Nucleic Acids Research, 1996, Vol. 24, No. 7 **1351–1359**

Folding of the HDV antigenomic ribozyme pseudoknot structure deduced from long-range photocrosslinks

Catherine Bravo, Franck Lescure[1], Philippe Laugâa, Jean-Louis Fourrey[2] and Alain Favre*

Laboratoire de Photobiologie Moléculaire, Institut Jacques Monod, CNRS, Université Paris 7, 2 Place Jussieu, 75251 Paris Cedex 05, France, [1]Laboratoire de Virologie Moléculaire, Université Paris 7, 2 Place Jussieu, 75251 Paris Cedex 05, France and [2]Institut de Chimie des Substances Naturelles, CNRS, 91198 Gif-sur-Yvette Cedex, France

Received November 13, 1995; Revised and Accepted February 12, 1996

ABSTRACT

A *trans*-acting system has been designed in order to explore the three-dimensional structure of the antigenomic HDV ribozyme. In this system, the substrate (SANT) is associated by base-pairing to the catalytic RNA (RzANT) forming helix H1. RzANT is able to cleave specifically the RNA substrate as well as a deoxysubstrate analogue containing a single ribocytidine at the cleavage site (position –1). This demonstrates that such deoxysubstrate analogues are valuable tools for structural studies of this ribozyme domain. They form however weak complexes with RzANT which is due in part to their ability to fold as stable hairpins unlike the RNA substrate. Using a set of full deoxy or of mixed deoxy-ribo substrate analogues site-specific substituted with the photoaffinity probe deoxy-4-thiouridine, ds^{4}U, at a defined position, we were able to determine a number of long range contacts between the substrate and the ribozyme core. In particular, crosslinks between substrate position –1 and position –2 with residues C15, G19 and C67, thought to be involved in the ribozyme catalytic site, were detected. A three dimensional model of the antigenomic ribozyme system, derived from the structure proposed by Tanner *et al.* [*Current Biol.* (1994) 4, 488–498] for the genomic system was constructed. Apart from residue deletion or insertion, only minor accommodations were needed to account for all photocrosslinks but one which is attributed to an alternative hybridization of the substrate with the ribozyme. This study therefore further supports the structure proposed by Tanner *et al.* for the pseudoknot model.

INTRODUCTION

Hepatitis delta virus (HDV) is an infectious human pathogen that requires the coat proteins of hepatitis B virus for encapsulation. Both the genomic and antigenomic HDV RNAs probably replicate by a RNA-dependent rolling-circle mechanism that generates linear multimers (1). These multimers exhibit self-cleaving activity which requires the presence of divalent cations and generates a 2′,3′-cyclic phosphate and a free 5′-hydroxyl at the cleavage site (2–5).

Under *in vitro* conditions, optimal catalytic activity is obtained with 85 nucleotide (nt) long genomic or antigenomic ribozymes, i.e. with only one nucleotide 5′ to the cleavage site (6). The corresponding sequences cannot be folded in the hammerhead or hairpin forms and therefore represent a distinct self-cleaving motif (7). At least five different secondary structures have been proposed (2,7–10) but the pseudo-knot model of Perotta and Been, which is common to both the genomic and antigenomic ribozymes, best fits the experimental data (7,13). It contains four base-paired regions (H1–H4) and the cleavage site is positioned at the 5′-end of the 84 nt ribozyme sequence. H3 and H4 are stems with hairpin structures and, together with H1 and H2, form a pseudo-knot. Hairpin IV can be reduced in size without complete loss of catalytic activity (11–13) and residues involved in the catalytic site are thought to be located within the loop of hairpin III and in the sequence connecting hairpin IV to helix II (11–12,14). Recently, a three dimensional model of the genomic HDV ribozyme based on the pseudo-knot secondary structure has been proposed (15). In this genomic model two conserved residues of loop III and a residue of junction II/IV essential for cleavage activity are positioned in close proximity to the scissile phosphodiester bond (for review see ref. 16).

In order to check this model we have attempted to identify couples of residues which are distant in the primary sequence but maintained in close contact within the folded structure of the ribozyme. This can best be achieved using an intrinsic photoaffinity methodology based on the site-specific insertion of thionucleotides. The corresponding thiobases can be selectively photoexcited with light of wavelength >310 nm and then form mixed photoadducts with neighbour residues. In particular, 4-thiouracil can substitute U residues in RNA (T in DNA) with minimal structural perturbation and is able, upon 365 nm irradiation, to yield stable photoadducts

* To whom correspondence should be addressed

with any of the current nucleosides (17–19). So far, this methodology has been successfully applied to the analysis of the conformation of various RNA in solution including the hairpin or hammerhead ribozyme domains (20–21).

In the present work we have taken advantage of the capability of the HDV ribozyme to undergo *trans*-cleavage reactions. Two forms of *trans* reaction have been generated (10,22). In the first form the self-cleaving molecule is interrupted at H4 and the two sequences reassociate via interactions within H4 and H2. In the alternative form, the sequence is separated between H1 and H2 and the two sequences reassociate by base-pairing within H1 to generate the cleavable structure. Our present studies were conducted with this latter form of the antigenomic ribozyme since (i) it exhibits higher catalytic activity in *trans*, (ii) the minimal substrate size is only 8 nt long (Fig. 1) and, accordingly, substrate or substrate analogues containing the deoxy-4-thiouridine probe (ds^{4}U) inserted at a selected position can readily be obtained by chemical synthesis. Here we show for the first time that a deoxy-substrate analogue containing a single ribocytidine at the cleavage site (position –1) can be enzymatically cleaved. This substrate analogue exhibits much lower binding affinity toward the ribozyme than its parent substrate which is due in part to its ability to fold as a hairpin. Using a series of full deoxy or mixed ds^{4}U containing substrate analogues, we have then determined a number of long-range contacts within the hybrid substrate analogue–antigenomic ribozyme complexes. Overall, our data are consistent with a three dimensional folding of the antigenomic ribozyme derived from the one proposed by Tanner *et al.* (15) for the genomic ribozyme.

MATERIALS AND METHODS

Synthesis of the HDV antigenomic ribozyme (RzANT) and its substrate (SANT)

The *trans*-cleaving ribozyme (RzANT) was synthesized using a large scale transcription method (24) for <100 nt long RNAs. We used the sequence that Perrotta *et al.* proposed (14), including the non-HDV derived nucleotides. The transcription buffer contained 40 mM Tris–HCl pH 8.0, 15 mM $MgCl_2$, 5 mM DTT and 1 mM spermidine (23). The template was a double-stranded DNA containing a T7 RNA polymerase promoter formed by the hybridization of two chemically synthesized (GENSET, France) and purified DNAs. The final concentrations of the bacteriophage T7 RNA polymerase and the DNA template were 5.6 U/ml and 200 nM, respectively. The 2 ml reaction mixture was incubated 4 h at 37°C. After a PAGE purification, we obtained 0.21 mg of unlabelled active RNA. The ribozyme was dephosphorylated using calf intestinal phosphatase (Pharmacia), and the mixture treated with 10 U Proteinase K in the presence of 2% SDS. After a phenol extraction, the ribozyme was precipitated with ethanol and then phosphorylated with T4 polynucleotide kinase and [γ-^{32}P]ATP at 37°C for 30 min. The reaction was terminated by adding an equal volume of formamide containing xylene cyanol and purified by gel electrophoresis in 10% acrylamide. The ribozyme was located by autoradiography, cut from the gel, eluted overnight in NaCl (0.15 M) and precipitated with ethanol. The corresponding 15mer RNA substrate SANT was synthesized using the same method.

