Chemical Probing Studies of Variants of the Genomic Hepatitis Delta Virus Ribozyme by Primer Extension Analysis†

P. K. R. Kumar, Kazunari Taira, and Satoshi Nishikawa*

National Institute of Bioscience and Human Technology, Agency of Industrial Science and Technology, MITI, Tsukuba Science City, Ibaraki 305, Japan

Received June 7, 1993; Revised Manuscript Received November 9, 1993®

ABSTRACT: We have investigated in detail the higher order structure of the genomic hepatitis delta virus (HDV) ribozyme using various base-specific chemical probes under native, semi-denaturing, and denaturing conditions. The bases of the HDV ribozyme were probed by treatment with dimethyl sulfate [which reacts with A (at N1) and C (at N3)] and a carbodiimide [which reacts with U (at N3) and G (at N1)]. In addition, for probing G residues (at N7), RNA samples were treated with $NaBH_4$ and aniline after modification by treatment with dimethyl sulfate. The sites of modified positions were identified by primer extension analysis with reverse transcriptase. In general, our results are consistent with the proposed pseudoknot model of secondary structure, a model that is based on data from ribonucleolytic cleavage experiments. Our results provide clues to the identification of interacting bases in the HDV ribozyme. Furthermore, using this method we identified local conformational changes in several stem variants.

Human hepatitis delta virus (HDV) is known to cause fulminant hepatitis or chronic hepatitis in the presence of cirrhosis due to hepatitis B virus (HBV) or as a result of superinfection of chronic carriers of HBV (Govindarajan *et al.*, 1984). The genome is a single-stranded circular RNA of about 1700 nucleotides. The RNA appears to replicate through a rolling circle RNA-to-RNA pathway, as observed in the case of certain plant-pathogenic RNA viruses. As a result of the replication, both genomic and antigenomic HDV RNAs are formed. The genomic and antigenomic HDV RNAs have self-cleavage activity (ribozyme activity) *in vitro* (Sharmeen *et al.*, 1988). As observed with other self-cleaving RNAs, the self-cleavage activity of the HDV ribozyme requires divalent cations and results in products with a 2′,3′-cyclic phosphate and a 5′-OH group (Kuo *et al.*, 1988; Sharmeen *et al.*, 1988). However, the primary sequence of the HDV ribozyme does not resemble those of other known types of ribozyme. Thus, the HDV ribozyme represents a unique type of self-cleaving RNA (Wu *et al.*, 1989; Taira & Nishikawa, 1992).

Recently, secondary-structure models have been proposed for both genomic and antigenomic HDV ribozymes (Perrotta & Been, 1991; Branch & Robertson, 1992; Wu *et al.*, 1992). In order to identify the important regions of the HDV ribozyme and to define the minimal molecule required for the self-cleavage activity, several bases were deleted, either internally or from both termini, to generate several shorter but still active ribozymes (Suh *et al.*, 1992). Perrotta & Been (1990) showed that a minimal self-cleaving molecule consists of 85 nucleotides (nt), namely, 1 nt at the 5′-end and 84 nt at the 3′-end of the molecule are sufficient for efficient cleavage of the HDV ribozyme. Therefore, if we assume that the 5′-end nucleotide is part of the "substrate", the active site of the HDV ribozyme must be present in the structure that is formed by folding the sequence 3′ to the cleavage site. Secondary-structure models have been generated using 3′-cleaved products of genomic and antigenomic HDV ribozymes (Rosenstein & Been, 1991; Perrotta & Been, 1991). In an attempt to understand the role of bases and to define the catalytic core structure of the genomic HDV ribozyme, we constructed several point and random variants in single-stranded regions (Kumar *et al.*, 1992; Suh *et al.*, 1993; Kawakami *et al.*, 1993). Recently, we constructed several variants of the HDV ribozyme with changes in the stem regions to examine the proposed secondary-structure models (Kumar *et al.*, 1993). The results of our mutational analysis are consistent with the proposed pseudoknot secondary structure for the genomic HDV ribozyme (Figure 1). Further clarification of the structure of the genomic HDV ribozyme depends, therefore, on the identification of the bases that interact to produce the tertiary structure.

Peattie and Gilbert (1980) described a method for identification of the sites of base modification by use of site-specific alkylating agents in which the positions are read directly from the sequencing gels. These chemical probes have been widely used since then to test structural models for RNA, to monitor conformational changes in RNA, and to localize ligand-binding sites on RNA (Inoue & Cech, 1985; Moazed & Noller, 1986; Moazed *et al.*, 1986; Stern *et al.*, 1986; Mougel *et al.*, 1987). Results obtained with the chemical probes support conclusions derived from crystallographic studies (Peattie & Gilbert, 1980). Thus, it appears that chemical probing is an experimental approach that allows the prediction of certain structural constraints that can help in the characterization of RNA structures for which no X-ray data are available. Therefore, chemical probes provide an image of the structure of RNAs in solution.

The sites of modification of nucleotides can be identified either by direct cleavage of the phosphodiester bonds of 3′-end-labeled RNAs or by mapping with reverse transcriptase of unlabeled RNAs. In the present study, we used mapping with reverse transcriptase to identify the positions in the genomic HDV ribozyme that had been modified by use of several structure-specific chemical probes. The four bases were monitored at their Watson–Crick pairing positions by reaction with dimethyl sulfate (DMS) at A (at N1) and C (at N3) and with a carbodiimide at U (at N3) and G (at N1). In addition, to determine the interactions of G (at N7), DMS-treated samples were treated further with $NaBH_4$ and aniline.

† P.K.R.K. is grateful to the Science and Technology Agency of Japan for the award of a fellowship.

* Author to whom correspondence should be addressed (Telephone, 0298-54-6085; FAX, 0298-54-6095).

® Abstract published in *Advance ACS Abstracts,* January 1, 1994.

0006-2960/94/0433-583$04.50/0

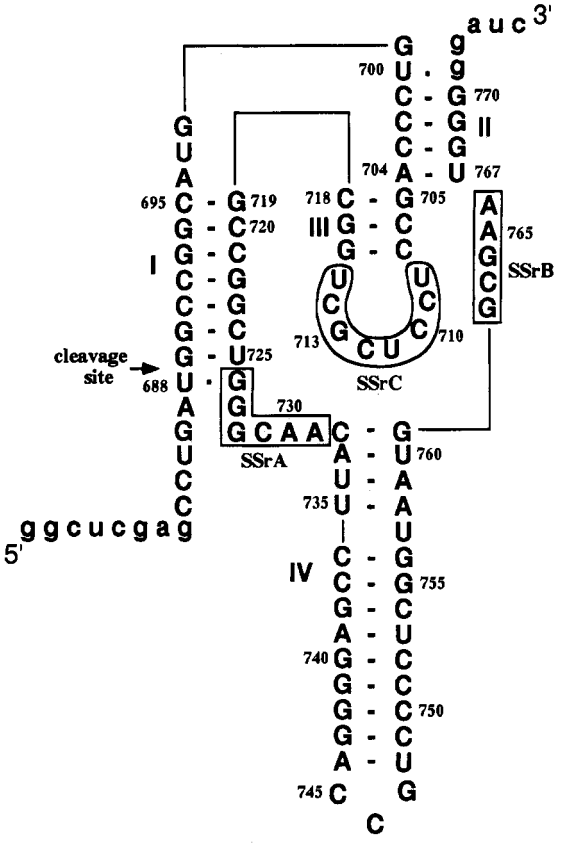

FIGURE 1: Proposed pseudoknot-like secondary structure for the genomic HDV ribozyme. The genomic sequence of HDV is numbered according to Makino *et al.* (1987). Small letters indicate vector-derived sequences. Important single-stranded regions, SSrA, SSrB, and SSrC (Kumar *et al.*, 1992; Suh *et al.*, 1993; Kawakami *et al.*, 1993), are enclosed by a solid line.

The stability of each structure was estimated by modifying the HDV ribozyme under native, semi-denaturing, and denaturing conditions. Using reverse transcriptase and synthetic deoxyoligonucleotide primers (designated 1 and 2) that were complementary to RNA, we were able to probe most of the nucleotides of HDV88. The results of this study provide data that permit a detailed and rigorous examination of the validity of the accepted secondary-structure model of the genomic HDV ribozyme, as well as clues to the nature of tertiary interactions.

