Nucleic Acids Research, 1993, Vol. 21, No. 18 ***4193–4199***

Mutagenesis analysis of a hepatitis delta virus genomic ribozyme

Huey-Nan Wu, Jin-Yu Lee, Huey-Wen Huang, Yi-shuian Huang and Tung-Guang Hsueh
Institute of Molecular Biology, Academia Sinica, Nankang, Taipei, Taiwan 11529, China

Received June 29, 1993; Revised and Accepted August 10, 1993

ABSTRACT

We conducted extensive mutagenesis analysis on a hepatitis delta virus (HDV) genomic ribozyme to study the sequence specificity of certain region and to derive the secondary structure associated with the catalytic core. The results confirmed that the autocatalytic domain of HDV genomic RNA contained four basepairing regions as predicted in the 'pseudo-knot' model [Perrotta & Been (1990) *Nature* 350, 434–436]. The size and sequence of one of the basepairing regions, i. e. stem-and-loop, could be flexible. Helix 3 and the first basepair of helix 1 required specific sequence to retain self-cleavage activity. The structural requirement of helix 2 was less stringent than the other basepairing regions. Moreover, the size of helix 1 affected self-cleavage whereas the length of hinge could be variable even though the first three residues of hinge had stringent sequence requirement.

INTRODUCTION

The genome of the hepatitis delta virus (HDV) is a 1.7 kb single-stranded circular RNA (1–3). The HDV RNAs of both genomic and antigenomic senses have self-cleavage activity, and this property has been proposed to be involved in generating monomeric-sized molecules when the RNAs are synthesized through a rolling-circle mechanism (4). The self-cleavage reaction requires magnesium ions or other divalent cations (5), and generates a 2′, 3′-cyclic phosphoryl group and a 5′-hydroxyl group (5, 6). The boundaries of the catalytic domains of HDV genomic and antigenomic RNAs have been defined, and the sequences of both domains are highly homologous (7). However, the domains do not have the consensus sequences of other well-characterized catalytic RNAs such as the hammerhead ribozyme and the hairpin (paperclip) ribozyme (8, 9).

The HDV genomic and antigenomic catalytic domains have been proposed to fold into structures similar to each other. Several structural models have been proposed; all have in common a helix next to the cleaving point (helix 1) and a stem-and-loop near the 3′ terminus (10–12). Moreover, the structure rather than the sequence of these two base-pairing regions has been identified to be essential for self-cleavage activity (12–14). In the 'axe-head' model, the sequence upstream of cleaving point basepairs with the sequence downstream of the stem-and-loop (3′ tail), whereas no specific structure is associated with the region enclosed by helix 1 (loop 1) (11). In the 'pseudo-knot' model, the loop 1 contains a 3-basepair helix (helix 3) which may stack co-axially with a helix between loop 1 and 3′ tail (helix 2) (10, 14).

The purpose of this study is to further identify the structural and sequence requirements of the autocatalytic domain of HDV genomic RNA. We conducted mutagenesis analysis of certain regions of a HDV genomic ribozyme and assayed the effect of site-specific mutation on self-cleavage activity. The results indicated that HDV genomic ribozyme contained the basepairing structures as described in the 'pseudo-knot' model (10), and HDV genomic and antigenomic ribozymes were very similar.

MATERIALS AND METHODS

Plasmid construction

Construct pD(683-770)D4 was derived from construct pD(683-770) that encoded a sequence containing the autocatalytic domain of HDV genomic RNA (12). In pD(683-770)D4, the region corresponds to nucleotides 729–759 of HDV RNA was partially deleted and substituted by non-HDV sequence (12). A TA dinucleotide was inserted to the 5′ end of pD(683-770)D4 cDNA to generate a Kpn I site to facilitate further constructions and the recombinant was named as DG-1 (Fig. 1). The ribozyme coding sequence of construct DG-1 contained a Kpn I site, an Eco RI site and a Bam HI site (Fig. 1). Variants in this study were constructed by cassette mutagenesis or by PCR-mediated mutagenesis. Variants DG-6 and DG-6-1 ~ 14 were constructed by replacing the region between Kpn I and Eco RI sites of DG-1 cDNA with synthetic DNA fragments that contained degenerated sequences in nucleotides 690–695 and 719–724 together with a G to C substitution to destroy the Eco RI site and facilitate mutant selection (Fig. 1). Variants DG-6-61 ~ 63 were constructed by replacing the sequence between two Sma I sites in helix 1 of construct DG-6-6 with desired synthetic DNAs (Fig. 1). All the other mutants were constructed by PCR-mediated mutagenesis with appropriated synthetic oligonucleotides as primers from constructs DG-1 and DG-5. The sequence of each construct was confirmed by DNA sequencing (15).

Synthesis of RNA and determination of extent of cleavage

The internally ^{32}P-labeled RNA of each construct was synthesized by T7 RNA polymerase run-off transcription reaction in the presence of [α-^{32}P]-CTP with Bam HI-digested plasmid as template as described (5). Usually, the transcription reaction

(10 μl) was conducted at 37°C for 1 hr and then the reaction was stopped by an equal volume of 50 mM EDTA/7 M urea. The RNA that retained self-cleavage activity usually underwent cleavage when it was synthesized by T7 RNA polymerase because of the presence of 12 mM $MgCl_2$ in the reaction mixture. In order to obtain more full-length RNA, the transcription reactions of some variants were conducted at 4°C for 20 hr to reduce the extent cleavage. The full-length RNA and the products of the cleavage reaction were resolved on a polyacrylamide gel containing 7 M urea, and then the remaining full-length RNA and 3′ cleavage product were isolated from gel separately. The radioactivity of each RNA was determined by Cerenkov counting. The extent of cleavage of each RNA (E) was determined as [$cpm^{3'\ product}/(cpm^{3'\ product}+cpm^{full\text{-}length\ RNA}$)]. The full-length RNA used for the kinetics study was purified by the method previously described (5).