Substrate analogues synthesis

Full deoxy or mixed substrate analogues were synthesized by GENSET (France) using the phosphoramidite method. Analogues containing a single ds^{4}U residue at a selected position were synthesized similarly, following the protocol which uses 4-*S*-pivaloyloxymethyl-4-thiodeoxyuridine phosphoramidite (17). After deprotection the oligomers were purified by 20% PAGE in the presence of 8 M urea. Fully thiolated oligomers were obtained by an additional affinity electrophoresis step (20,25). These thiolated oligomers were referenced in Table 1. The substrates used in the cleavage reaction are either oligoribonucleotides SANT, 5′-GGUUCGGUCGGCAU-3′; SANT′, 5′-UUGGUUCGGUCGGCAU-3′; or a mixed DNA–RNA analogue S′, 5′-ccaaaCgggtcggt-3′. In native gel electrophoresis experiments C, 5′-UGCCCGUCUGUUGU-3′, was used as a control.

Trans-cleavage reactions

The incubation buffer was usually 10 mM $MgCl_2$ and 50 mM Tris–HCl (pH 8.0) and *trans*-cleavage reactions were initiated by adding the 5′-end-labelled substrate and incubating at 37 or 55°C. Aliquots were removed at various time intervals and reactions were stopped by adding an equal volume of 50 mM EDTA, 0.1% xylene cyanol and 80% formamide. Samples were electrophoretically separated on a denaturating 20% acrylamide gel and autoradiographed. The bands corresponding to full-length precursors and 5′ cleaved products were quantified by counting Cerenkov emission of the excised gel slices in a Beckman LS 6000IC scintillation counter. The percentage of cleavage was determined by taking the ratio of product to total radioactivity for each time point.

Photocrosslinks formation and separation

5′-^{32}P-labelled thiolated substrate analogues were dissolved at defined concentrations in the cleavage buffer. The solution (5 µl) was then introduced in a siliconized glass capillary and placed at 2 cm from the exit slit of a Bausch and Lomb monochromator equipped with a HBO 200W super pressure mercury lamp and a Schott WG 345 nm filter. The sample was irradiated for 45 min (100 kJ/m^2) i.e. to completion of the reaction. The crosslinked and uncrosslinked oligomers were then separated by 15% PAGE.

The same general procedure was applied to ribozyme–substrate analogue complexes. Each complex was formed by incubating 100 nM 5′-^{32}P-labelled ribozyme with 2 µM cold substrate analogue in the cleavage buffer at 4°C. Irradiation was performed at 15°C and the crosslinked forms of the ribozyme were separated from uncrosslinked ribozyme by 12% PAGE. The crosslinked species were detected by autoradiography quantified using a phosphorimager (Molecular Dynamics), eluted and precipitated with ethanol.

Identification of ribozyme crosslinked residues

Limited alkaline hydrolysis of end-labelled intact or crosslinked species was carried out in 50 mM carbonate buffer pH 9.2 in the presence of tRNA (0.5 mg/ml). Each sample (5 µl) was heated at 100°C for 1 min, cooled and loaded onto a sequencing gel (10% PAGE) at 60 W for 1.5 h. The correspondence between the individual bands of the ladder obtained with the intact ribozyme and its sequence was established using limited T1 RNase digestion.

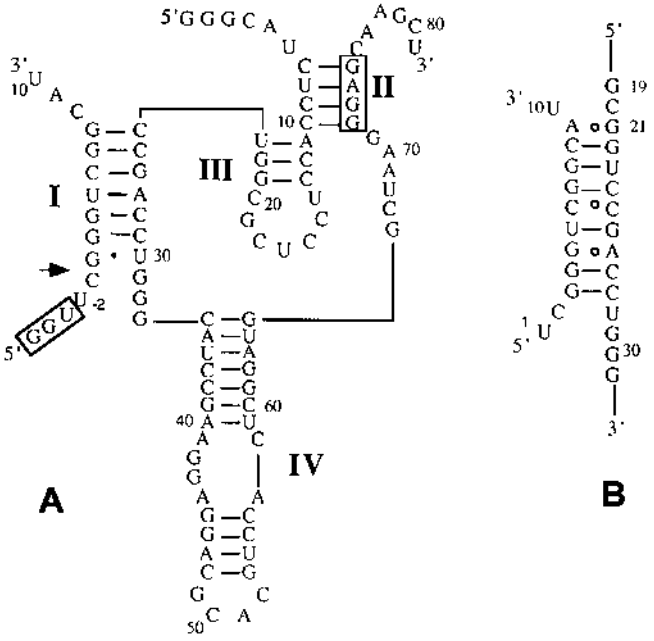

Figure 1. (**A**) The two-dimensional structure model of the RzANT–SANT complex as proposed by Perrotta *et al.* (8). The different domains are designated by roman numerals, the arrow indicates the cleavage site. Boxed sequences are base-paired in the axehead model (32). (**B**) Alternative hybridization of RzANT–SANT.

Native gel electrophoresis

5′-^{32}P-labelled substrate analogues and control oligonucleotides pre-incubated in the cleavage buffer for 15 min at 55°C were analysed by a non-denaturating 15% PAGE in the presence of 10 mM Mg^{2+}, 50 mM Tris–HCl, pH 8. Samples were run for 15 h at 11 W at various temperatures (see Results).

Three-dimensional modeling

The three-dimensional modeling was based on the structure proposed by Tanner *et al.* (15) for the genomic ribozyme. Coordinates were kindly provided by E. Westhof. Structural elements were manipulated on a PS 390 using SYBYL software package (Tripos, St Louis). At various steps, the results of the manipulations were energy minimized to satisfy stereochemical constraints. Coordinates are available from the authors by e-mail to laugaa@ccr.jussieu.fr.

RESULTS

Trans-cleavage of a RNA substrate and of a deoxy-substrate analogue by the antigenomic derived HDV ribozyme (RzANT)

The sequences of the 15mer RNA substrate (SANT) and of the corresponding *trans*-acting 81mer ribozyme (RzANT) are shown in Figure 1. The cleavage reaction was performed at a constant initial substrate concentration (0.1 μM) and at three different ribozyme to substrate ratios, *r*. The half-reaction times at 55°C were found to be respectively 0.9, 3 and 120 min at *r* values 10, 1 and 0.1, respectively. The cleavage reaction is less efficient at

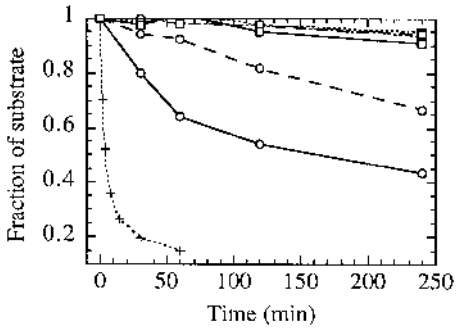

Figure 2. +, Cleavage of SANT by RzANT at 37°C and with *r* = 10. ○, Cleavage of S′ by RzANT at 37°C; plain line, *r* = 10; broken line - - - , *r* = 1; broken line ----, *r* = 0.1. □, Cleavage of S′ by RzANT at 55°C; plain line, *r* = 10; broken line - - -, *r* = 1; broken line ----, *r* = 0.1.