MATERIALS AND METHODS

Chemicals and Enzymes. 1-Cyclohexyl-3-[2-(*N*-methylmorpholino)ethyl]carbodiimide *p*-toluenesulfonate (CMCT) and aniline were purchased from Wako Pure Chemical Industries (Osaka, Japan). Dimethyl sulfate was from Nacalai Tesque Inc. (Kyoto, Japan). T4 polynucleotide kinase was from Nippon Gene (Toyama, Japan), and avian myeloblatosis virus (AMV) reverse transcriptase was from Life Sciences Inc. (St. Petersburg, FL). Other chemicals of higher purity were purchased from Wako Pure Chemical Industries.

Plasmid DNA. The plasmid, pUHD88 (HDV88), was used in the present study, and the construction of this vector has been reported elsewhere (Suh *et al.*, 1992). It contains the genomic HDV sequence from nucleotides (nt) 683 to 770, which has high self-cleavage activity during transcription relative to our various other constructs. Construction of several stem variants of HDV88 that were used in the present study, namely, the U767A, C695G:G719C, and C718G variants, has been reported previously (Kumar *et al.*, 1993). All experiments were carried out using *Escherichia coli* MV1184 as the host. The plasmid DNA for transcription of HDV88 was prepared from an overnight culture.

Transcription in Vitro. The plasmid DNA was isolated by the alkaline lysis method, with subsequent treatment with 20% PEG (M_r, 6000) and CsCl density gradient ultracentrifugation. The DNA was linearized with *Bam*HI and used for *in vitro* transcription. To prepare larger quantities of RNA, transcription was carried out without [α-^{32}P]CTP for 4 h at 42 °C using an Ambion MEGAscript *in vitro* transcription kit (Ambion Inc., Austin, TX). Under these conditions, the HDV88 ribozyme undergoes self-cleavage to nearly 100%, mainly resulting in the 3′-cleaved product (689–770 plus five vector-derived nucleotides, 87 nt) of HDV88. The stem-variant ribozymes U767A and C695G:G719C have about 90% and 80% self-cleavage activity relative to that of the HDV88 ribozyme, respectively (Kumar *et al.*, 1993). Therefore, the product observed in the case of both the U767A and the C695G:G719C variants after transcription was mainly the 3′-product, since these were active variants. By contrast, the C718G variant (inactive variant) was isolated as the full-length transcript (101 nt). After each reaction, an equal volume of stop solution (9 M urea, 50 mM EDTA, 0.1% bromophenol blue, and 0.1% xylene cyanol FF) was added, and then heat-denatured RNA was fractionated by electrophoresis on an 8% (w/v) gel that contained 7 M urea. The 3′-product was located by UV shadowing, which was followed by excision and extraction of the product from the gel and recovery of the RNA as described earlier (Kumar *et al*, 1993). The isolated RNA (60 pmol) was first danatured by heating in deionized water for 4 min at 92 °C. One-tenth volume of buffer [50 mM NaCl and 25 mM Tris-HCl (pH 7.5)] was added to the ribozyme solution to renature the RNA. The renatured RNA (5 pmol/μL) was used to probe the RNA under different conditions and was also used for the control ladders.

Chemical Modification. The general procedures were essentially those described by Krol and Carbon (1989). For each modification reaction, 5 pmol of unlabeled 3′-cleaved product RNA from the HDV88, the U767A, or the C695G:G719C variant was used. In the case of the C719G variant, the full-length unlabeled product of transcription was used for the modification studies. Controls were treated in parallel with omission of the chemical reagent.

Modification with Dimethyl Sulfate. For probing under native conditions, modification was carried out at 30 °C in 200 μL of buffer I [200 mM HEPES (pH 8), 10 mM $MgCl_2$, and 50 mM KCl] and 1 μL of DMS for 5 min. For probing under semi-denaturing conditions, buffer I was replaced by buffer II [200 mM HEPES (pH 8) and 1 mM EDTA], but all other conditions remained the same. For reactions under totally denaturing conditions, 300 μL of buffer II was used, and reaction mixtures were incubated at 90 °C for 1 min. To identify the methylation of G residues at the N7 position, cleavage at modified guanosine residues was achieved by treatment with aniline, as described by Krol and Carbon (1989).

Modification with the Carbodiimide. For modification under native conditions, we combined 4 μL of the 3′-cleaved product of the HDV88 ribozyme (total, 5 pmol), 50 μL of buffer III [50 mM sodium borate (pH 8), 10 mM $MgCl_2$, and 50 mM KCl], 2 μL of 1 M $MgCl_2$, 10 μL of 1 M KCl, 84 μL of H_2O, and 50 μL of a freshly prepared solution of CMCT in H_2O (42 mg/mL). The reaction products were incubated at 30 °C for 10 min. For probing under semi-denaturing conditions, we combined 4 μL of the 3′-cleaved product of the

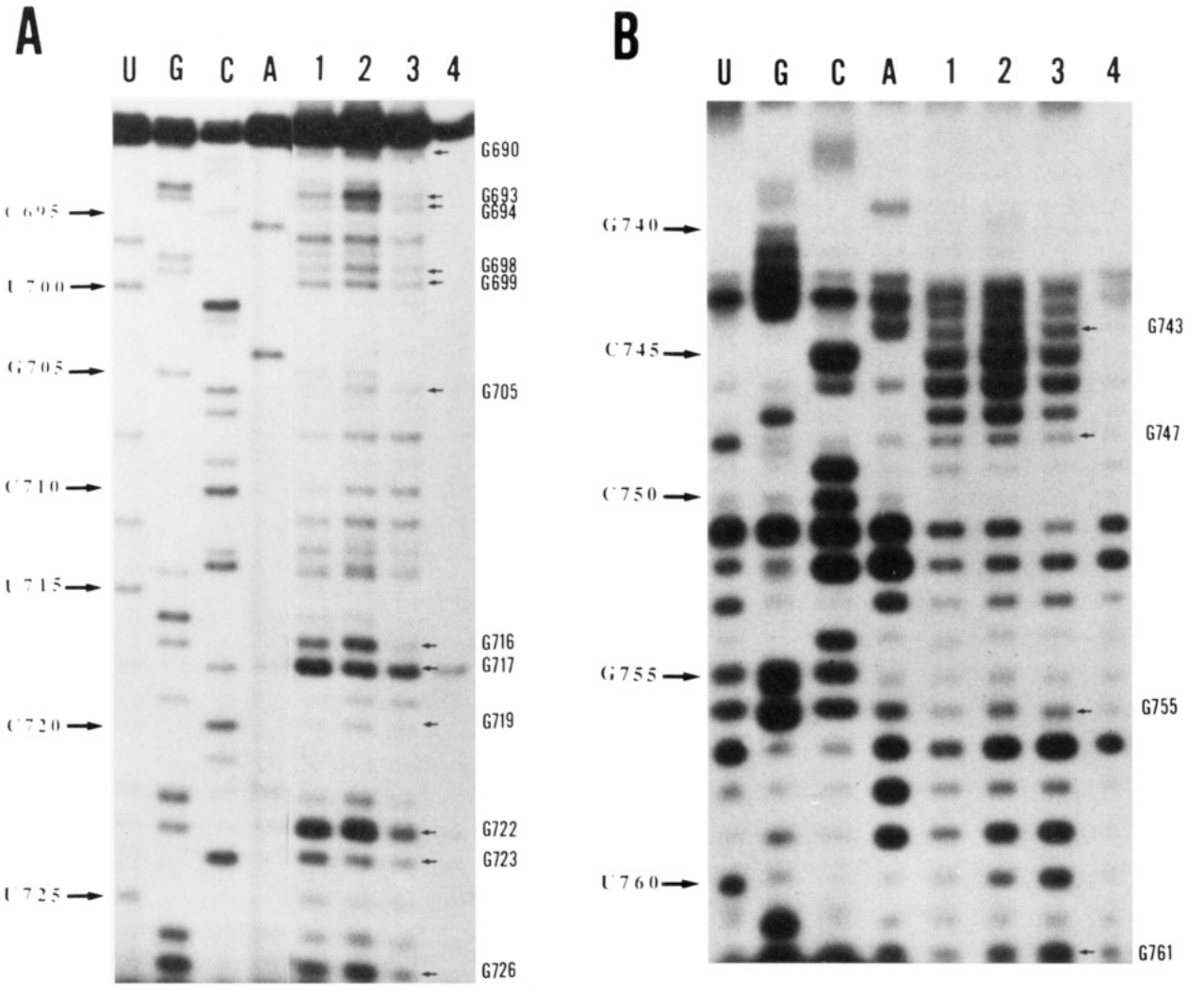

FIGURE 2: Autoradiography showing sites of chemical modification by dimethyl sulfate, with subsequent treatment by $NaBH_4$ and aniline of the 3′-product of HDV88. Sites were identified after hybridization with primer 2 (A) or primer 1 (B) and reverse transcription. Bold arrows indicate the numbering of bases on the RNA molecule. Small arrows indicate the modifications of G residues (at N7). U, G, C, and A indicate control ladders that were generated by dideoxynucleotide sequencing of unmodified, cleaved HDV88 RNA. Lane 1, probing under native conditions; lane 2, probing under semi-denaturing conditions; lane 3, probing under denaturing conditions; lane 4, RNA treated in the same way as the sample in lane 1, except that dimethyl sulfate was not included in the reaction mixture.