Determination of the half-life of the self-cleavage reaction

The kinetics of the self-cleavage reaction of the purified RNA was determined at 37°C or 50°C in the presence of 12 mM $MgCl_2$ and 40 mM Tris−HCl, pH 8.0. After denaturation at 90°C for 2 min and cooling down to room temperature, the ^{32}P-labeled full-length RNA was pre incubated at 37°C or 50°C in the presence of 40 mM Tris−HCl, pH 8.0, for at least 5 min. $MgCl_2$ solution that was pre warmed at the reaction temperature for at least 5 min with a final concentration of 12 mM, was then added to initiate the self-cleavage reaction. For each time point, 10 μl of the reaction mixture was removed and mixed with an equal volume of 50 mM EDTA/7 M urea to stop the self-cleavage reaction. The extent of cleavage of each RNA was determined without Mg^{2+} (E_0), at a certain time point in the presence of Mg^{2+} (E_t), and when the self-cleavage reaction had leveled-off in the presence of Mg^{2+} (E_f). We assumed that the Ef value might reflect the fraction of RNA that possessed active conformation for self-cleavage. For the RNAs of this study, the Ef of one RNA might be different from the Ef of the other RNA under a specific incubating condition e.g. 37°C with 12 mM Mg^{2+}. In addition, for a specific RNA, the Ef would vary form one incubation condition to the other. The fraction of the RNA cleaved at a certain time point (F_t) was normalized for only the active species according to the equation of $[(E_t-E_0)/(E_f-E_0)]$. The half-life of the self-cleavage reaction was obtained from the plot of log $(1-F_t)$ versus reaction time.

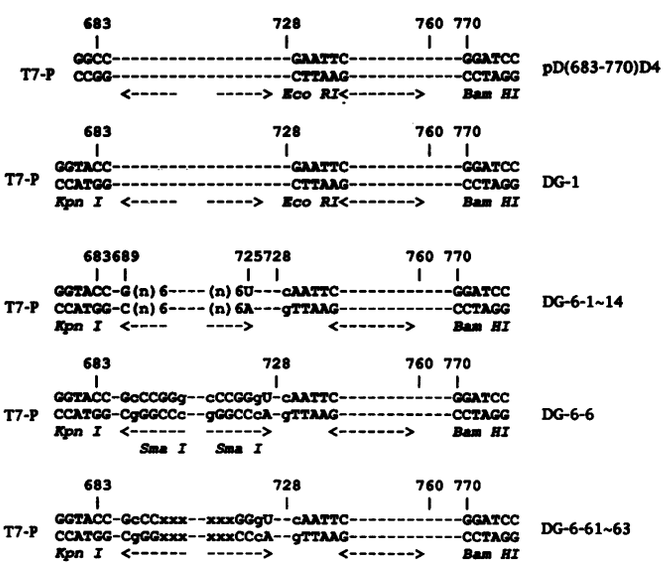

Figure 1. Diagrams to show the restriction sites of different constructs. Construct pD(683-770)D4 contains nucleotides 683−727 and 760−770 of HDV genomic RNA, and other residues are non-HDV sequences. Sequences in each construct that are different from the corresponding sequences of construct DG-1 are shown in small letters, in addition, '<- - - ->' represents the complementary sequences of nucleotides 690−695/719−724 and '<- - - - - - ->' represents the stem-and-loop region.

RESULTS AND DISCUSSION

We conducted mutagenesis analysis on a ribozyme (DG-1) derived from the autocatalytic domain of HDV genomic RNA, studied the effect of site specific mutation on self-cleavage reaction, and tried to derive the structural features and sequence requirements of the catalytic core. The self-cleavage activity of an RNA was justified according to several criteria. Since the T7 RNA polymerase *in vitro* transcription reaction is conducted in the presence of 12 mM $MgCl_2$, the RNA is able to undergo cleavage when it is synthesized. Therefore, an RNA is defined to possess high self-cleavage activity if it cleaves very efficiently when synthesized, i.e. the extent of cleavage is 70% to 90%. However, for certain RNA, the molecule may not fold properly and only a small amount of RNA cleaves when it is synthesized. The purified full-length RNA is heat denatured and then renatured before the kinetics analysis to provide the opportunity for the 'inactivated' RNA to refold itself, adopt 'active conformation' and undergoes self-cleavage reaction. An RNA possesses high self-cleavage activity if it has similar $t_{1/2}$ as that of the wild type ribozyme.