37°C since the half-reaction time was close to 5 min at *r* = 10. A turn-over was observed only at 55°C and in our conditions (*r* = 0.1) 3.5 molecules of SANT were cleaved per molecule of RzANT, after 1 h incubation. These data are in line with those previously reported by Perrotta and Been (14). Under the conditions used here, RzANT was far more active than its genomic counterpart in the *trans*-cleavage reaction (data not shown). Therefore, one should expect RzANT to be a more valuable tool for structural analysis of the active form of the HDV catalytic RNA.

In order to examine whether a deoxy substrate analogue can be catalytically cleaved we have synthesized a DNA oligomer S′ which contains a ribocytidine at the cleavage site. S′ is cleaved in the presence of RzANT at both 37 and 55°C (Fig. 2) generating a 5′-^{32}P-labelled product that co-migrates with the 5′ fragment (7mer) obtained upon alkaline hydrolysis of S′. As expected, the *trans*-cleavage reaction is abolished when 50 mM EDTA is added to the ribozyme prior to S′ addition. Perrotta and Been (14) have shown that a deoxycytidine 5′ to the cleavage site blocked the reaction. Taken together, these data demonstrate for the first time that the presence of a 2′-hydroxyl group i.e. a ribonucleotide at position –1 of the substrate is strictly required for the catalytic cleavage by RzANT as previously established in the case of hairpin and hammerhead ribozymes (27–26). The cleavage of S′ is highly sensitive to temperature; the substrate is apparently not stably associated with the RzANT at 55°C, which dramatically reduces the rate of the reaction. Consistent with such a view, at 55 or 37°C no competitive inhibition of SANT cleavage in the presence of large excess of a deoxy substrate analogue ds (up to 0.1 mM) was detected. Under the same conditions the 10mer product of SANT cleavage behaved as an inhibitor (data not shown). It can thus be concluded that S′ is able to bind to the correct site on RzANT albeit with a low affinity.

Photocrosslinking between substrate analogues and RzANT

In order to unravel long-range contacts within the ribozyme domain (Fig. 1) we first examined the behaviour of a series of photoactivable deoxysubstrate analogues and then extended this study to a few selected mixed ribo-deoxy oligomers. The full deoxy analogues used here are derived from the 12mer analogue ds which is expected to form a complete hybrid helix H1 when combined with RzANT. Each dsi contains a single deoxy-4-thiouridine residue, ds^{4}U, at a position occupied either

Table 1. Yield of formation and localization of the crosslinks

Substrate analogues	Position of crosslink	Yield (%)
Controls		
ds 5′ tcgggtcggcat 3′ dc 5′ ggxcgctcatgg 3′	No crosslink detected	
Uncleavable analogues		
ds4 5′ tcgggxcggcat 3′	No crosslink detected	
ds5 5′ tcgggtxggcat 3′		
ds8 5′ tcgggtcggxat 3′		
ds10* 5′ tcgggtcggcaxa 3′	C7 G21 U23	0.5 0.5 1
ds-1 5′ xgggtcggcat 3′	C15 U23 G31	1.5 4 0.5
ds-2 5′ xcgggtcggcat 3′	G19 U23	3.5 4.5
s″-2 5′ ccaaaxcGGGUCGG 3′	C15 G19 G31	6.5 3.6 3.8
Cleavable analogue		
s‴-2 5′ ccaaaxCGGGUCGG 3′	C67	5

dsi, full deoxy substrate analogues containing the ds^{4}U probe (X) at position i relative to the putative cleavage site.
s″-2 and s‴-2, mixed ribo-deoxy substrate analogues with a, t, c and g standing for deoxyribonucleotides and U, C, G for ribonucleotides.
dc, control ds^{4}U containing deoxyoligomer unable to bind to RzANT and to adopt a hairpin structure.
ds, unsubstituted deoxy substrate analogue.
*, notice that ds10 was extended at the 3′-end by one a residue to allow its chemical synthesis.

by U (–2, +4, +10) or by C (–1, +5, +8) in ds (Table 1). The substituted positions were selected in order to explore extensively the long range interactions within the dsi/RzANT complex and to minimally affect the stability of the helix I which is common to all the proposed secondary structures. Indeed, residues at positions –2, –1, +8 and +10 are not involved in base pairs (Fig. 1). Substitutions at position +4 and +5 will introduce either a ds^{4}U–A and or wobble ds^{4}U–G base pair in H1, respectively, with unchanged (18) or a decreased stability of this helix. In order to establish appropriate conditions for the *trans*-photocrosslinking experiments various concentrations of the ds-2 analogue, ranging from 100 nM to 2 µM, were mixed with 100 nM 5′-^{32}P-labelled RzANT in the cleavage buffer at 15°C. The mixtures were irradiated with 365 nm light at 15°C. Each sample was then run on a denaturing electrophoresis gel to separate the retarded crosslinked species from the unreacted RzANT. The fraction of RzANT crosslinked to ds-2 increased in parallel with ds-2 concentration reaching 8% at 2 µM. While the same crosslink patterns were obtained at 15 and 25°C, the overall crosslinking yield appears sensitive to the temperature at which the irradiation step is performed. It decreases with an increasing temperature and for ds-2 at temperatures >37°C no crosslinks could be detected. Further experiments were thus performed at 15°C in the presence of 2 µM dsi since under these conditions the control thiolated 12mer dc, which is unable to anneal to RzANT, yielded no detectable photocrosslink (Table 1).

As shown in Table 1, some dsi oligomers (ds10, ds-1, ds-2) yielded two to three crosslinked species (Fig. 3A) while others (ds4, ds5, ds8) generated no products. Some crosslinks were formed in low yield however, ≤1%, and could possibly have resulted from side photoreactions not involving the ds^{4}U probe. To check this point, the unsubstituted deoxysubstrate ds was irradiated together with 5′-^{32}P-labelled RzANT and no crosslink could be detected under these conditions (Table 1). Hence the covalent bridges analysed below are necessarily formed between ds^{4}U and a ribozyme residue. In order to determine the nature of this residue in crosslinked species each of them was isolated, repurified by denaturing PAGE if necessary and sequenced using limited alkaline hydrolysis. A control ladder was obtained starting from 5′-^{32}P-labelled RzANT and scaled using limited RNases digestions. When RzANT was crosslinked at position Y, random cuts 5′ to the crosslink generated a partial ladder that paralleled the control one, while all the cuts 3′ to the crosslink resulted in bands with retarded migration due to the attached substrate analogue. This led to a gap in the alkali ladder thus allowing an unambiguous identification of Y (20–21) (Fig. 3B).