HDV88 ribozyme (total, 5 pmol) 50 μL of buffer IV [50 mM sodium borate (pH 8) and 1 mM EDTA], 95 μL of H_2O, 1 μL of 200 mM EDTA, and 50 μL of a solution of CMCT (42 mg/mL) and incubated the mixture for 5 min at 30 °C. For probing under completely denaturing conditions, the reaction components were similar to those for semi-denaturing conditions, but mixtures were incubated at 90 °C for 1 min.

After the alkylation, all reactions were stopped by precipitation with 100 μL of 0.3 M sodium acetate, 5 μg of carrier tRNA, and 3 vol of ethanol at −70 °C for 15 min. The pellet was washed with 70% ethanol and dried.

Primer Extension. Two primers were used to identify the sites of modification by each type of reaction in HDV88 and other stem variants (C695G:G719C, U767A, and C718G). Primer 1 (5′-GATCCCCCATTC-3′) and primer 2 (5′-CCTCGGAATGTT-3′) are complementary to residues 764–775 and 730–741, respectively. These primers were synthesized by the phosphoramide method on an automated DNA synthesizer (Model 380A, Applied Biosystems, Foster City, CA). Both primers were 5′-end-labeled in a 20-μL reaction mixture that contained 20 pmol of either primer 1 or primer 2, 20 pmol of [γ-^{32}P]ATP (3000 Ci/mmol), 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, and 10 units of T4 polynucleotide kinase. The reaction mixture was incubated for 1 h at 37 °C and terminated by the addition of EDTA to 25 mM. The labeled primers were purified by electrophoresis on a 20% polyacrylamide gel under denaturing conditions and extracted as described elsewhere (Kumar *et al.*, 1993). The hybridization mixture, containing either modified or unmodified RNA (1.0 pmol) and primer 1 (2.5×10^4 cpm) in 4 μL of water, was incubated at 90 °C for 1 min and slowly cooled to room temperature. For hybridization with primer 2, 1.0 pmol of either modified or unmodified RNA was mixed with primer 2 (1.0×10^5 cpm) in 4 μL of H_2O, incubated at 90 °C for 2 min, and then immediately placed on ice for 15 s. Then the reaction mixtures were returned to 90 °C for 2 min and the above procedure was repeated twice. Once this process was completed, the tubes were gradually cooled to room temperature. This repeated heating and cooling was essential for the binding of primer 2 since its complementary sequence is denatured with difficulty at higher temperatures. To 2 μL of the hybridizing mixture we then added 1.5 μL of dNTP mix that contained 750 μM of each dNTP and 1.5 μL of buffer solution [100 mM Tris-HCl (pH 8.3), 150 mM KCl, 20 mM dithiothreitol, and 34 mM $MgCl_2$] that contained 1.1 units of AMV reverse transcriptase. Extension was performed at 37 °C for 30 min. The reactions were stopped by the

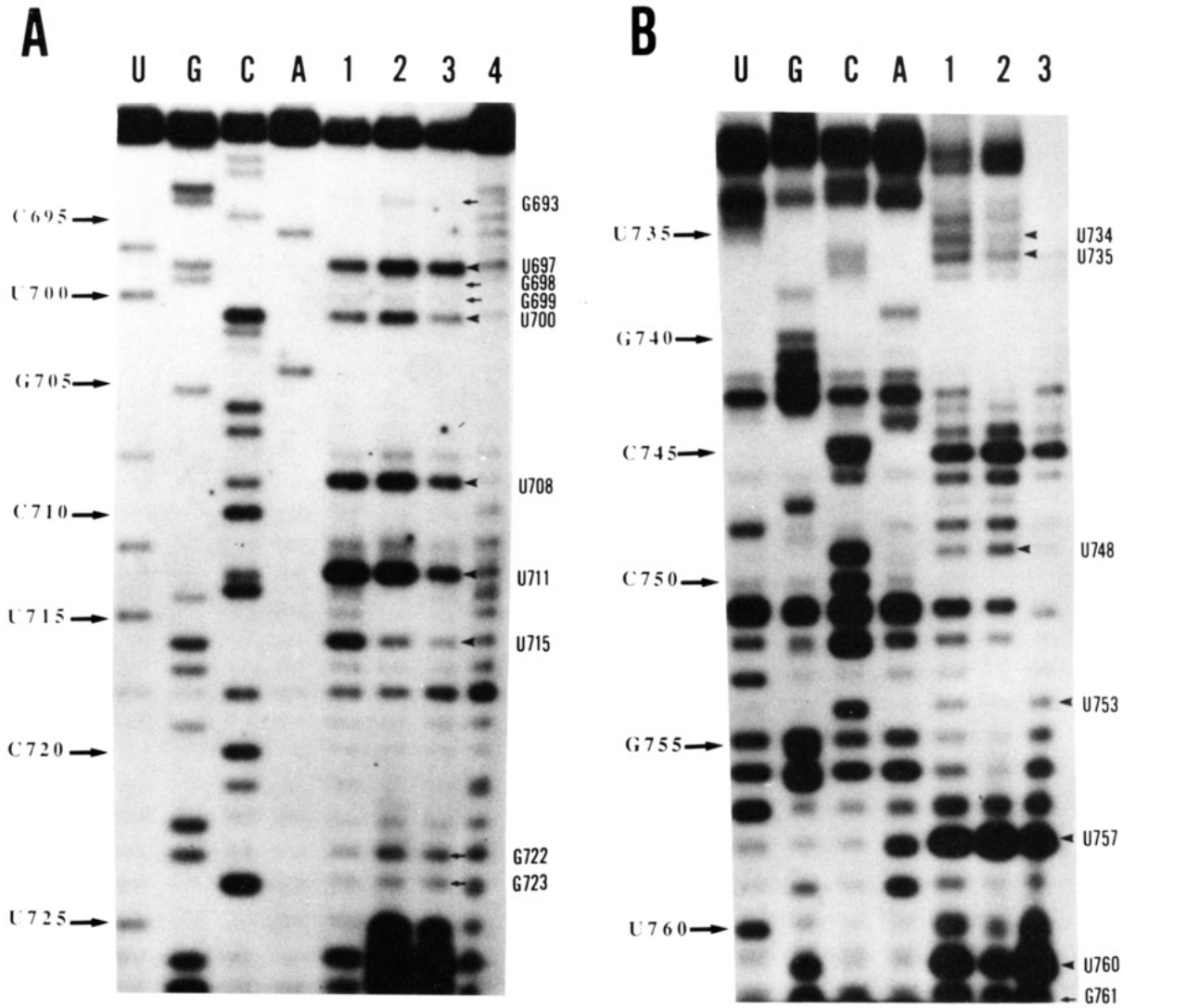

FIGURE 3: Autoradiography showing sites of chemical modification by CMCT of the 3′-product of HDV88. Sites were identified after hybridization with primer 2 (A) or primer 1 (B) and reverse transcription. Bold arrows indicate the numbering of bases on the RNA molecule. Arrowheads indicate the modifications of G residues (at N1). Small arrows indicate the modifications of U residues (at N3). U, G, C, and A indicate control ladders that were generated by dideoxynucleotide sequencing of unmodified, cleaved HDV88 RNA. Lane 1, probing under native conditions; lane 2, probing under semi-denaturing conditions; lane 3, probing under denaturing conditions; lane 4, RNA treated in the same way as the sample in lane 1, except that CMCT was not included in the reaction mixture.

addition of 5 μL of stop solution (9 M urea, 50 mM EDTA, 0.1% BPB, and 0.1% XCFF). Dideoxynucleotide sequencing of unmodified HDV88 RNA was carried out in the same manner, except that the reactions contained 150 μM ddNTP (Sanger *et al.*, 1977).