Ribozyme DG-1

The DG-1 RNA is a self-cleaving RNA (ribozyme) derived from a subfragment of HDV genomic RNA that has autocatalytic activity, i.e. pD(683-770) RNA, which contained nucleotides 683-770 of HDV RNA plus several nucleotides derived from cloning vector (12). The DG-1 RNA has the nucleotides 729−759 of pD(683-770) RNA replaced by a non-HDV sequence and as a result the DG-1 RNA contains 76 nucleotides downstream of cleaving point, i. e. between nucleotides 688 and 689 (Fig. 2B). Regardless of the differences in stem-and-loop and the residues around it, DG-1 and pD(683-770) RNAs seemed to have similar self-cleavage activities since both RNAs cleaved very well when they were synthesized at 37°C for 1 hr with 12 mM Mg^{2+}. In addition, the $t_{1/2}$'s of the self-cleavage reactions of both of the purified full-length RNAs were less than 1 min at 37°C and 50°C in the presence of 12 mM Mg^{2+} (Fig. 2B). Therefore, it is likely that DG-1 RNA retains the essential structure and sequence elements of the autocatalytic domain. Moreover, to obtain more full-length RNA for kinetics analysis, we synthesized RNA by a 4°C-transcription reaction. The full-length DG-1 RNA cleaved at similar rates regardless of the RNA that was synthesized by a transcription reaction conducted at 4°C or at 37°C. Nevertheless, the extent of cleavage was higher if the RNA was originally synthesized at lower temperature (data not shown). In this report, we conducted extensive mutagenesis analysis on ribozyme DG-1 instead of the autocatalytic domain of HDV genomic RNA by taking the advantage of the Eco RI site in the

cDNA of ribozyme DG-1 which would facilitate recombinant screening and cassette mutagenesis.

Hinge mutants/stem-and-loop mutants

The hinge region corresponds to the sequence connecting helix 1 and stem-and-loop which contains eight residues in ribozyme DG-1 (5'-GGGGAAUU-3') (Fig. 2A). For the hinge of ribozyme DG-1, the two U residues were derived from Eco RI site that were not present in HDV RNA, in addition, the fourth residue corresponds to nucleotide 729 is a C in wild type HDV RNA (Fig. 2A). To understand the sequence requirement of hinge, we altered this region of DG-1 RNA one nucleotide at a time and study the effect on self-cleavage reaction. As shown in Table 1, only the first three G residues, i. e. nucleotides 726–728, had stringent sequence requirement. The G726C and G728A substitutions affected self-cleavage reaction slightly, whereas the G726U, G727U, G728C and G728U substitutions decreased self-cleavage activity severely, and the remaining three single point substitutions abolished self-cleavage ability almost completely. In contrast, the individual sequence alterations from the fourth to the eighth residues of hinge of ribozyme DG-1 had almost no effect on self-cleavage reaction since each of the variants had the extent of cleavage of 70% to 90% when the RNA was synthesized by T7 RNA polymerase (Table 1 & data not shown). Thus, these five residues of hinge seemed to be dispensable for self-cleavage reaction. Furthermore, we deleted the 5'-GAAUU-3' of hinge from ribozyme DG-1 (variant DG-2). This DG-2 RNA self-cleaved and the $t_{1/2}$'s of the reaction were 5 min and 2 min at 37°C and 50°C, respectively (Fig. 2B, Table 1). Thus, the hinge region requires no more than three residues, in addition, coincidentally these residues have stringent sequence requirement and the preferred sequence is 5'-G/C-G-G/A-3' (Table 1). Kumar and co-workers (1993) have illustrated that the first several residues of hinge of the autocatalytic domain of HDV genomic RNA require specific sequence and the substitution of G726 to other sequence abolishes self-cleavage activity when the RNA is synthesized by T7 RNA polymerase at 42°C with 6 mM $MgCl_2$ (16). The differences in the acceptable sequence of the first residue of hinge (G/C or only G) may be due to the differences in the assay conditions used to determine the self-cleavage activity. Moreover, the corresponding region of the autocatalytic domain of HDV antigenomic RNA shares similar sequence specificity as that of ribozyme DG-1 regardless of the sequence differences in these two self-cleaving RNAs (13). All the studies suggest that the three G residues of hinge are involved in sequence specific interactions whereas extra nucleotides downstream of these G residues may not affect the overall conformation of the ribozyme. Recently, Belinsky *et al.* (17) suggested that A730 and A731 bulged out of stem-and-loop and the bulge was critical for self-cleavage activity. Mutant G3-61 had the AA bulge deleted and had the stem-and-loop extended to five continuous GC pairs (including G728/C763 and C729/G762) did not cleave. Our studies indicated that these two A's together with the other three residues of hinge of ribozyme DG-1 could be deleted simultaneously (variants DG-2 and DG-5,