In order to identify long-range contacts nearby the cleavage site in a full RNA helix I, we have synthesized two mixed RNA–DNA substrate analogues: s″-2 and s‴-2. The ds^{4}U residue was inserted at position –2 shown to be reactive in the case of full DNA analogues. The 5′ moiety of these oligomers is a DNA segment added to facilitate the localization of the crosslinks, while their RNA 3′ moiety can anneal to the corresponding complementary region of RzANT and yield H1. s″-2 contains a deoxycytidine at position –1 and, accordingly, is resistant to RzANT cleavage, while s‴-2 harbors a ribocytidine at this position and is readily cleaved at both 37 and 55°C. For example, the initial rate of cleavage of s‴-2 (1 µM) measured at 55°C in the presence of 0.1 µM RzANT was found to be 30–50% slower than that observed with SANT under the same conditions. When 2 µM s″-2 was irradiated in the presence of 0.1 µM RzANT in the cleavage buffer, three crosslinks were formed. The same three crosslinks were found whatever the irradiation temperature, however the overall yield of photocrosslinking was found to be 14% at 15°C, 4% at 37°C and 2% at 55°C. The crosslinks species were separated and sequenced as described above (Table 1). In order to avoid cleavage during the irradiation step, s‴-2 was combined with RzANT in the cleavage buffer in which Mg^{2+} was replaced by Na^+ (1 M). Under these conditions a single crosslink involving residue C67 was formed (Table 1). When s″-2 was irradiated under the same conditions, it yielded the three crosslinks observed in the standard cleavage buffer albeit with a lower yield (2%). It was of interest to check whether the s‴-2-RzANT crosslinked complex retained some cleavage activity. No evidence for the formation of the expected 5′-labelled 6mer product was obtained, however, when this complex was incubated at 37 or 55°C in the standard buffer (data not shown).

Folding of SANT′ and of its ds-2 under cleavage conditions

A full comparison of the behaviour of the RNA and DNA substrate analogues in the presence of RzANT should take into account their eventual structured forms which, for short oligonucleotides, can be either hairpins or dimers (28). Thus, we first compared by native gel electrophoresis (pH 8, 10 mM Mg^{2+}) the migration rates of SANT′ and of a control C, unable to fold as a

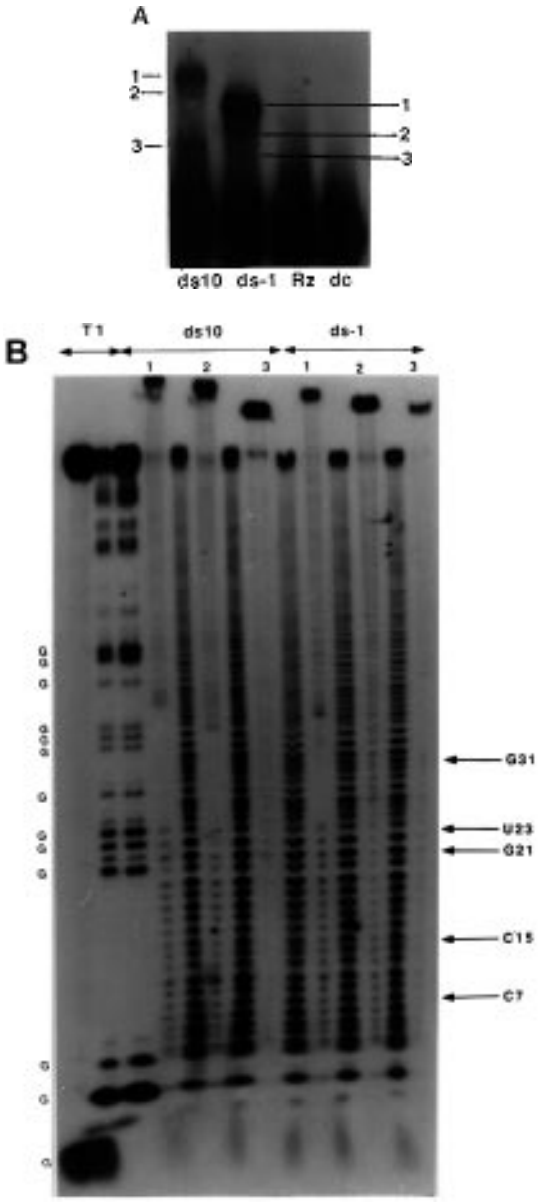

Figure 3. (**A**) Separation on 15% denaturing PAGE of the crosslinked species, noted 1, 2 and 3, formed upon irradiation of 100 nM 5′-^{32}P-labelled RzANT with 2 μM of either ds10 or ds-1 in the cleavage buffer at 15°C. Lane Rz, ribozyme irradiated alone. Lane dc, ribozyme irradiated in the presence of the control thiolated oligomer dc (2 μM). (**B**) Localisation of the RzANT crosslinked residues in the crosslinked species (1, 2 and 3) obtained respectively with the RzANT/ds10 and RzANT/ds-1 species. Lane T1, limited digestion of 5′-^{32}P-labelled RzANT with RNase T1 showing the G positions on the gel. The arrows indicate the ribozyme nucleotides immediately 5′ to the crosslinked residues.

hairpin or to give rise to self-associated forms. The gel was run at 37°C and irrespective of the applied concentrations (5 nM to 1 μM) SANT′ and C yielded a single spot that migrates at closely similar rates indicating that SANT′ is devoid of any kind of secondary structure.

We then compared, by the same procedure, the behaviour of ds-2 with that of a control 12mer dc of similar nucleotide composition (Table 1). Again, dc was designed in order to minimize its ability to fold as a hairpin or to self-associate. Figure 4A illustrates the data obtained at 55°C. ds-2 yielded a minor spot comigrating with dc and a major faster migrating spot. The ratio of counts in the two bands was found to be constant, 11 ± 2%, irrespective of the applied ds-2 concentration. The data therefore strongly suggest that at 55°C, ds-2 is in equilibrium between an unfolded form and a major fast migrating hairpin form. The existence of this hairpin form was confirmed by the photoaffinity methodology (28). 5′-^{32}P-labelled ds-2 was irradiated to completion of the reaction with 365 nm light at 15°C in the cleavage buffer and the irradiated mixtures subjected to denaturing PAGE analysis. At ds-2 concentrations ≤1 μM only two species were detected on the gel (Fig. 4B). One of them behaves as a single-stranded 12mer co-migrating with irradiated dc or non-irradiated ds-2 (not shown) and is expected to occur from unproductive ds^{4}U photolysis. The major fast migrating band corresponds to a crosslinked hairpin since: (i) it adopts a compact structure in denaturing conditions; (ii) its yield of formation, 70 ± 3%, remains constant whatever the initial ds-2 concentration <1 μM. When ds-2, 1 μM, was irradiated at different temperatures the yield of formation of this crosslinked hairpin remained practically constant, up to 55°C (data not shown), again indicating that ds-2 is predominantly in the hairpin form at this temperature. Irradiation of 5 or 10 μM solution of ds-2 at 15°C generated new strongly retarded species on the gel (Fig. 4B) that formed in competition with the crosslinked hairpin (Fig. 4C). These species certainly correspond to crosslinked ds-2 dimers as shown by their migration behaviour and their concentration dependent formation. Thus ds-2 is able to self-associate as a duplex form under cleavage conditions.

