetics of Ag94, although it did return the cleavage in the presence and absence of formamide to the levels seen with the Ag94gag mutant (Figure 2c). Again, only 50% cleavage was seen, with no enhancement by formamide. Thus, the new pairing, in which UC/GG pairs were replaced by GG/CC pairs, was not equivalent either in the cleavage reaction or in formamide stability to Ag94. This effect on the cleavage rate is not surprising, since changing a pyrimidine–purine pair to a purine–pyrimidine pair can alter the base stacking, affecting the configuration of the cleavage site. Mutations on the hammerhead RNA have shown that pyrimidine–purine to purine–pyrimidine changes near the cleavage site affected catalytic activity (Ruffner & Uhlenbeck, 1990; Perrault et al., 1991). The increased activity of the Ag94dbl mutant over the Ag94cs mutant, in the absence of formamide (Figure 2c), may be due to the stabilizing effect of five base pairs just upstream of the cleavage site (Figure 1b).

Mutagenesis of the GGAGA and the Cleavage Site Limited Possible Alternative Pairings. Our results support a preferential base-pairing between the GGAGA and region 1 in Ag94. This may be due to other secondary and tertiary foldings within the molecule. When the GG in the GGAGA is mutated to a CC (Ag94gag), only 3 base pairs are possible with region 1, 3 base pairs with region 2, and 2 base pairs with region 3 (Figure 5b). Region 1 may still be preferred, since inactive foldings are not apparent, and the other pairings are also diminished, but the rate of cleavage is slower, and the stability in high formamide has been lost.

The Ag94cs mutant has only 3 base pairs with the GGAGA, while regions 2 and 3 have the original 5 and 4 base-pairings, respectively. Thus, inhibitory pairings may be formed with either region 2 or region 3, or both, and formamide denaturation is needed to restore the active configuration, as indicated by Figure 2c. Finally, the dbl mutant allows significant pairing only with region 1, but the cleavage is impaired because of the alteration in the sequence of the bases near the cleavage site.

The Cleavage in High Formamide Is Not Dependent upon High Concentrations of Magnesium. The cleavage in high formamide occurs even when the concentration of magnesium ions is reduced to 1 mM. Since the RNA was preincubated for 5 min in the formamide alone, before the magnesium was introduced, this indicated that either the addition of low magnesium was sufficient to restabilize the active structure or the structure itself was resistant to denaturation in formamide. The predicted folding of the RNA projects a structure in which three RNA helices join at a junction which includes the cleavage site (Figure 5a), similar to the overall structure of the hammerhead ribozymes. In Ag94, two of the delta helices are 7 and 9 contiguous base pairs in length. The third helix is only 5 base pairs, yet it is this structure, in conjunction with the other two helices, or in an unknown tertiary interaction, which allows the proper configuration for stability in formamide, even in low magnesium concentrations.

It may be of interest that the region spanning the cleavage site, 5′ UCUUCGGG, has the sequence of a highly conserved tetraloop, UUCG, with a two-base-pair stem (Figure 6a; Tuerk et al., 1988). Melting studies of the tetraloop with a three-base stem have shown unusual stability at high temperatures (Tuerk et al., 1988), and NMR studies have revealed unique base stacking of the four nucleotides in the tetraloop (Cheong et al., 1990). The structure that we propose would compete with the internal stacking of the tetraloop (Figure 6a), and by reducing the potentials for base-pairing in the mutant RNAs, for example, in the Ag94gag mutant, we may have induced tetraloop formation, which would inhibit self-cleavage through the formation of an inactive RNA conformation.

The Nature of the GGAGA Interaction. Denaturing agents such as formamide can disrupt nucleic acid structure by interfering with the ability of nucleotides to form noncovalent

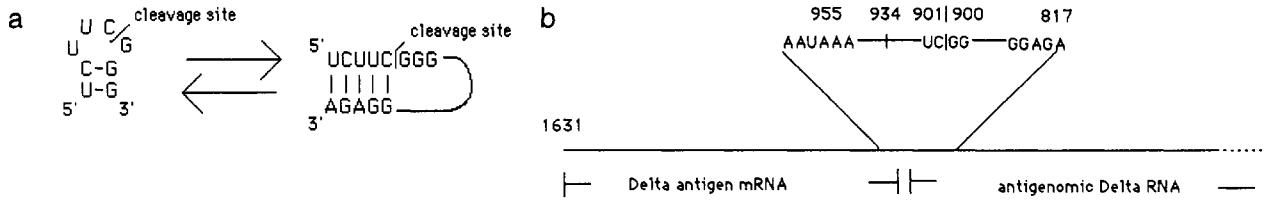

FIGURE 6: (a) Interconversion between the region 1 pairing with GGAGA and the formation of a tetraloop encompassing region 1 and the cleavage site. (b) Significant regions in the HDV and viral genome concerned with replication of the virus and transcription of the mRNA for the delta antigen. The initiation site of the mRNA and its polyadenylation site indicate that it is transcribed from approximately half the circular RNA genome.