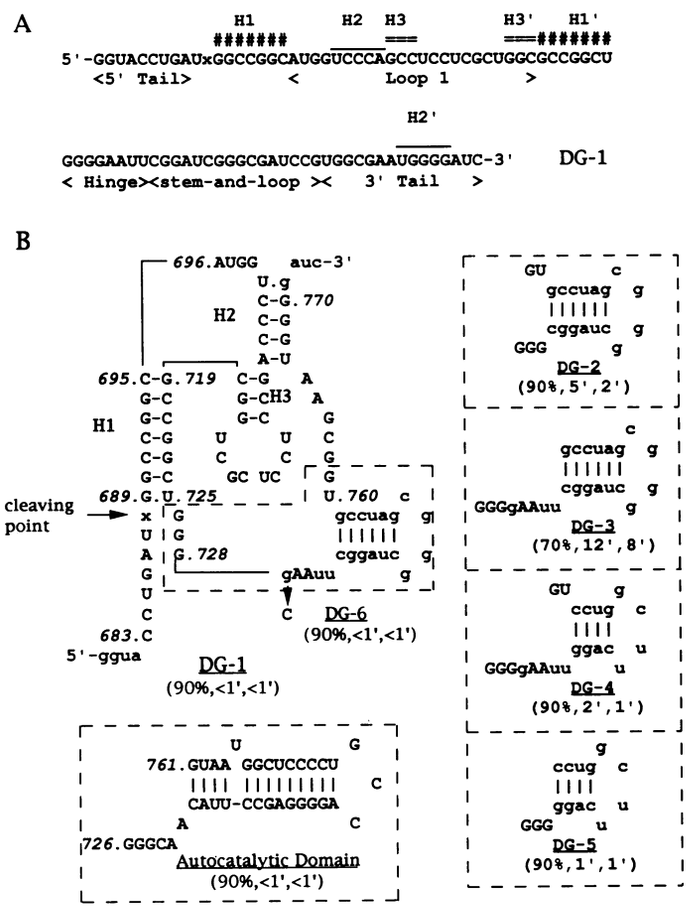

Figure 2. Sequence, proposed secondary structure and self-cleavage activity of different ribozyme. **(A)** Different sequence domains of ribozyme DG-1. H1/H1', H2/H2' and H3/H3' represent helix 1, helix 2 and helix 3, respectively. **(B)** The sequences of the autocatalytic domain of HDV genomic RNA and ribozymes DG-1~6. The sequence of hinge and stem-and-loop regions of each RNA is indicated in the dashed boxes. The sequences derived from HDV RNA are shown in capital letters and non-HDV sequences are shown in small letters. x is the cleaving point. H1, H2 and H3 correspond to the three helixes predicted in 'pseudo-knot' model (10). (x%, y', z') presents the percentage of full-length RNA cleaved when it is synthesized by T7 RNA polymerase at 37 °C for 1 hr, the $t_{1/2}$'s of the cleavage reaction of purified full-length RNA in the presence of 12 mM Mg^{2+} at 37°C and 50°C, respectively.

Table 1. Summary of the self-cleavage activity of hinge mutants of DG-1 RNA

Mutants	% cleaved	$t_{1/2}$ (37°C)	$t_{1/2}$ (50°C)
DG-1(W.T.)	90%	<1'	<1'
G726A	–	CS	CS
G726U	30%	50'	30'
G726C	90%	7'	6'
G727A, C	–	CS	CS
G727U	25%	55'	25'
G728A	80%	12'	2'
G728U	15%	75'	20'
G728C	30%	75'	30'
G729A	90%	1'	ND
G729U	75%	1'	<1'
G729C*1 (DG-6)	90%	<1'	<1'
A730G,U,C	75%	2'	<1'
A731G,U,C	75%	ND	ND
DG-2*2	90%	5'	2'

% cleaved indicates the extent of cleavage when the RNA was synthesized at 37°C for 1 hr. $t_{1/2}$: the half life of self-cleavage reaction of purified full-length RNA at 37°C and 50°C in the presence of 12 mM Mg^{2+}, 40 mM Tris–HCl, pH 8.0. ND: not determined. '–': the self-cleavage reaction could not be detected. CS: the RNA cleaved slowly and less than 5% of the RNA cleaved after more than 2 hr of incubation. *1: G729C is the wild type HDV sequence. *2: sequence of hinge of DG-2 RNA is GGG.

Fig. 2). Moreover, G762 and C763 were not able to form basepairs with the residues of hinge region to extend the stem-and-loop of DG-2 RNA, DG-5 RNA and other derivatives of ribozyme DG-1 (Fig. 2). The inconsistency of the role of these two A residues may due to the deletion of A730 and A731 by itself is not deleterious, instead, whether G762 and C763 are involved in stable basepairing affects self-cleavage reaction. This hypothesis is probably correct since mutants G3-84, G3-70 and G3-63 of Belinsky *et al.* that would have the basepairing of G728/C763 and C729/G762 destabilized by the bulge AA self-cleaved (17).

Unlike in the autocatalytic domain, residues U760 and G761 of ribozyme DG-1 are not involved in basepairing (Fig. 2B). We deleted these two residues from ribozyme DG-1 (variant DG-3, Fig. 2B) to test whether they are not important for self-cleavage reaction. The DG-3 RNA cleaved very well and the $t_{1/2}$'s of self-cleavage reaction were 12 min and 8 min at 37°C and 50°C, respectively, which were slightly slower than those of the DG-1 RNA. The result suggests that residues U760 and G761 can be deleted without severely altering the structure of catalytic core.

Since it has been shown that the structure rather than the sequence of stem-and-loop is critical for self-cleavage reaction, we tried to replace the corresponding region of ribozyme DG-1 with a superstable hairpin loop 5′-GGAC(UUCG)GUCC-3′ (variant DG-4, Fig. 2B) (18). As shown in Fig. 2B, the replacement had little effect on self-cleavage activity. Therefore, the stem-and-loop region might not interact with the other sequence and structural domains of ribozyme molecule.

To minimize the size of ribozyme, we constructed variant DG-5 that contained all the characteristics of DG-2, DG-3 and DG-4 RNAs, i.e. replaced nucleotides 729–761 of the autocatalytic domain with a hairpin loop (Fig. 2B). This DG-5 RNA that contained only 65 nucleotides downstream of cleaving point cleaved to similar extent as ribozyme DG-1 when it was synthesized, and it cleaved at similar rates in comparison to ribozyme DG-1 (Fig. 2). Therefore, the long self-complementary sequence can be replaced by a twelve-nucleotide hairpin loop without affecting the catalytic core, in addition, the hinge region requires only three nucleotides. Thus, the stem-and-loop region is a structural element and its length and sequence may not affect the overall structure of catalytic core.