The contrasted behaviour of SANT′ and ds-2 led us to examine whether a bipartite mixed DNA–RNA oligomer s′′-2 could fold as a hairpin. After irradiation at 15°C no evidence could be found by denaturing PAGE for a fast migrating species. However, s′′-2 retained the ability to self-dimerize since a crosslinked dimeric species formed in low yield (1–2 %) at the highest concentration checked (10 μM).

DISCUSSION

Stability of the ribozyme–substrate (analogues) complexes

The data of Figure 4 indicate that ds-2 either folds as a hairpin or self-associates to yield imperfect dimeric forms. As the hairpin form (unlike the dimers) is stable up to 55°C, it should largely predominate over unfolded ds-2 at lower temperatures. At 15°C it can thus be considered that the dimeric and hairpin forms are in equilibrium with an apparent dissociation constant, K_{ap} (~10 μM), derived from the data of Figure 4C according to Tanaka *et al.* (28). The structures proposed for the hairpin (H, H′) and the dimeric forms (D, D′) of ds-2 are shown in Figure 5. A single crosslinked hairpin species can be detected (Fig. 4B) and we favor form H as its stem is far more stable that the corresponding stem in H′. Both dimeric forms D and D′ contain several mismatched base-pairs of type g–t, g–g, g–a that are known to slightly destabilize a duplex (29). The presence in D of three mismatched pairs and two couples of alternative g–c, c–g pairs, nevertheless, favors this form. As the photochemical data (Fig. 4B) indicate the formation of multiple crosslinked dimers, it cannot be excluded that D′ may be present in low relative amounts. In contrast to the behaviour of the dsi family, both SANT′ and s′′-2, which contains a ribonucleotide sequence potentially involved in the hairpin H, revealed unable to adopt

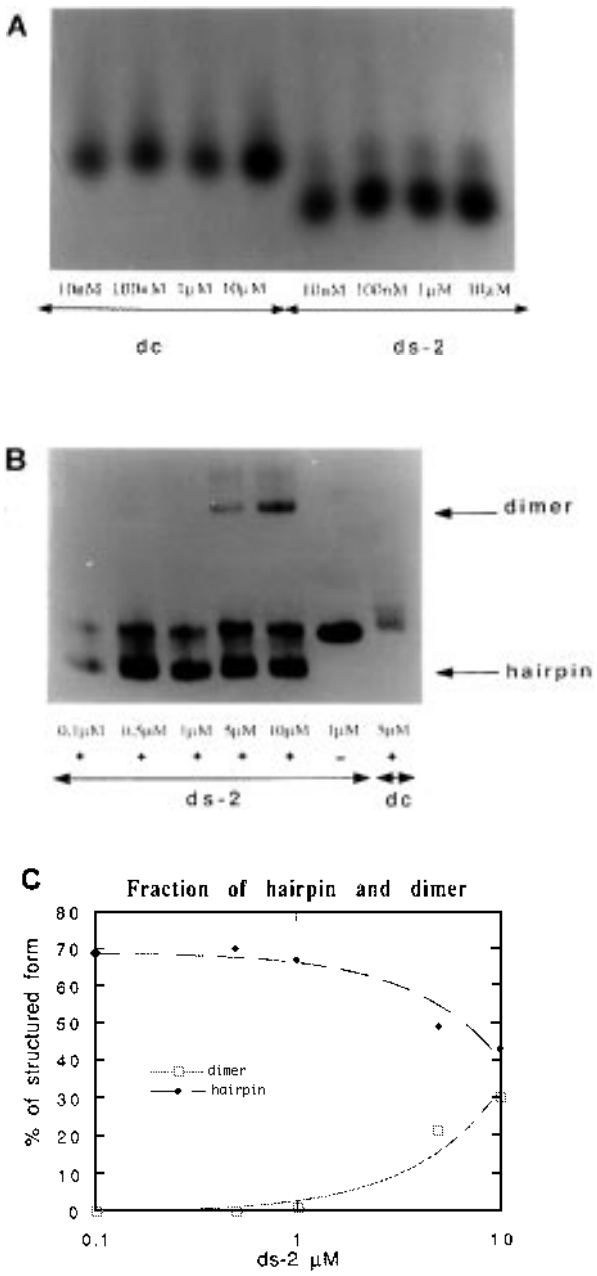

Figure 4. (**A**) Native gel electrophoresis of ds-2 at 55°C in the presence of 10 mM Mg^{2+}. dc is a control 12mer unable to adopt hairpin structure. (**B**) Denaturating gel electrophoresis of the 12mer substrate analogue ds-2, either irradiated (+) in the cleavage buffer (15°C) at different concentrations or not irradiated (–). dc is a control thiolated 12mer irradiated in the same conditions. (**C**) Yield of crosslinked species as a function of the concentration of the substrate analogue ds-2: ◇, hairpin form, □, dimer form.

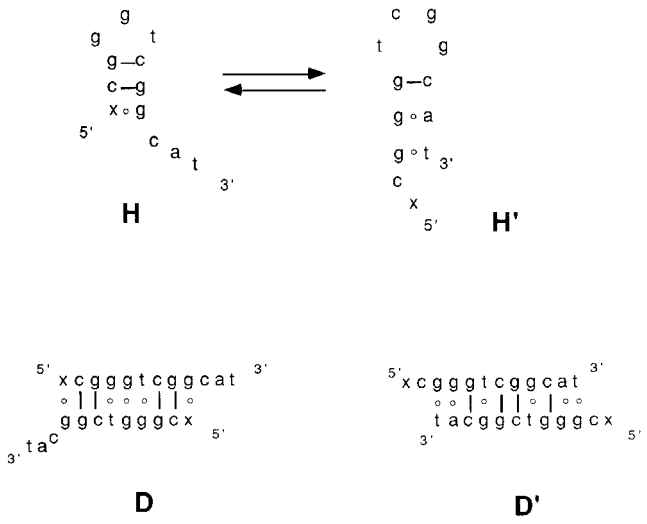

Figure 5. Plausible hairpin (H, H′) and dimeric (D, D′) forms adopted in solution by ds-2 with X standing for ds^{4}U. The symbols – and o indicate Watson–Crick or mismatched base-pairs formation, respectively.

such a structure. They behave as unfolded oligomers retaining the potential to form dimeric structures at 15°C.