RESULTS AND DISCUSSION

In tRNAs, a number of bases contribute to secondary and tertiary interactions both by participating in hydrogen bonding with other bases or phosphates and also by participating in the coordination of metal ions. Base interactions determined by the use of various chemical probes have been explained by reference to crystallographic data for tRNAs. While specific tertiary interactions in other RNAs are very difficult to identify when crystallographic data are not available, chemical probes can help us to locate bases that may be involved in such interactions.

Modification Strategy. In our studies, we monitored the four bases at their Watson–Crick base-pairing positions by treatment with dimethyl sulfate at A(N1) and C(N3) and by treatment with CMCT at U(N3) and G(N1). Positions of G(N7) residues were probed by dimethyl sulfate, with subsequent treatment with $NaBH_4$ and aniline. The extent of modifications was examined under native, semi-denaturing, and denaturing conditions. Although these chemical probes react specifically with various nucleotides, the reactions proceed at different rates. Thus, dimethyl sulfate reacts with C(N3) more slowly than with A(N1), and CMCT reacts with G(N1) more slowly than with U(N3). The sites of modification were identified by the reverse transcriptase reaction after hybridization of primers with modified 3′-cleaved products of the active variants (HDV88, C695G:G719C, and U767A) of HDV ribozyme. Bands resulting from the terminations of reverse transcriptase reactions each correspond to the DNA that extends from the 5′-end of the primer to the nucleotide that precedes the modified nucleotide. Identification of the modified base was carried out by the inclusion of four parallel sequencing ladders on the same gel. Since primer extension does not extend to the modified bases, modifications can be recognized as occurring one position below the corresponding band, in sequencing lanes, when base modification has taken place. When the transcribed HDV RNA was used for 3′-end-labeling with [α-^{32}P]pCp, heterogeneity was often observed at the 3′-ends (one- or two-base

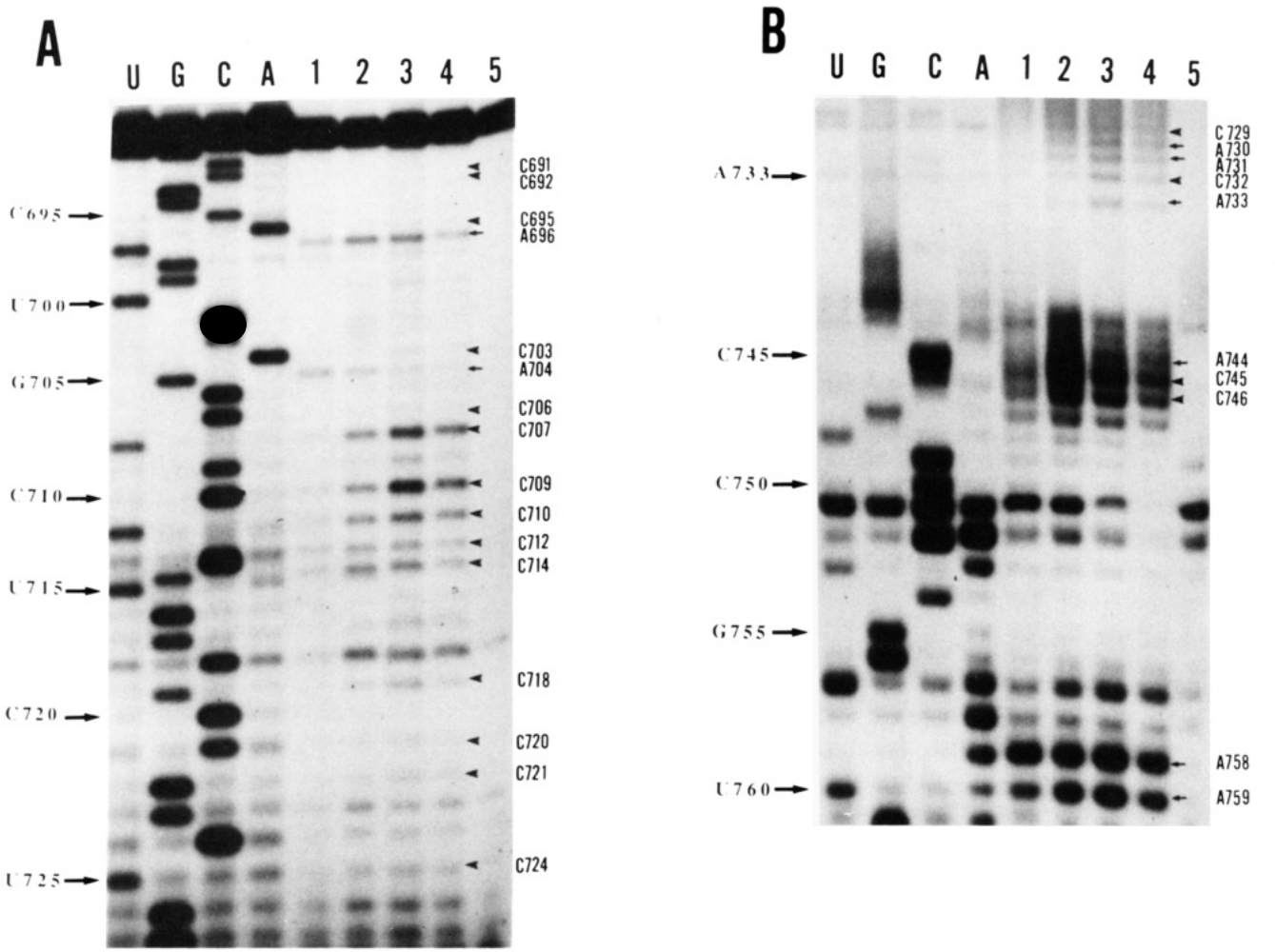

FIGURE 4: Autoradiography showing sites of chemical modification by dimethyl sulfate, without strand scission by $NaBH_4$ and aniline, of the 3′-product of HDV88. Sites were identified after hybridization with primer 2 (A) or primer 1 (B) and reverse transcription. Bold arrows indicate the numbering of bases on the RNA molecule. Arrowheads indicate the modifications of C residues (at N3). Small arrows indicate the modifications of A residues (at N1). U, G, C, and A are control ladders that were generated by dideoxynucleotide sequencing of unmodified, cleaved HDV88 RNA. Lane 1, probing under native conditions; lane 2, probing under semi-denaturing conditions; lane 3, probing under denaturing conditions; lane 4, probing under strongly denaturing conditions (same as denaturing conditions but the reaction was carried out at 90 °C for 5 min); lane 5, RNA treated in the same way as the sample in lane 1, except that dimethyl sulfate was not included in the reaction mixture.

difference), resulting in doubling of the sequence ladders. Therefore, we used the primer extension analysis method to identify the sites of base modifications. Primer 1 (which hybridizes the ribozyme between nt 764 and 775) did not allow the reaction to proceed to the 5′-end of the molecule under our conditions (Figure 1). We attempted to improve the extension conditions, for example, by varying ratios of primer to template, annealing conditions, components of the buffer used for reverse transcription, and the concentrations of dNTPs. No significant improvement was observed. However, using primer 2 (which hybridizes the ribozyme between nt 730 and 741) we could monitor the alkylation reaction as far as the 5′-end of the 3′-cleaved RNA product of the HDV ribozyme. Since it is very difficult to examine the structure of the native self-cleaving RNA under conditions that allow rapid self-cleavage, for chemical probing studies we used the 3′-cleaved product instead of the original enzymatically active transcript. This approach is justified because, in the case of the HDV ribozyme, only one nucleotide 5′ to the cleavage site is sufficient for efficient cleavage (Perrotta & Been, 1990). Therefore, the original structure of the HDV ribozyme is most probably maintained, even after dissociation of the single 5′-cleaved nucleotide. Nuclease-probing data have been presented for the 3′-cleaved products of genomic and antigenomic HDV ribozymes on the basis of this very assumption (Rosenstein & Been, 1991). In fact, our nuclease-probing data suggest that both the original transcripts and the 3′-products of active variants have nearly the same patterns of sensitivity to ribonucleases (unpublished data).