Helix of nucleotides 690–695 and 719–724

The basepairing structure of nucleotides 690–695/719–724 has been shown to be critical for self-cleavage activity (12). To further investigate the sequence requirement of these six basepairs, we systematically changed the corresponding sequence on ribozyme DG-1 as well as on ribozyme DG-5.

Derivatives of ribozyme DG-1, i.e. variants DG-6 and DG-6-1 ~ 14, had a G to C substitution in hinge region to disrupt the Eco RI site and facilitate selection of recombinants (Fig. 1, 2B, 3A & 3C). The substitution would restore the wild type HDV sequence at nucleotide 729 and would not affect self-cleavage activity [variant DG-6 (Fig. 2B)]. Variants DG-6-1 ~ 13, all retained the basepairing structure intact and had one, two or four of the six basepairs substituted (Fig. 3A). These RNAs cleaved

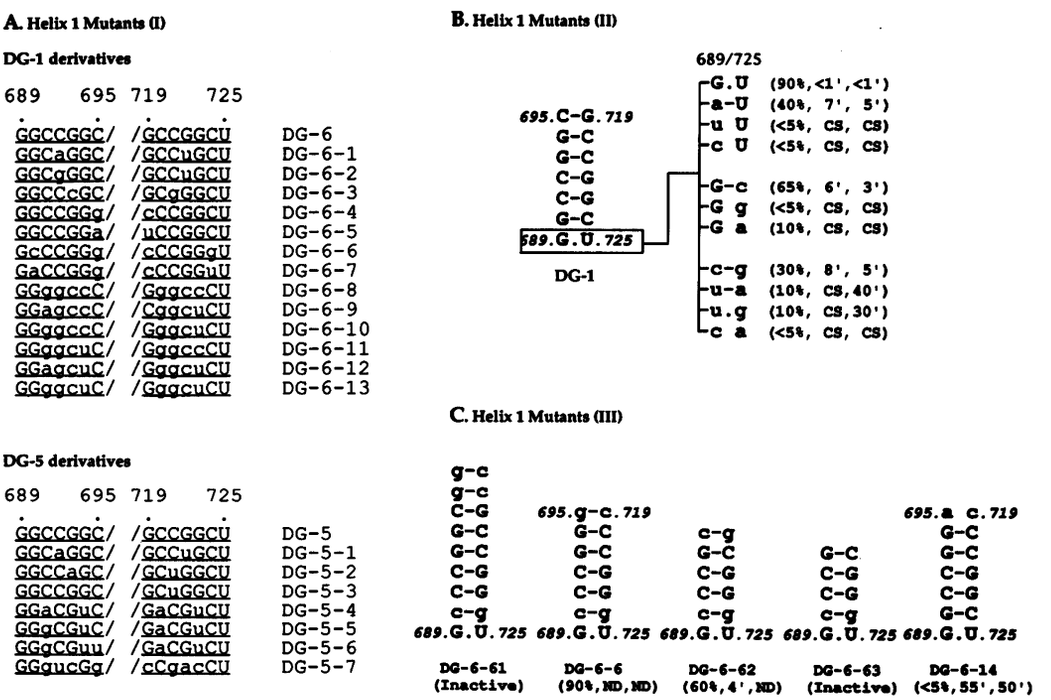

Figure 3. Different helix 1 mutants. **(A)** Variants derived from ribozymes DG-1 and DG-5 that contain sequence substitutions in nucleotides 690–695/719–724. **(B)** Variants of ribozyme DG-1 that have substitution(s) at nucleotides 689 and 725. **(C)** Variants of ribozyme DG-1 that contain helix 1 of different length. Only the sequences of the helix 1 region of each variant are shown, and the small letters indicate the substituted residues. The% of full-length RNA cleaved when it is synthesized and the $t_{1/2}$'s of self-cleavage reactions of full-length RNA are indicated. CS: less than 5% RNA cleaved after 2 hr of incubation, ND: not determined, Inactive: the RNA do not cleave.

to similar extent as those of the DG-1 and DG-6 RNAs when they were synthesized, which indicated that all of them were very active (data not shown). Thus, these 6 GC/CG pairs of nucleotides 690−695/719−724 can not only be flip flopped (variants DG-6-3, 4, 6 & 8) but also be substituted by AU/UA pairs (variants DG-6-1, 5, 7, 9 & 12) and by GU/UG wobble pairs (variants DG-6-2, 10, 11, 12 & 13) at almost every position, and at four basepairs simultaneously. In contrast, variants that had one or several basepairs mismatched either did not cleave or cleaved poorly (e. g. variant DG-6-14, Fig. 3C & data not shown).

Like in ribozyme DG-1, nucleotides 690−695/719−724 of ribozyme DG-5 could be flip flopped as well as be substituted by AU/UA/GU/UG at one or more than one basepairs without affecting self-cleavage activity [variants DG-5-1 ~ 7, (Fig. 3A) & data not shown]. Moreover, these six basepairs of ribozyme DG-5 can be substituted by the corresponding sequence of the autocatalytic domain of HDV antigenomic RNA (variant DG-5-7, Fig. 3A) (13).