SANT and S′ are cleaved by RzANT at 37 and 55°C demonstrating that these oligomers are able to form the correct catalytic structure implying helix H1. Both the cleavage kinetics and crosslinking data, however, show that the deoxy substrates have a much lower affinity to RzANT than their RNA counterparts. Cleavage of S′ is faster at 37°C than at 55°C, in contrast to SANT cleavage. The DNA substrate analogue, ds, is not a competitive inhibitor in the cleavage reaction in contrast to the 10mer product of SANT cleavage. ds-2-RzANT crosslinks are not formed at temperatures >37°C while under the same experimental conditions s″-2 yields crosslinks up to 55°C. The situation observed here is very similar to the one encountered with the (–)TRSV hairpin system, where deoxy substrates analogues were shown to bind the ribozyme according to the consensus structure but with low affinity (28). Obviously this is due primarily to the lower stability of hybrid DNA–RNA minihelices as compared to the corresponding RNA–RNA duplexes (30–31). In addition the DNA substrate analogues fold as stable hairpins, thus disfavouring their binding to the ribozyme.

A three-dimensional model for the antigenomic ribozyme

Recently, three-dimensional models have been proposed for the genomic pseudoknot (15) and both antigenomic pseudoknot and axehead (32) ribozymes. A basic difference between the pseudoknot and axehead lies in the Watson–Crick base-pairing involving the ribozyme 3′-end sequence which belongs to helix II in the former while it forms a helix with residues 5′ to position –2 in the latter (Fig. 1). The models of Branch and Polaskova (32) were designed after co-mutation data interpreted in terms of base-pairing or proximity between couples of residues, though no upper limit was set to define proximity. They claimed that their data were largely compatible with the model of Tanner *et al.* (15) but it should however be noticed that, in this model, distance measurements between C1′ atoms yielded values ranging from 20 to 33 Å for four out of their eight sets of nearby residues. As neither coordinates, nor details concerning software and building methodology were given, the models of Branch and Polaskova (32) are inadequate for structural purpose and precludes any comparison with our photocrosslinking data. It is questionable whether the *trans*-acting ribozyme constructions described here could adopt the axehead structure. None of the substrate

analogues (cleavable or not) used here have a 5′-end sequence able to form a Watson–Crick minihelix with the 5′-GGAG sequence at the 3′-end of the ribozyme as required for the axehead. It is also noteworthy that position –1 yielded a crosslink with position 31, in agreement with the pseudoknot model, but none with positions 70–72 which would have supported the axehead. With regards to our system, the pseudoknot secondary structure appears to be the more likely.

It cannot be excluded, however, that alternative hybridization occurs. For example, the crosslink between ds10 and G21 is best explained by the structure depicted in Figure 1B, whereas those between s″-2, ds-1 and G31 can be accounted for by both. The lack of reactivity of ds4 and ds5 is expected since, in helix H1, the ds^{4}U probe form standard (ds4) or wobble base pair (ds5) and the minihelix geometry and flexibility disfavor inter-strand photocrosslink formation as observed in a number of model systems (unpublished data). The long range crosslinks obtained with ds^{4}U at position –1 or –2 of the substrate indicate a tertiary folding of the pseudoknot structure placing C67, G19 and C15 in close proximity to the cleavage site. The antigenomic and genomic ribozymes can adopt similar pseudoknotted secondary structures (23). Some residues, located in the same secondary patterns, are strictly conserved in both ribozymes: G31, G32 and G33 in junction I/IV; G66, C67, A69 and A70 in junction II/IV; U14–C20 in loop III (antigenomic numbering). Mutational studies performed on the genomic ribozyme have shown that substitution of any of these residues strongly decreases the cleavage efficiency and that the residue corresponding to C67 cannot be changed to any other nucleotide. On this basis, a three-dimensional model has been proposed for the genomic ribozyme by Tanner *et al.* (15). In this model the residue of the cleavage site is within 12 Å of the residues corresponding to C67 and C15. Given these similarities, both ribozymes are expected to have very closely related three-dimensional folding. We therefore derived a model for the antigenomic ribozyme from the structure proposed for the genomic one, assuming that the conserved aforementioned residues have similar spatial arrangement.

Apart from the sequence, the main differences between both ribozymes result from either deletion of non-conserved residues in loop III and junction I/IV or insertion of a residue at the 3′-end of junction II/IV. Compared to the genomic model, the antigenomic construction (Fig. 6) is as follows, (i) Loop III is one nucleotide shorter, which results from the deletion of the dispensable U residue at the 3′-end of the genomic loop (33). The 5′ side of the loop which contains catalytically important residues is left unchanged whereas, on the 3′ side, G19 is now 4 Å closer to the cleavage site. (ii) Junction I/IV is three nucleotides long (GGG) instead of six (GGGCAA). The gap formed by the deletion of the CAA trinucleotide is now filled by G32 and G33 which adopt an unwinded A-type stacking between helix I and helix IV. The relative spatial disposition of helix I with respect to helix IV remains almost unchanged. It should be noticed that residues G32 and G33 are stacked between helices I and IV whereas they are tilted and point outwards in the genomic ribozyme. (iii) In order to accommodate the additional guanine at the 3′-end of junction II/IV, helix II has been slightly tilted (~15°) with respect to helix III via a +20° rotation about the ε angle of C10, at the junction between both helices. Therefore, the remainder of junction II/IV which contains residues important for cleavage is left unchanged. Starting from its 5′-end, the conformational path of the junction can be seen as a right handed helical extension of helix IV, the first three residues yielding close contacts between their sugars and helix/loop III. Their bases point towards junction I/IV and helix I. There is a turn 3′ to the third (variable) residue, induced by steric repulsion with helix I which induces the last residues to wrap around helix III in a right helical way, therefore insuring continuity with helix II. Such a way of incorporating the extra G is devoid of effect on the relative positions of the residues important for cleavage and induces only minor effects on the structure of junction I/II of *cis*-cleaving ribozymes. This is consistent with data showing that a genomic ribozyme in which junction II/IV has been replaced by the antigenomic one is still active (Lescure, unpublished results).

Model fitting to photocrosslinking data

While cleavage experiments were performed at 37 or 55°C, most of our crosslinking data were obtained at 15°C. In a few cases ds-2 and s″-2, we have shown however that the crosslink pattern was largely temperature independent (up to 55°C for s″-2). Formation of a photoadduct between ds^{4}U and an acceptor residue implies that the two reactants come in close proximity, <4.5 Å, within the complex (18).