Figures 2–4 show representative autoradiograms from chemical probing experiments with the 3′-cleaved product of HDV88. These results are summarized in Table 1. Among the various bases of the HDV88 ribozyme, structural interactions were detected in the case of 52 bases. Modifications were not detected at the C(N3) and G(N1) positions, even under denaturing conditions for bases C691–C692, G694, C695, G719 (in the stem I region), C701, C702 (in the stem II region), G705, G716 (in the stem III region), and C736–G740, C749, C752, and C754 (in the stem IV region). The results suggest that strong interactions exist in these base-pairing regions because the base-pairings were not disrupted under the conditions tested. Thus, several shorter forms of the HDV ribozyme, which had previously been shown to possess self-cleavage activity under denaturing conditions (Perrotta & Been, 1990; Prasad *et al.*, 1992), appeared to persist as active structures under our denaturing conditions. We also observed variations in the mobility of 3′-cleaved products (with the same 5′- and 3′-ends) in the case of stem

Table 1: Results of Chemical Probing of the Structure of the HDV88 Ribozyme (3′-Product of HDV88) and the Prediction of Base Interactions[a]

base	chemical modification	native	semi-denaturing	denaturing	prediction
G689			end base		
G690	DMS G(N7)	–	+	+	interaction at N7G
G691	DMS C(N3)	–	–	–	strong interaction at N3C
C692	DMS C(N3)	–	–	–	strong interaction at N3C
G693	DMS G(N7)	+	++	+	interaction at N7G
G693	CMCT G(N1)	–	+	–	interaction at N1G
G694	DMS G(N7)	–	+	+	interaction at N7G
G694	CMCT G(N1)	–	–	–	strong interaction at N1G
G695	DMS C(N3)	–	–	–	strong interaction at N3C
A696	DMS A(N1)	++	++	++	no interaction at N1A
U697	CMCT U(N3)	++	++	++	no interaction at N3U
G698	DMS G(N7)	+	++	+	interaction at N7G
G698	CMCT G(N1)	+	+	+	no interaction at N1G
G699	DMS G(N7)	++	++	++	no interaction at N7G
G699	CMCT G(N1)	+	+	+	no interaction at N1G
U700	CMCT U(N3)	++	++	+	no interaction at N3U
C701	DMS C(N3)	–	–	–	no interaction at N3C
C702	DMS C(N3)	–	–	–	no interaction at N3C
C703	DMS C(N3)	–	–	+	W–C base-pairing at N3C
A704	DMS A(N1)	++	+	+	no interaction at N1A
G705	DMS G(N7)	–	++	++	interaction at N7G
G705	CMCT G(N1)	–	–	–	strong interaction at N1G
C706	DMS C(N3)	–	–	+	W–C base-pairing at N3C
C707	DMS C(N3)	–	++	+++	interaction at N3C
U708	CMCT U(N3)	++	++	++	no interaction at N3U
C709	DMS C(N3)	+	++	+++	interaction at N3C
C710	DMS C(N3)	+	++	++	interaction at N3C
U711	CMCT U(N3)	+++	+++	++	no interaction at N3U
C712	DMS C(N3)	+	+	+	no interaction at N3C
G713					migrates with adjacent bases
C714	DMS C(N3)	+	++	++	interaction at N3C
U715	CMCT U(N3)	++	+	+	no interaction at N3U
G716	DMS G(N7)	+	++	+	interaction at N7G
G716	CMCT G(N1)	–	–	–	strong interaction at N1G
G717	DMS G(N7)	++	++	++	no interaction at N7G
C718	DMS C(N3)	–	–	+	W–C base-pairing at N3C
G719	DMS G(N7)	–	++	++	interaction at N7G
G719	CMCT G(N1)	–	–	–	strong interaction at N3C
C720	DMS C(N3)	–	–	+	W–C base-pairing at N3C
C721	DMS C(N3)	–	–	+	W–C base-pairing at N3C
G722	DMS G(N7)	+++	+++	++	no interaction at N7G
G722	CMCT G(N1)	+	++	++	no interaction at N1G
G723	DMS G(N7)	++	++	+	no interaction at N7G
G723	CMCT G(N1)	++	++	+	no interaction at N1G
C724	DMS C(N3)	+	+	+	no interaction at N3C
U725	CMCT U(N3)	++	+++	+++	interaction at N3U
G726	DMS G(N7)	++	+++	+	interaction at N7G
G726	CMCT G(N1)	++	+++	+++	no interaction at N1G
G727					band hidden in the primer 2 portion
G728					band hidden in the primer 2 portion
C729	DMS C(N3)	–	+	++	interaction at N3C
A730	DMS A(N1)	–	+	++	interaction at N1A
A731	DMS A(N1)	–	+	++	interaction at N1A
C732	DMS C(N3)	–	+	++	interaction at N3C
A733	DMS A(N1)	–	+	++	interaction at N1A
U734	CMCT U(N3)	++	+	–	no interaction at N3U
U735	CMCT U(N3)	++	+	–	no interaction at N3U
C736	DMS C(N3)	–	–	–	strong interaction at N3C
C737	DMS C(N3)	–	–	–	strong interaction at N3C
G738	CMCT G(N1)	–	–	–	strong interaction at N1G
A739	DMS A(N1)	–	–	–	strong interaction at N1A
G740	CMCT G(N1)	–	–	–	strong interaction at N1G
G741					natural reverse transcriptase stop
G742					natural reverse transcriptase stop
G743	DMS G(N7)	++	+++	++	interaction at N7G
A744	DMS A(N1)	–	++	++	interaction at N1A
C745	DMS C(N3)	+	+++	++	interaction at N3C
C746	DMS C(N3)	+	+++	++	interaction at N3C
G747	DMS G(N7)	++	++	+	no interaction at N7G
G747	CMCT G(N1)	++	++	+	no interaction at N1G
U748	CMCT U(N3)	+	++	–	interaction at N3U
C749	DMS C(N3)	–	–	–	strong interaction at N3C
C750					natural reverse transcriptase stop
C751					natural reverse transcriptase stop
C752	DMS C(N3)	–	–	–	strong interaction at N3C

Table 1 (Continued)

base	chemical modification	native	semi-denaturing	denaturing	prediction
U753	CMCT U(N3)	+	–	+	no interaction at N3U
C754	DMS C(N3)	–	–	–	strong interaction at N3C
G755	DMS G(N7)	+	++	++	interaction at N7G
G756					natural reverse transcriptase stop
U757	CMCT U(N3)	+++	+++	+++	no interaction at N3U
A758	DMS A(N1)	+++	+++	+++	no interaction at N1A
A759	DMS A(N1)	++	+++	+++	interaction at N1A
U760	CMCT U(N3)	+++	+++	+++	no interaction at N3U
G761	DMS G(N7)	+	++	+++	interaction at N7G
G761	CMCT G(N1)	+++	+++	+++	no interaction at N1G

[a] Strong interaction: Base-pairing regions were not disrupted under the conditions tested herein.