The lack of sequence specificity of the 6 basepairs between nucleotides 690−695 and 719−724 confirms that the structure rather than the sequence of this helix is critical for self-cleavage activity. In addition, the result also implies that this helix may not have sequence specific tertiary interactions with other regions in ribozymes DG-1 and DG-5.

Basepairing of nucleotides 689 and 725

Variants that had either one or both of the G689 and U725 of ribozyme DG-1 substituted were constructed to investigate the sequence requirements, and the importance of basepairing interaction between nucleotides 689 and 725 on self-cleavage reaction. The G689A, U725C and G689C/U725G substitutions of ribozyme DG-1 affected the rate of self-cleaved reaction only slightly (Fig. 3B). In contrast, variants with other substitutions lost self-cleavage activity almost completely whether they retained the basepairing interaction (G689U/U725A and G689U/U725G) or not (G689U, G689C, U725G, U725A and G689C/U725A) (Fig. 3B). In addition, we conducted similar mutagenesis analysis on ribozyme DG-5-7 and found the effect of sequence alterations on self-cleavage reaction was in consistent with that of ribozyme DG-1 (data not shown). Thus, nucleotides 689 and 725 have stringent sequence requirement and the basepairing interaction between these two residues is critical for self-cleavage activity. It is possible that some sequence specific tertiary interactions are associated with these two residues or/and the functional groups of these nucleotides 689 and 725 are directly involved in cleavage reaction.

Helix 1 of different length

To investigate the length requirement of helix 1 (nucleotides 689−695/719−725), we constructed three mutants that contained a 9- (DG-6-61), a 6- (DG-6-62), or a 5- (DG-6-63) GC/CG basepair helix 1 and had nucleotides G689 and U725 of each variant unaltered (Fig. 3C). Variants DG-6-61 and DG-6-63 did not cleave whereas variant DG-6-62 self-cleaved very well. Thus the 7-basepair helix 1 could be deleted to 6-basepair long. However, the increase or the decrease of the size of helix 1 by two basepairs might alter the overall folding of the ribozyme and resulted in the loss of self-cleavage activity. Furthermore, variant DG-6-14, which had nucleotides 695/719 mismatched and as a result the sequence enclosed by helix 1 was elongated by 2 nucleotides, cleaved poorly (Fig. 3C). This result was consistent with the finding of Kumar et al., (19). Therefore, the disruption of the last basepair (nucleotides 695/719) of helix 1 would affect self-cleavage activity severely even though helix 1 could be 6- or 7-basepair long.

In summary, the size of the helix next to the cleaving point (helix 1) is restricted, a helix 1 of inappropriate size may affect the global structure of the autocatalytic domain and causes the loss of self-cleavage activity.

Intra-loop 1 helix (helix 3)

Loop 1 is the sequence enclosed by helix 1 and it corresponds to nucleotides 696−718 (Fig. 1). The loop 1 region contains several self-complementary sequences that may be involved in the formation of a 5-, a 4- or a 3-basepair helix (Fig. 4A). The study of sequence requirement of the entire loop 1 region of ribozyme DG-1 indicates that nucleotides 705−707 and 716−718 are invariable that strongly favors the presence of the 3-basepair helix (helix 3) as predicted in the 'pseudo-knot' model (10) (data from saturated single point mutagenesis analysis that is not shown here). The results also illustrate that the disruption of any of the basepair of helix 3 would abolish self-cleavage activity and each of the GC/CG pairs can not be replaced by GU/UG wobble pairs.

Furthermore, we studied the sequence requirement of helix 3 by constructing variants that retained the basepairing structure and determined the effect of sequence substitution on self-cleavage reaction. Mutants A−E had all six residues of helix 3 substituted. Mutant B that contained a UA pair, and mutants C, D and E, which contained one or two GU or UG wobble pair, lost self-cleavage activity completely (Fig. 4B). The results confirmed that each of the basepair of helix 3 could not be substituted by GU/UG wobble pairs, and indicated that the basepair of nucleotides 707/716 could not be UA either. Variants 705a718u and 706g717c are as active as the wild type ribozyme (Fig. 4B). Variant 706g707g716c717c and mutant A cleaved reasonably well

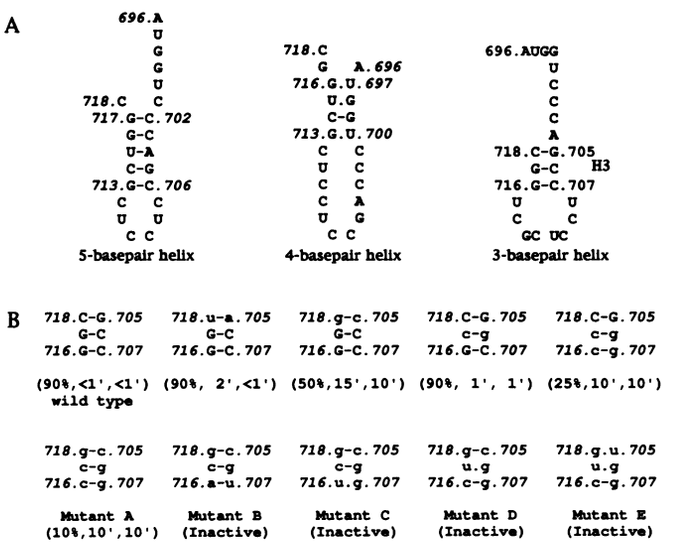

Figure 4. **(A)** Possible basepairing interactions within loop 1 of ribozyme DG-1. Only the nucleotides 696−718 of ribozyme DG-1 RNA are shown. **(B)** Variants to confirm the existence of the 3-basepair helix. Only the sequence of nucleotides 705−707/718−716 of each variant are shown and the remaining sequences of each RNA are the same sequence as DG-1 RNA.