Photocrosslink with s‴-2 and s″-2. Both s‴-2 and s″-2 form an A-type RNA–RNA helix I. In the model the residue at position –1 is closely surrounded by residues G1, G31–G33 (junction I/IV) and residues U14–C15 (loop III). The strand 5′ to C-1 (dc-1) is constrained to go out either on loop III side, at the back of the molecule or on junction I/IV side at the front of the molecule. The presence of a C3′-endo ribocytidine in s‴-2 orients its 5′-phosphate on the front of the molecule. The residue at position –2 slips in between G33 on the one side and C67, G66 on the other side and the remainder of the 5′ strand lies along the major groove of helix IV. These steric constraints bring residue –2 in close proximity to C67. The carbon–sulfur bond of ds^{4}U is parallel to the C5–C6 double bond of C67 and at a distance of 3.8 Å which nicely accounts for the observed crosslink (Fig. 6, top). Furthermore, none of the other residues is liable to form a photocrosslink. Interestingly, the Rz/s‴-2 complex becomes inactive upon formation of the –2 to C67 crosslink which could well be due to the alteration of the chemical structure and spatial disposition of C67. This is in line with the statement discussed above that C67 is a key residue of the antigenomic catalytic site. The non-cleavable substrate analogue s″-2 differs from s‴-2 only by the presence of a deoxy residue at the cleavage site. It however yields a distinct crosslink pattern which involves C15 and G19 in loop III, at the back of the molecule, whether in Mg^{2+} or high Na^+ containing media. Its behaviour is necessary due to the presence of a deoxy sugar at position –1: C′2-endo preferred conformation, greater flexibility and absence of 2′-hydroxyl. The model shows that, when the residue of the cleavage site is a deoxyribonucleotide with a C2′ endo conformation its 5′-phosphate is in a less favorable configuration to position the 5′-end of the substrate at the front of the molecule. The substrate 5′-end is easily built running on loop III side, at the back of the molecule (Fig. 6, bottom). Residue –2 now points towards residues C15 and G19. Both photoreacting bonds are sterically accessible to residue –2 and are at a distance of 6–7 Å. A slight rotation (+20°) of helix and loop III was performed to bring these bonds in a more favorable orientation. The rotation axis was perpendicular to helix III average bases plane, and passed through the center of helix III atomic coordinates. The distance between ds^{4}U-2 and

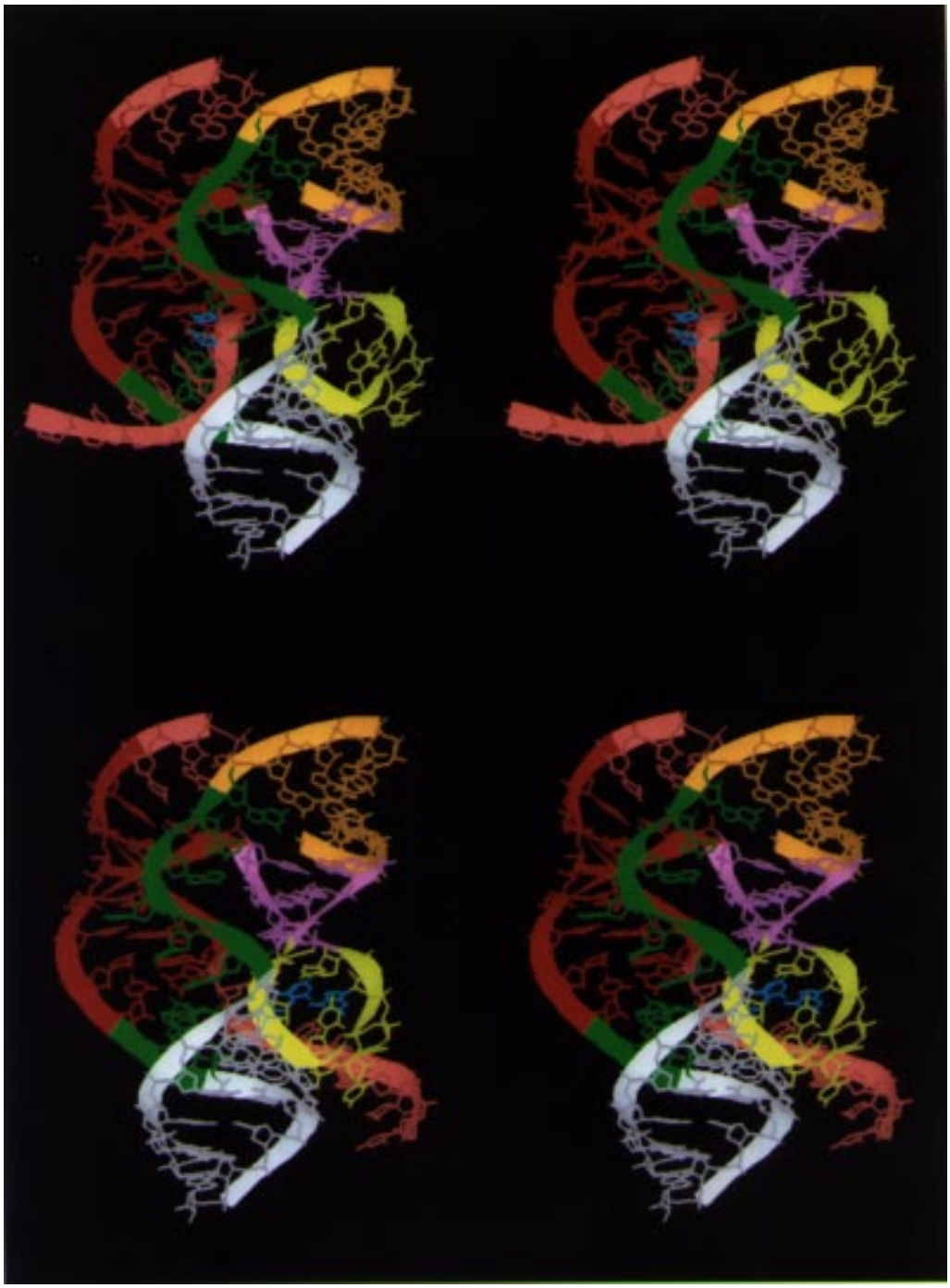

Figure 6. Stereo views of the antigenomic HDV ribozyme. The color code is the same as in reference 15 to allow comparison. Helices I, II, III and IV are drawn red, orange, violet and grey, respectively. Loop III is drawn yellow and junctions I/IV and II/IV are drawn green. The substrate 5′ and 3′-end are drawn pink. (**Top**) Front view of the ribozyme complexed with a substrate containing a ribonucleotide at the cleavage site (s″-2). Residue ds^{4}U-2 and the conserved residue C67 are drawn cyan. (**Bottom**) Front view of the ribozyme complexed with a substrate containing a deoxy residue at the cleavage site (s″-2). Residue ds^{4}U-2 and the conserved residues C15, G19 are drawn cyan. The drawing is produced by RASMOL2 (R. Sayle, rasmol@ggr.co.uk).

G19 photoreacting bonds average to ~4 Å with an antiparallel orientation which accounts for the photocrosslink detected. The distance between ds^{4}U-2 and C15 is larger (8 Å) but a distance as low as 4 Å can be reached without steric hindrance, at the expense of that between G19. Thermal motions of the substrate 5′-end and of loop III may well account for such behaviour. Therefore, both photocrosslinks are in agreement with the proposed three-dimensional structure (Fig. 6).