interactive bonds (Zaug et al., 1988). Such interactions are necessary for catalytic RNAs to form stable secondary and tertiary interactions which maintain a catalytically competent structure. Our results indicate that the presence of the GGAGA sequence at the 3′ terminus of antigenomic HDV RNAs helps to maintain the catalytic core so that self-cleavage occurred even in 20 M formamide. This element introduces unusual stability to the RNA molecule, so that, under denaturing conditions which ordinarily melt RNA structures, these delta RNAs are now resistant to the mechanism of denaturation. This resistance to formamide was destroyed by mutagenesis of two of the bases within the sequence. The precise nature of this interaction remains unclear. Throughout the HDV RNAs, as well as in other self-cleaving or self-splicing RNAs such as the group I introns, there are extended helical regions which are required for the formation of an active structure. Why should this particular interaction, consisting of a five-nucleotide purine-rich region paired with a five-nucleotide pyrimidine-rich region, confer such unusual stability? The calculated free energy of the putative helix is only −3.1 kcal/mol. However, Bevilacqua and Turner (1991) have recently measured the themodynamic stability of the interaction of polypyrimidines such as CUCU with GGAGAA polypurine molecules. These measurements were performed on both the duplex formation and binding of CUCU to the ribozyme L-21 *Sca*I, which contains the sequence GGAGGA. They showed that the calculated $\Delta G°$ at 15 °C in 50 mM magnesium for the tetramer CUCU bound to GGAGGA is 1000 times weaker than that for the CUCU oligomer bound to the L-21 *Sca*I form of the *Tetrahymena* ribozyme. Single substitutions of deoxynucleotides in the CUCU demonstrated that putative tertiary interactions between the ribozyme and the middle two 2′OH groups on the CUCU oligonucleotide contributed approximately 1 kcal/mol of favorable free energy for each sugar. These data were consistent with gel mobility studies carried out by Pyle and Cech (1991), which showed that an unusual stability existed between the CUCU molecule and L-21 *Sca*I enzyme at 42 °C at 5 mM magnesium. The extraordinary stability that is apparent in the delta self-cleaving RNA, and which allows it to cleave in 20 M formamide, is presumed to reside, in part, in the pairing of GGAGA sequence with available polypyrimidine tracts. However, the stability of the antigenomic delta RNA is not derived from sequence alone, but rather the correct juxtaposition of sequence elements in a particular structural orientation that confers stability in formamide. This is supported by our observation that, when the GGAGA sequence is located near the 5′ end of the molecule, rather than at the 3′ end, the rapid cleavage activity remains, but not the stability in 20 M formamide, although cleavage in the unusually high 15 M formamide occurs. Complete removal of the GGAGA element (Ag96-v) leaves an RNA with reasonably efficient cleavage but with no stability even in 2.5 M formamide (Figure 3).

Our functional studies on the genomic RNA, using a modification-interference approach to determine the nucleotides essential for the self-cleavage reaction (Belinsky and Dinter-Gottlieb, submitted), indicate that these nucleotides are clustered in the regions analogous to stems II and III (Belinsky & Dinter-Gottlieb, 1991; Smith & Dinter-Gottlieb, 1991; Figure 5a). Furthermore, stem IV in the antigenomic RNA could be truncated to leave only seven contiguous base pairs below the GG bulge and still retained the ability to cleave in 18 M formamide (Prasad and Dinter-Gottlieb, in preparation). These studies, in addition to more direct physical analysis of the RNA, will allow us to pinpoint the significant regions of structural stability.

Significance of the GGAGA Interaction in Hepatitis Delta Replication. The structure of RNA is intimately related to its varied functions. Research over the past 10 years has begun to elucidate the nature of the catalytic cores of some of the RNA enzymes. Even when the minimal catalytic structures have been defined, however, the subtle interplay between secondary, tertiary, and alternative interactions is not obvious. The distinct interactions seen in two different tetraloop structures (Cheong et al., 1990; Heus & Pardee, 1990) were unexpected and novel.

Hepatitis delta virus may use self-cleavage to generate progeny virus molecules during its replication cycle (Taylor & Mason, 1987). We have defined a region between 80 and 84 nucleotides downstream of the cleavage site which can interact with the cleavage site to ensure efficient cleavage, even under nonphysiological, denaturing conditions. This may be useful to ensure the correct processing of the HDV in living cells. Presumably, the cleavage event cannot occur until the polymerase copying HDV has passed this point. The presence of the polymerase might, in fact, prevent correct interactions from forming, and thus prevent cleavage. Multimeric RNAs detected in vivo might be failure products of the self-cleavage reaction.