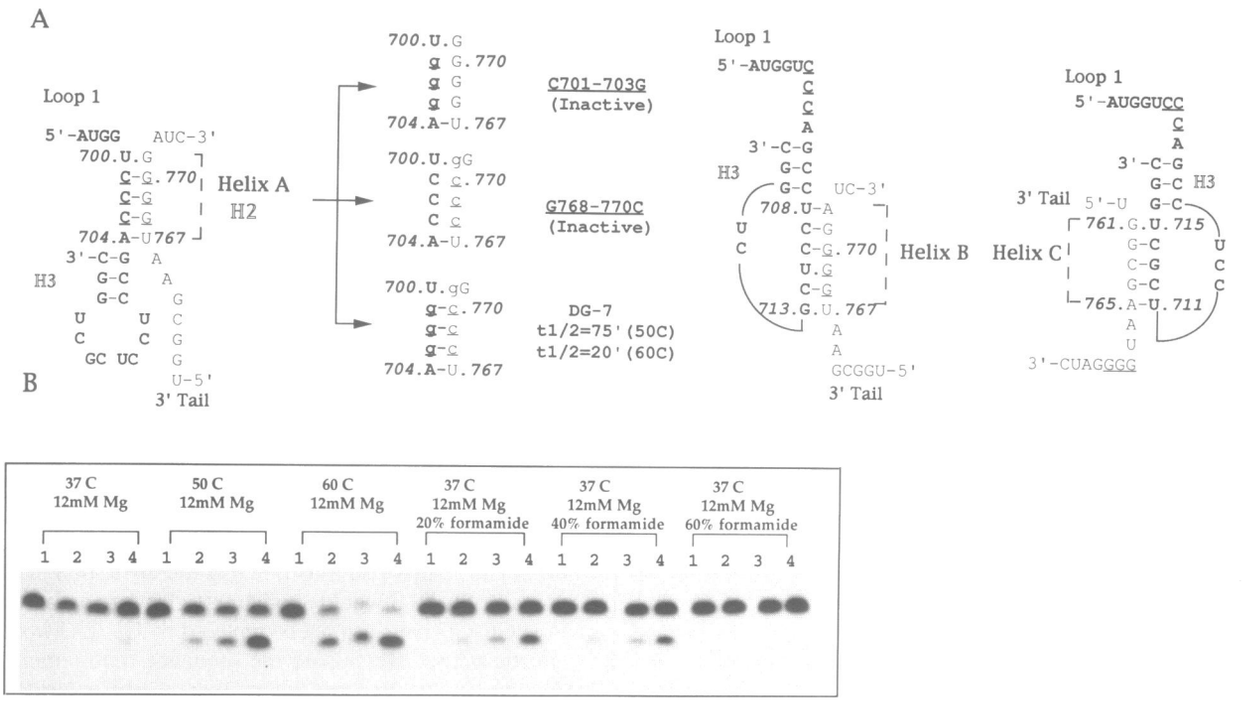

Figure 5. (A) Possible basepairing interactions between the loop 1 and 3′ tail regions of DG-1 RNA. Residues of loop 1 are shown in bold letters, in addition, residues correspond to nucleotides 701–703 and 768–770 are underlined. The sequence and structure of the remaining parts of the molecule are the same as for DG-1 RNA. (B). The self-cleavage reaction of purified full-length DG-7 RNA under different condition. Lanes 1, 2, 3 and 4 indicate the DG-7 RNA are incubated at indicated conditions for 0, 30, 60 and 120 min respectively. Upper bands are full-length RNAs, lower bands are 3′ cleavage products, and the 5′ cleavage products run off the 7% polyacrylamide gel containing 7 M urea.

only after the RNAs had been heat denatured and then renatured but they cleave at slightly slower rates (Fig. 4B). Variant 705c718g cleaved to higher extent when it is synthesized but it still cleaved at slower rate (Fig. 4B). The corresponding mutants of variants 705c718g and 706g707g716c717c of the autocatalytic domain of HDV genomic RNA constructed by Kumar and co-workers cleaved very well when they were synthesized by T7 RNA polymerase at 42°C with 6 mM $MgCl_2$ (19). The difference between the results of two laboratories may come from the difference in incubating temperature (42°C vs. 37°C). The results suggest that certain sequence substitutions in helix 3 may favor the formation of 'inactive conformations' and as a result, RNA cleaves poorly unless it has the chance to refold itself and adopts active conformation. In addition, higher incubating temperature may prevent the formation of alternative inactive conformations and facilitate self-cleavage reaction. Furthermore, sequence specificity of helix 3 suggested that certain sequence substitution may alter some tertiary interactions and decreases self-cleavage activity. Nevertheless, helix 3 of ribozyme DG-1 has sequence specificity that is like the corresponding helix in the autocatalytic domain of HDV antigenomic RNA (14).