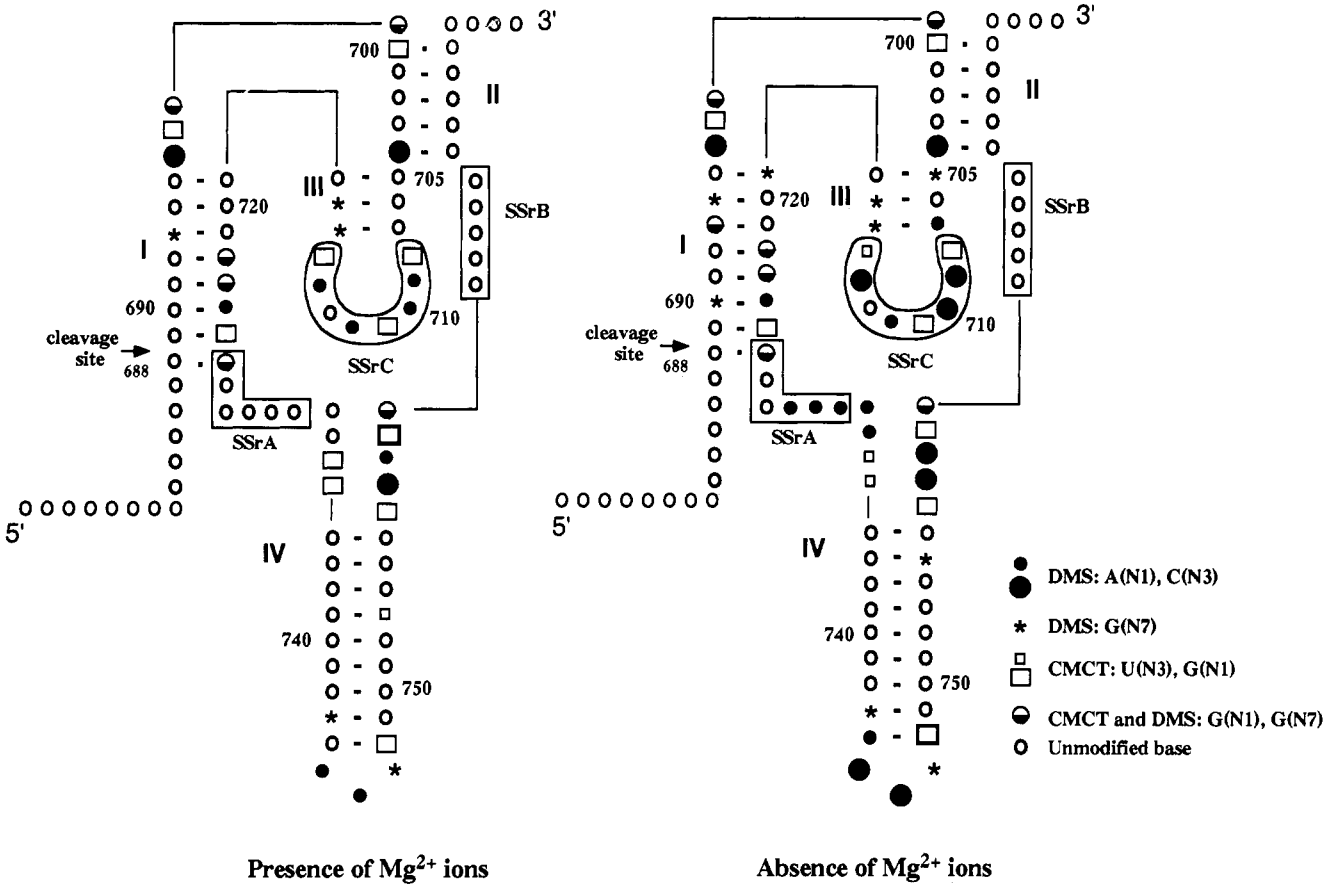

FIGURE 5: Schematic diagrams of the secondary structure of the HDV88 ribozyme (Perrotta & Been, 1991), with summaries of the results of chemical probing experiments under native (presence of Mg^{2+} ions) and semi-denaturing conditions (absence of Mg^{2+} ions) at 30 °C. Small symbols indicate lower levels of modification and large symbols indicate effective modifications by the chemical probes. The reactivities that are indicated here were derived from the results of two or more experiments and are based on visual estimates of the intensity of the bands.

and SSrC (bases between nt 708–715) variants on denaturing polyacrylamide gels (Kumar *et al.*, 1993; Kawakami *et al.*, 1993). The prediction of interactions at G741, G742, C750, C751, and G756 was difficult because of the presence of natural stops of reverse transcriptase that were observed even with unmodified RNA. We could not analyze the bases between nt 762 and 770 (8 nt) at the 3'-end of the HDV88 molecule, because this region includes the sequence complementary to primer 1 (Figure 1).

Identification of Structural Interactions within the HDV Ribozyme. We predicted that the secondary structure remains undisturbed under native (presence of Mg^{2+} ions, 30 °C) and semi-denaturing (absence of Mg^{2+} ions, 30 °C) conditions and that single-strand-specific chemical probes would not attack the base-paired regions under these two sets of conditions. Inspection of Table 1 and Figure 5 suggests that, under these conditions, the chemical probes mainly attacked single-stranded regions of the HDV ribozyme, as predicted by the pseudoknot structure, with some exceptions in the stem I and stem IV regions. The bases within A696–G699, C707–U715, G722–U735, A744–U748, and U757–G761 were readily attacked by dimethyl sulfate and CMCT. Since these reagents primarily react with A(N1), C(N3), and G(N1), it can, thus, be assumed that these bases are in the single-stranded regions of the HDV88 ribozyme. When these results are analyzed by reference to the pseudoknot secondary structure, they support the proposed pseudoknot structure, for the most part. Our present results suggest that the C732–U735 and U757–G761 regions are in single-stranded regions (or, at least, they are in equilibrium with the single-stranded form), and therefore, these results are not in complete agreement with the pseudoknot model in which these bases are involved in the

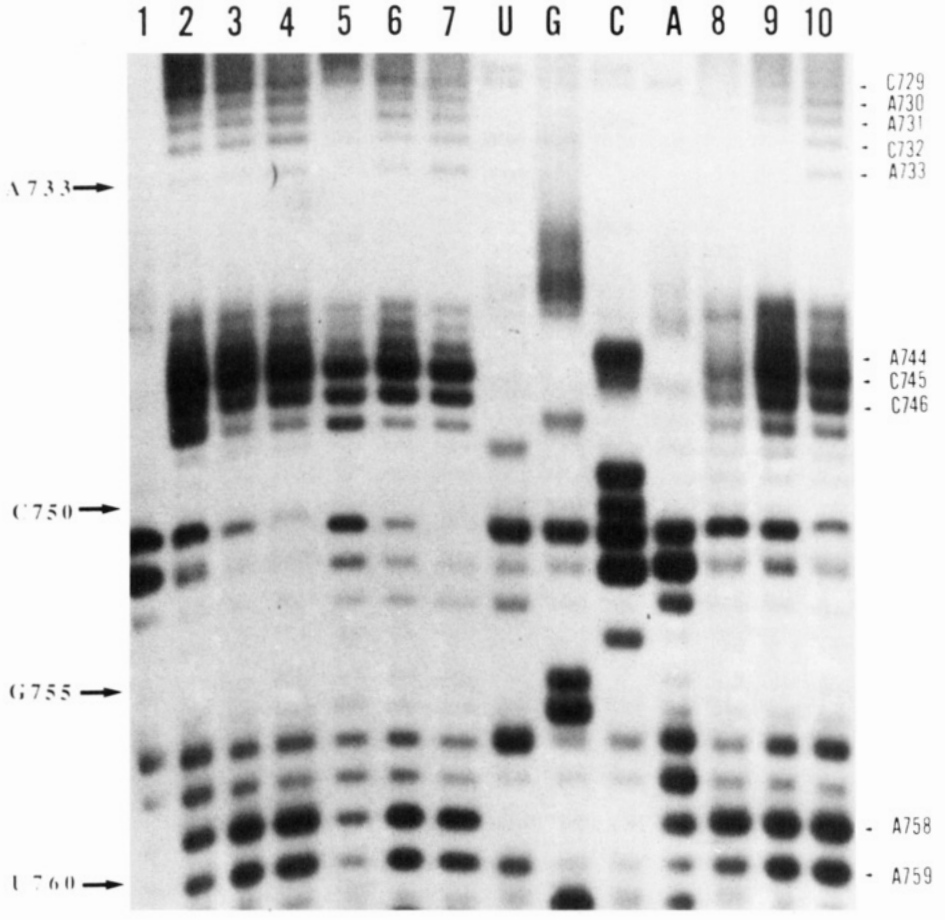

FIGURE 6: Autoradiography showing sites of chemical modification by dimethyl sulfate, without strand scission by $NaBH_4$ and aniline, of active stem variants (C695G:G719C and U767A) of the 3′-product of HDV88. Primer 1 was used for primer extension analysis. Bold arrows indicate the numbering of bases on the RNA molecule. Thin arrows indicate the modifications of A residues (at N1) and C residues (at N3). U, G, C, and A indicate control ladders that were generated by dideoxynucleotide sequencing of the unmodified, cleaved fragment of HDV88 RNA. Lane 1, HDV88 RNA treated in the same way as the samples in lanes 2, 5, or 8 except that dimethyl sulfate was not included in the reaction mixture; lanes 2–4, probing of the 3′-cleaved product of the C695G:G719C variant under native conditions (lane 2), semi-denaturing conditions (lane 3), and denaturing conditions (lane 4); lanes 5–7, probing of the 3′-cleaved product of the U767A variant under native conditions (lane 5), semi-denaturing conditions (lane 6), and denaturing conditions (lane 7); lanes 8–10, probing of the 3′-cleaved product of the HDV88 variant under native conditions (lane 8), semi-denaturing conditions (lane 9), and denaturing conditions (lane 10). The radioactivity was balanced for lanes 1–10 before loading the sample and was about 9000 cpm/lane.

formation of the stem IV region. Bases between C736 and G756 appeared to form a very stable stem IV region. Furthermore, our results indicate some discrepancies in the stem I region at the G722–U725 positions that cannot be easily explained.