However, the self-cleavage to produce antigenomic viral RNAs must be balanced with the role of the antigenomic RNA as the mRNA for the delta antigen. Approximately 25 times more antigenomic delta RNA is detected than mRNA. The 5′ end of the mRNA has been mapped at nucleotide 1631, the adenylation signal at 954, and the adenylation site at 934 (Hsieh et al., 1990). The cleavage site is between 900 and 901 (Kuo et al., 1988), and the GGAGA, between 820 and 817 (Figure 6b). Self-cleavage and polyadenylation must be mutually exclusive events; otherwise all of the antigenomic RNAs would be polyadenylated, since the signal and site for polyadenylation are present in each copy of the antigenomic RNA. The mRNA represents only half of the full antigenomic circle, and the polyadenylation event would leave 34 nucleotides upstream of the self-cleavage site, sufficient for activity. Self-cleavage at the first possible site 34 nucleotides downstream of the polyadenylation site would "synchronize" subsequent self-cleavage events. However, once the antigenomic RNAs folded to allow further self-cleavage, polyadenylation of these full-length RNAs might be prevented either by the rapidity

of the self-cleavage or by the inaccessibility of the polyadenylation signal or site in the folded RNAs.

Physical and biological studies on the hepatitis delta RNA, both as virus and as ribozyme, will permit these biochemical interactions to be better understood.

REFERENCES

Belinsky, M., & G. Dinter-Gottlieb (1991) *Nucleic Acids Res. 19*, 559–564.

Behlen, L. S., Sampson, J. R., DiRenzo, A. B., & Uhlenbeck, O. C. (1990) *Biochemistry 29*, 2515–2523.

Bevilacqua, P. C., & Turner, D. H (1991) *Biochemistry 30*, 10632–10640.

Branch, A. D., & Robertson, H. D. (1984) *Science 223*, 450–455.

Chen, P.-J., Kalpana, G., Goldberg, J., Mason, W., Werner, B., Gerin, J., & Taylor, J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8774–8778.

Cheong, C., Varani, G., & Tinoco, I. (1990) *Nature 346*, 680–682.

Forster, A. C., & Symons, R. H. (1987) *Cell 49*, 211–220.

Heus, H. A., & Pardee, A. (1990) *Science 253*, 191–194.

Heus, H. A., Uhlenbeck, O. C., & Pardee, A. (1990) *Nucleic Acids Res. 18*, 1103-1108.

Hsieh, S.-Y., Chao, M., Coates, L., & Taylor, J. (1990) *J. Virol. 64*, 3192–3198.

Kos, A., Dikjema, R., Arnberg, A. C., van der Meide, P. H., & Schellekens, H. (1986) *Nature 323*, 558–560.

Kuo, M., Sharmeen, L., Dinter-Gottlieb, G., & Taylor, J. (1988) *J. Virol. 62*, 4439–4444.

Kuo, M. Y.-P., Chao, M., & Taylor, J. (1989) *J. Virol. 63*, 1945–1950.

Perrault, J.-P., Labuda, D., Usman, N., Yang, J.-H., & Cedergren, R. (1991) *Biochemistry 30*, 4020–4025.

Perrotta, A. T., & Been, M. D. (1990) *Nucleic Acids Res. 18*, 6821–6827.

Perrotta, A. T., & Been, M. D. (1991) *Nature 350*, 434–436.

Pyle, A. M., & Cech, T. R. (1991) *Nature 350*, 628–631.

Rosenstein, S., & Been, M. D. (1990) *Biochemistry 29*, 8011–8016.

Ruffner, D., & Uhlenbeck, O. C. (1990) *Nucleic Acids Res. 18*, 6025–6029.

Sharmeen, L., Kuo, M., Dinter-Gottlieb, G., & Taylor, J. (1988) *J. Virol. 62*, 2674–2679.

Smith, J. B., & Dinter-Gottlieb, G. (1991) *Nucleic Acids Res. 19*, 1285–1289.

Tang, C. K., & Draper, D. E. (1990) *Biochemistry 29*, 4434–4439.

Taylor, J., Mason, W., Summers, J., Goldberg, J., Aldrich, C., Coates, L., Gerin, J., & Gowans, E. (1987) *J. Virol. 61*, 2891–2895.

Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., D'Auberton-Carafa, Y., Uhlenbeck, O. C., Tinocco, I., Jr., Brody, E. N., & Gold, L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1364–1368.

Wu, H.-N., Lin, Y.-J., Lin, F.-P., Makino, S., Chang, M.-F., & Lai, M. C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1831–1835.

Zaug, A. J., Grosshans, C. A., & Cech, T. R. (1988) *Biochemistry 27*, 8924–8931.

Registry No. Formamide, 75-12-7.