Helix between loop 1 and 3′ tail (helix 2)

It is likely that a helix instead of several discontinuous basepairs is formed between loop 1 and 3′ tail to stabilize the overall folding of the ribozyme molecule since complementary sequences are present between these two regions. As shown in Fig. 5A, there are at least three kinds of basepairing interaction between loop 1 and 3′ tail that may stack with helix 3 co-axially. We studied the effect of the disruption of one or two basepairs of each of the hypothetical helix of ribozyme DG-1 on self-cleavage activity to investigate which of the helixes was critical for self-cleavage activity. However, the results did not provide information on which of the helixes was critical since all the variants we constructed self-cleaved (data not shown). It is possible that either none of the hypothetical helix is required for self-cleavage reaction or the structural requirement of the helix between loop 1 and 3′ tail is non-stringent.

We constructed multiple mutants of ribozyme DG-1 to detect the existence of the hypothetical helix A, B or C. They are variants C701-703G, G768-770C and the compensatory mutant variant DG-7, and the last two RNAs have a G residue inserted to near the 3′ end of the RNAs to restore the Bam HI restriction site (Fig. 5A). The mutations in these three variants would not affect helix C at all, whereas the mutations in variant C701-703G would disrupt helix A, the mutations in variant G768-770C would disrupt helix A as well as helix B, and variant DG-7 would have helix B disrupted and have the basepairing of helix A restored (Fig. 5A). Variants C701-703G and G768-770C did not cleave under any of the condition tested (Fig. 5A). Variant DG-7 cleaved poorly at 37°C in the presence of 12 mM Mg^{2+}, however the addition of 20% or 40% (v/v) formamide, or the elevation of the reaction temperature to 50°C or 60°C could enhance self-cleavage activity of DG-7 RNA (Fig. 5A & 5B). These results give direct evidence on the presence of the hypothetical helix A in ribozyme, which is the helix 2 of the 'pseudo-knot' model. Moreover, the studies of the corresponding mutants of ribozyme DG-5-7 gave the same conclusions. The requirement of the

maintenance of helix 2 is less stringent in comparison to that of helixes 1 and 3 since the self-cleavage activity will not be abolished unless three basepairs of helix 2 are mismatched simultaneously (Fig. 5A & unpublished data of this lab). This feature of helix 2 is very similar to the corresponding helix of the autocatalytic domain of HDV antigenomic RNA in which the basepair mismatch(es) in helix 2 does not affect self-cleavage reaction unless in the presence of denaturant (10).

CONCLUSION

These studies confirm that the catalytic domain of HDV genomic RNA contains four basepairing regions in which helix 3 as well as the first basepair of helix 1 may be involved in some tertiary interactions that are important for catalytic activity, whereas the remaining six basepairs of helix 1, stem-and-loop and helix 2 are likely the structural elements of the catalytic core. Moreover, the common features of these double stranded regions and the similarity in sequence specificity of hinge strongly suggest that the catalytic cores of HDV genomic and antigenomic RNAs share similar secondary and tertiary structures.

ACKNOWLEDGEMENTS

This work was supported by grant NSC-80-0203-B001-14 from National Science Council of the Republic of China.

REFERENCES

1. Kos, A., Dijkema, R., Arnberg, A. C., van der Meide, P. H., and Schelleken, H. (1986) *Nature (London)*, **323**, 558–560.
2. Makino, S., Chang, M. F., Shieh, C. K., Kamahora, T., Vannier, D. M., Govindarajan, S., and Lai, M. M. C. (1987) *Nature (London)*, **329**, 343–346.
3. 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. (1987) *Nature (London)*, **323**, 508–515.
4. Kuo, M. Y.-P., Sharmeen, L., Dinter-Gottieb, G., and Taylor, J. (1988) *J. Viol,*. **62**, 4439–4444.
5. 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.
6. Perrotta, A. T. and Been, M. D. (1992) *Biochemistry*, **31** 16–21.
7. Rosenstein, S. P. and Been M. D. (1991) *Nucleic Acids Res.*, **19**, 5409–5416.
8. Uhlenbeck, O. C. (1987) *Nature (London)*, **328**, 596–600.
9. Hampel, A., Tritz, R., Hicks, M., and Crutz, P. (1990) *Nucleic Acids Res.*, **18**, 299–304.
10. Perrotta, A. T. and Been, M. D. (1990) *Nature (London)*, **350**, 434–436.
11. Branch, A. D. and Robertson, H. D. (1991) *Proc. Natl. Acad. Sci., USA* **88**, 10163- 10167.
12. Wu, H. N., Wang, Y. J., Hung, C. F., Lee, H. J., and Lai, M. M. C. (1992) *J. Mol. Biol.*., **223**, 233–245.
13. Wu, H. N. and Huang, Z. S. (1992) *Nucleic Acids Res.* **20**, 5937–5941.
14. Been, M. D., Perrotta, A. T. and Rosenstein, S. P. (1992) *Biochemistry* **31**, 11843–11852.
15. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci., *USA* **74**, 1074–1078.
16. Kumar, P. K. R., Suh, Y. -H., Miyashiro, H., Nishikawa, F., Kawakami, J., Taira, K., & Nishikawa, S. (1992) *Nucleic Acids Res.* **20**, 3919–3924.
17. Belinsky, M. G., Britton, E. and Dinter-Gottlieb, G. (1993) *FASEB J.*, **7**, 130–136.
18. Cheong, C., Varani, G. and Tinoco, I. Jr. (1990) *Nature* (London), **346**, 680–682.
19. Kumar, P. K. R., Suh, Y. A., Taira, K. and Nishikawa, S. (1993) *FASEB J.*, **7**, 124–129.