To identify the interactions at N7 of G residues in the HDV88 ribozyme, we subjected the 3′-product of HDV88 to modification by dimethyl sulfate (under native, semi-denaturing, and denaturing conditions) with subsequent treatment with $NaBH_4$ and aniline. Since N7 is not involved in Watson–Crick base-pairing, it is possible to modify all G residues in single-stranded regions and in base-paired regions, unless they are involved in other interactions (either base–base or base–phosphate or –magnesium coordination) or are hidden inside the molecule. We found that, for most of the G residues in the HDV ribozyme, N7 was reactive either in the presence or absence of Mg^{2+} ions (Figure 2A,B). To generalize the dimethyl sulfate reactivities with the N7 of G residues in HDV ribozyme, we compared the modification sites in the presence or absence of Mg^{2+} ions at 30 °C and found that bases G690, G693, G694, G698, G705, G719, G743, G755, and G761 are protected to some extent when Mg^{2+} ions are present. The above interacting residues probably play some role in the formation of the active structure of the HDV88 ribozyme. Similarly, when N7 of A was probed with diethyl pyrocarbonate, none of the A residues showed evidence of interactions in the 3′-cleaved product of HDV88 (data not shown). In summary, both stem I and stem–loop region III include bases that are more likely to be involved in tertiary interactions than the bases in stem II and stem IV.

Structures Formed in the Presence or Absence of Mg^{2+} Ions. When the reactivities of nucleotides with chemical probes are superimposed on the proposed pseudoknot model under native (presence of Mg^{2+} ions) and semi-denaturing conditions (absence of Mg^{2+} ions), some of the bases in proposed single-stranded regions and the N7 of G residues are protected under native conditions in the HDV88 ribozyme. These results are summarized in Figure 5. When modification of the 3′-product of HDV88 by dimethyl sulfate was carried out in the presence or absence of Mg^{2+} ions, bases C729, A730, A731, A732, and A733 in single-stranded regions were protected in the presence of Mg^{2+} ions. Similarly, bases C709, C707, and C714 were protected to a greater extent than when there were no Mg^{2+} ions in the reaction mixture. As mentioned earlier, N7 of G690, G693, G694, G698, G705, G719, G743, G755, and G761 in base-paired regions were also protected in the presence of Mg^{2+} ions, indicating that some interactions exist in the native structure. Modifications of bases in other stems (II and IV) were found to be similar under the two different conditions.

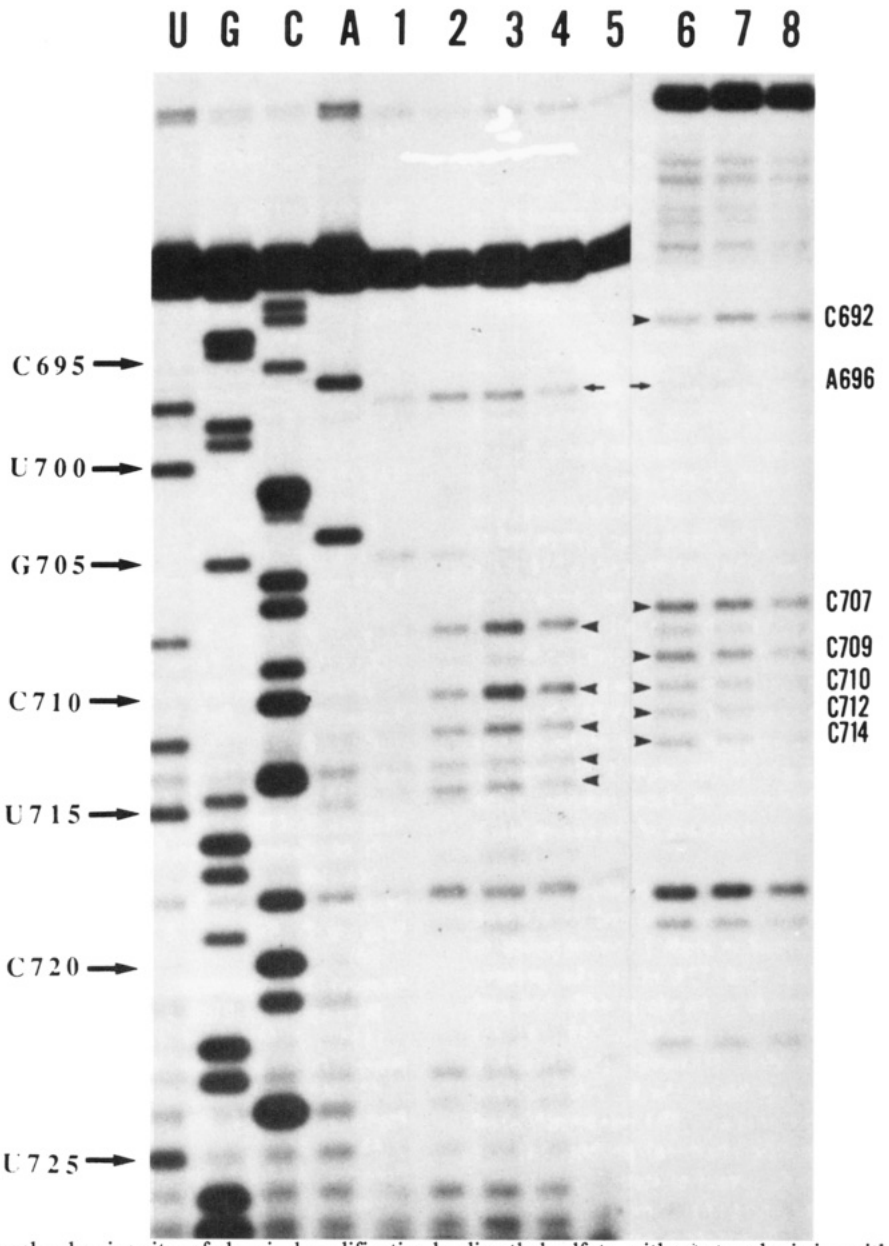

FIGURE 7: Autoradiography showing sites of chemical modification by dimethyl sulfate, without strand scission with $NaBH_4$ and aniline, of HDV88 (3′-cleaved product) and the C718G variant (the original transcript). Primer 2 was used for primer extension analysis. Bold arrows indicate the numbering of bases on the RNA molecule. Small arrows indicate the modifications of A residues (at N1), and modifications of C residues (at N3) are indicated by arrowheads. U, G, C, and A indicate control ladders that were generated by dideoxynucleotide sequencing of the unmodified, cleaved fragment of the HDV88 RNA. Lanes 1–5, probing of the 3′-cleaved product of the HDV88 under native conditions (lane 1), semi-denaturing conditions (lane 2), denaturing conditions (lane 3), strongly denaturing conditions (same as the denaturing conditions, but the reaction was carried out at 90 °C for 5 min; lane 4), and the 3′-product of HDV88 RNA was treated in the same way as the sample in lane 1, except that dimethyl sulfate was not included in the reaction mixture (lane 5); lanes 6–8, probing of the original transcript of the C718G variant under native conditions (lane 6), probing under semi-denaturing conditions (lane 7), and probing under denaturing conditions (lane 8).