Photocrosslinking with ds-2 and ds-1. ds-2 is made of deoxyribonucleotides and contains only one residue 5′ to the cleavage site. A deoxy substrate is not expected to induce structural modifica-

tions of the ribozyme since it has been shown by NMR that DNA–RNA hybrids retain most of the structural features of A-type RNA helices (34). Furthermore, the two crystal structures of the hammerhead ribozyme are almost identical though one of them has been performed with a deoxy substrate (35–36). A photocrosslink is obtained with G19, as in the case of s″-2. Besides, the crosslink with U23 may well be explained on the basis of thermal motions. Indeed, the lack of any nucleotide 5′ to residue –2 is likely to increase its mobility, hence it can reach U23 which is on the same side of the ribozyme and accessible at the top of helix III. Furthermore, U23, located 5′ to a sharp turn, is expected to experience a large degree of fraying. Unlike to s″-2, no photocrosslink is observed with C15 which might arise from the larger movement of residue –2 between G19 and U23. The ds-1 to C15 photocrosslink is well accounted for by the structure. The photocrosslink observed with U23 can be explained by the mobility of the short 5′-end of the substrate, including G1, allowing a back folding which brings residue –1 in the vicinity of U23.

ds10 photocrosslinking with C7, G21 and U23. Residue 10 is the penultimate residue of the 3′-end of the substrate. This residue located on a dangling strand is expected to have a large degree of freedom. The occurrence of photocrosslinks with C7 and U23 which are both sterically accessible is thus fully compatible with the three-dimensional structure of the ribozyme. Besides, residue G21, located at the junction between helix III and loop III, is much less accessible and too far from residue 10 to yield a photocrosslink. This adduct might be explained by an alternative co-folding of the ribozyme and the substrate as shown in Figure 1B.

The herein proposed model for the antigenomic HDV ribozyme accounts for the above photocrosslinking data. Apart from residue deletion or insertion, the only minor difference with the previously described model of the genomic HDV ribozyme (15), resides in a slight rotation of helix and loop III with respect to helix II and helix I. It is noteworthy that, for the first time, the model of Tanner *et al.* (15) as well as the pseudoknot secondary structure receives a strong experimental support due to the possibility to obtain information on inter-residue distances.

ACKNOWLEDGEMENTS

We thank Kyle Tanner and Ilya Pashev for constructive comments on the manuscript, Prof. Marc Vasseur for his constant interest in this work and GENSET (France) for its kind gift of synthetic oligonucleotides. This work was partly sponsored by grants from ANRS and Interface Chimie-Biologie (CNRS) to A.F. C.B. was supported by a fellowship from Conseil Général d'Ile de France, GENSET France and Institut de Formation Supérieure Biomédicale.

REFERENCES

1 Wang,K.S., Choo,Q.L. and Weiner,A.L. (1986) *Nature* **323**, 508–514.
2 Lai,M.M.C. (1995) *Annu. Rev. Biochem.* **64**, 251–286.
3 Wu,H.N., Lin,Y.J. and Lin,F.P. (1989) *Proc. Natl Acad. Sci. USA* **86**, 1831–1835.
4 Sharmen,L., Kuo,M.Y.P. and Dinter-Gottlieb,G. (1988) *J. Virol.* **62**, 2674–2679.
5 Kuo,M.Y.P., Sharmeen,L., Dinter-Gottlieb,G. and Taylor,J. (1988) *J. Virol.* **62**, 4439–4444.
6 Perrota,A.T. and Been,M.D.(1990) *Nucleic Acids Res.* **18**, 6821–6827.
7 Rosenstein,S. P. and Been,M.D. (1991) *Nucleic Acids Res.* **19**, 5409–5416.
8 Belinsky,M. and Dinter-Gottlieb,G. (1991) *Nucleic Acids Res.* **19**, 559–564.
9 Perotta,A.T. and Been,M.D.(1991) *Nature* **350**, 434–436.
10 Branch,A.D. and Robertson,H.D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10163–10167.
11 Thill,G., Blumenfed,M., Lescure,F. and Vasseur,M., (1991) *Nucleic Acids Res.* **23**, 6519–6525.
12 Suh,Y.A., Kumar,P.K.R., Nishikawa,F., Kayano,E., Nakai,S., Odai,O., Uesugi,S., Taira,K. and Nishikawa, S. (1992) *Nucleic Acids Res.* **20**, 747–753.
13 Wu,H.N. and Huang,Z.S. (1992) *Nucleic Acids Res.* **20**, 5937–5941.
14 Perrotta,A.T. and Been,M.D. (1992) *Biochemistry* **31**, 16–21.
15 Tanner,N.K., Scaff,S., Thill,G., Petit-Koskas,E., Crain-Denoyelle,A.M. and Westhof,E. (1994) *Current Biol.* **4**, 488–498.
16 Tanner,N.K. (1995) In Dinter-Gottlieb,G. (ed.), *The Unique Hepatitis Delta Virus.* R.G. Landes Company, pp. 11–29.
17 Clivio,P., Fourrey,J.L., Gashe,J., Audic,A., Favre,A., Perrin,C. and Woisard,A. (1992) *Tetrahedron Lett.* **33**, 65–68.
18 Favre,A. (1990) In Morrison (ed.) *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*, Inc. Wiley & Son, New York, pp. 379–425.
19 Favre,A., Lemaigre-Dubreuil,Y. and Fourrey,J.L. (1991) *New J. Chem.* **15**, 593–599.
20 Dos Santos,D., Vianna,A.L., Fourrey,J.L. and Favre,A. (1993) *Nucleic Acids Res.* **21**, 201–207.
21 Woisard,A., Fourrey,J.L., and Favre,A. (1994) *J. Mol. Biol.* **239**, 366–370.
22 Perrota,A.T. and Been,M.D. (1993) *Nucleic Acids Res.* **21**, 3959–3965.
23 Milligan,J.F. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
24 Milligan,J.F. and Uhlenbeck,O.C. (1989) *Methods in Enzymology* **180**, 51–62.
25 Igloï,G.L. (1988) *Biochemistry* **27**, 3842–3849.
26 Perreault,J.P. Wu,T., Cousineau,B., Oglivie,K.K. and Cedergren,R. (1990) *Nature* **344**, 565–567.
27 Chowrira,B.M. and Burke,J.M. (1991) *Biochemistry* **30**, 8518–8522.
28 Tanaka,M., Dos Santos,J., Laugâa,P., Fourrey,J.L. and Favre,A. (1995) *New J. Chem.* **19**, 469–478.
29 Ikuta,S, Takagi,K., Bruce Wallace,R. and Itakura,K. (1987) *Nucleic Acids Res.* **15**, 797–811.
30 Hall,K.B. and Mc Laughlin,L.W. (1991) *Biochemistry* **30**, 10606–10613.
31 Bevilacqua,P.C. and Turner,D.H. (1991) *Biochemistry* **30**, 10632–10640.
32 Branch,A.D. and Polaskova,J.A. (1995) *Nucleic Acids Res.* **23**, 4180–4189.
33 Thill,G., Vasseur,M. and Tanner,N.K. (1993) *Biochemistry* **32**, 4254–4262.
34 Salazar,M., Fedoroff,O.Y., Miller,J.M., Ribeiro,N.S. and Reid,B.R. (1993) *Biochemistry* **32**, 4207–4215.
35 Scott,W.G., Finch,J.T. and Klug,A. (1995) *Cell* **81**, 991–1002.
36 Pley,H.W., Flaherly,K.M. and McKay,D.B. (1994) *Nature* **372**, 68–74.