Probing Structural Dissimilarities in Stem Variants. We have previously noted the differences in mobility on denaturing polyacrylamide gels of 3′-cleaved products of stem variants, despite the presence of the same respective 5′-and 3′-ends (Kumar *et al.*, 1993). These differences suggest that conformational variations lead to differences in mobility. Nuclease-probing studies have allowed us to identify, to some extent, the conformational differences that exist among stem variants (unpublished data). However, because of their size, nucleases are very sensitive to steric hindrance. Moreover, enzymatic cleavage at a particular site may alter the structure of the resulting RNA, thereby inducing secondary cuts in the RNA molecule. To avoid these problems, we used chemical probes to examine structural differences between stem variants because of their small size and because of the lower sensitivity of chemical reagents to steric hindrance.

When modification by dimethyl sulfate of C(N3) and A(N1) was performed on two 3′-products of stem variants, namely, U767A and C695G:G719C, we observed that the bases between C729 and A733 gave the same pattern of reactivity (i.e., bases were protected in the presence of Mg^{2+} ions) in the U767A variant as in the HDV88 variant (compare lanes 5–7 with lanes 8–10 in Figure 6). By contrast, in the C695G:G719C variant, bases C729–A733 had the same reactivity in the presence and in the absence of Mg^{2+} ions (compare lanes 2–4 with lanes 8–10). These results imply that mutations at the C695G:G719C positions in the mutant ribozyme (C695G:G719C variant) destroyed the interactions in the C729–A733

region. Further, it is also of interest to note that the A744–C745 and A758 positions showed similar patterns of reactivity (compare lanes 2 and 5 in Figure 6) in two active variants (C695G:G719C and U767A variants). Moreover, in these two regions, they differ from the 3′-product of HDV88. Among the active variants, different patterns of reactivity were observed only in regions between C729 and A759, which mainly constitute the stem IV region in the pseudoknot molecule. These results suggest that variations in the mobility of active variants of the HDV ribozyme are probably due to differences in interactions in the stem IV regions without any loss of stability of the active conformation.

When we extended our studies to probe the differences between HDV88 and the C718G variant (an inactive variant with about 20% of the cleavage activity of HDV88) with dimethyl sulfate, we observed that the molecules had different reactivities (Figure 7). Although C692 and C707–C714 showed similar reactivity under native, semi-denaturing, and denaturing conditions in the C718G variant as compared to HDV88, we noted with interest that A696 and A704 were protected under native, semi-denaturing, and denaturing conditions in the C718G variant, while in all of the other active variants, A696 and A704 were reactive under all conditions tested (compare lanes 1–3 with lanes 6–8 in Figure 7). In contrast to the active variants of the HDV88 ribozyme, differences in reactivity in the alkylation reaction of the inactive ribozyme were observed mainly in the stem I–III regions, probably indicating the importance of these regions in the overall folding that generates the structure of the active HDV ribozyme.

With the genomic ribozyme (3′-end product), strong ribonuclease V1 (which recognizes base-paired or stacked regions) digestions were observed in stem I (corresponds to bands at 691, 692, 721, 722, and 724 nt in the present numbering), stem II (corresponds to bands at 699–704 and also at 769 nt), and stem IV (corresponds to 738–741 nt) regions in the pseudoknot model. However, the sequence forming stem III was not susceptible to the V1 ribonuclease (Rosenstein & Been, 1991; Kumar *et al.*, 1993; our unpublished data). Our single-strand-specific chemical probes that are used in the present study failed to react at stem I–IV regions in the pseudoknot model, suggesting that they are involved in Watson–Crick base-pairings. Rosenstein and Been (1991) found that nuclease sensitivity of HDV ribozyme (3′-cleaved product) toward single-strand-specific ribonucleases appeared to be different for some regions in the presence or absence of Mg^{2+} ions. For C745 and C746 in hairpin loop IV, the sensitivity to RNaseA was reduced in the presence of Mg^{2+} ions compared to the absence of Mg^{2+} ions. Similarly, when dimethyl sulfate modification [for probing the C(N3) position] was carried out on HDV ribozyme, the sensitivity for C745 and C746 in the presence of Mg^{2+} ions was reduced (Figures 4B and 5). These results suggest that structural rearrangements may occur in HDV ribozyme in the presence of Mg^{2+} ions, especially in inducing tertiary contacts.

Our present chemical probing studies provide data that support the postulated pseudoknot structure of the genomic HDV ribozyme, as well as clues to the identity of possible interacting bases. We also found that the C732–U735 and U757–G761 regions were at least in equilibrium with a single-stranded form, a finding that may be in conflict with the pseudoknot model wherein these bases are involved in a very stable stem IV region. Using our chemical probes, we observed differences in accessibility between active and inactive variants that suggested conformational variations, in particular around the stem I and III regions. These studies provide us with a reasonable starting point for future attempts at understanding the tertiary folding of the HDV ribozyme.

REFERENCES

Been, M. D., Perrotta, A. T., & Rosenstein, S. P. (1992) *Biochemistry 31*, 11843–11852.

Branch, A. D., & Robertson, H. D. (1991) *Proc. Natl. Acad. Sci. U.S.A., 88*, 10163–10167.

Govindarajan, S., Chin, K. P., Redeker, A. G., & Peter, R. L. (1984) *Gastroenterology 86*, 1417–1420.

Inoue, T., & Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 648–652.

Kawakami, J., Kumar, P. K. R., Suh, Y.-A., Nishikawa, F., Kawakami, K., Taira, K., Ohtsuka, E., & Nishikawa, S. (1993) *Eur. J. Biochem. 217*, 29–36.

Krol, A., & Carbon, P. (1989) *Methods Enzymol. 180*, 212–227.

Kuchino, Y., & Nishimura, S. (1989) *Methods Enz*ymol. *180*, 154–163.

Kumar, P. K. R., Suh, Y.-A., Miyashiro, H., Nishikawa, F., Kawakami, J., Taira, K., & Nishikawa, S. (1992) *Nucleic Acids Res. 20*, 3919–3924.

Kumar, P. K. R., Suh, Y.-A., Taira, K., & Nishikawa, S. (1993) *FASEB J. 7*, 124–129.

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

Makino, S., Chang, M.-F., Shieh, C.-K., Kamahora, T., Vannier, D. M., Govindarajan, S., & Lai, M. M. C. (1987) *Nature 329*, 343–346.

Moazed, D., & Noller, H. F. (1986) *Cell 47*, 985–994.

Moazed, D., Stern, S., & Noller, H. F. (1986) *J. Mol. Biol. 187*, 399–416.

Mougel, M., Eyermann, F., Westhof, E., Romby, P., Expert-Bezançon, A., Ebel, J.-P., Ehresmann, B., & Ehresmann, C. (1987) *J. Mol. Biol. 198*, 91–107.

Peattie, D. A., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A., 77*, 4679–4682.

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

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

Prasad, Y., Belinsky, M. G., Smith, J. B., & Dinter-Gottlieb, G. (1992) *Antisense Strategies 660*, 277–279.

Rosenstein, S. P., & Been, M. D. (1991) *Nucleic Acids Res. 19*, 5409–5416.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A., 74*, 5463–5467.

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

Stern, S., Wilson, R. C., & Noller, H. F. (1986) *J. Mol. Biol. 192*, 101–110.

Suh, Y.-A., Kumar, P. K. R., Nishikawa, F., Kayano, E., Nakai, S., Odai, O., Uesugi, S., Taira, K., & Nishikawa, S. (1992) *Nucleic Acids Res. 20*, 747–753.

Suh, Y.-A., Kumar, P. K. R., Kawakami, J., Nishikawa, F., Taira, K., & Nishikawa, S. (1993) *FEBS Lett. 326*, 158–162.

Taira, K., & Nishikawa, S. (1992) in *Gene Regulation: Biology of Anti-Sense RNA and DNA* (Erickson, R. P., & Izant, J. G., Eds.) Vol. 1., pp 35–54, Raven Press, Ltd., NY.

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

Wu, H.-N., Wang, Y.-J., Hung, C.-F., Lee, H.-J., & Lai, M. M. C. (1992) *J. Mol. Biol. 223*, 233–245.