The structure of the isolated, central hairpin of the HDV antigenomic ribozyme: novel structural features and similarity of the loop in the ribozyme and free in solution

M.H.Kolk, H.A.Heus and C.W.Hilbers

Nijmegen SON Research Center for Molecular Structure, Design and Synthesis, Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

The structure of an RNA hairpin containing a seven-nucleotide loop that is present in the self-cleaving sequence of hepatitis delta virus antigenomic RNA was determined by high resolution NMR spectroscopy. The loop, which is composed of only one purine and six pyrimidines, has a suprisingly stable structure, mainly supported by sugar hydroxyl hydrogen bonds and base–base and base–phosphate stacking interactions. Compared with the structurally well-determined, seven-membered anticodon loop in tRNA, the sharp turn which affects the required 180° change in direction of the sugar–phosphate backbone in the loop is shifted one nucleotide in the 3' direction. This change in direction can be characterized as a reversed U-turn. It is expected that the reversed U-turn may be found frequently in other molecules as well. There is evidence for a new non-Watson–Crick UC base pair formed between the first and the last residue in the loop, while most of the other bases in the loop are pointing outwards making them accessible to solvent. From chemical modification, mutational and photocrosslinking studies, a similar picture develops for the structure of the hairpin in the active ribozyme indicating that the loop structure in the isolated hairpin and in the ribozyme is very similar.

Keywords: central loop/hepatitis delta virus/ribozyme/
RNA hairpin/solution structure

Introduction

Hepatitis delta virus (HDV) has a single-stranded circular RNA genome, which is believed to replicate through a rolling-circle mechanism which generates linear multimers. Single copies of the genomic and the antigenomic strand are generated through self-cleavage, and this ribozyme activity has been demonstrated *in vitro* for sequences that are largely homologous in both strands (Kuo *et al.*, 1988; Sharmeen *et al.*, 1988; Wu *et al.*, 1989; Macnaughton *et al.*, 1993). As in other ribozyme systems, this cleavage reaction requires the presence of divalent cations, generally Mg²⁺.

A secondary structure model for the genomic and antigenomic ribozymes has been proposed in which a pseudoknot structure is formed by four helical regions, three single-stranded regions and two hairpin loops (Rosenstein and Been, 1991). This model—depicted in Figure 1—has gained strong support by data from chemical probing experiments (Kumar *et al.*, 1994). Extensive

mutational studies have shown that parts of the *cis*-acting ribozyme are redundant and that double mutations in most helical regions do not affect ribozyme activity (Been *et al.*, 1992). There are several hot spots, however, in which no mutations are allowed without impairing cleavage activity. Residues in these regions are believed to be located near the cleavage site in the active ribozyme. They may provide the necessary tertiary interactions that are required to constrain the conformation around the active site. Based on mutational studies, Tanner and co-workers (Tanner *et al.*, 1994) have proposed a model for tertiary folding of the HDV ribozyme. In this model structure, all single-stranded regions that were shown to be indispensable for activity are positioned in close proximity to the cleavage site. The possible formation of such a tertiary core has been confirmed recently by photo-crosslinking experiments (Bravo *et al.*, 1996; Rosenstein and Been, 1996).

The central part of the antigenomic pseudoknot model consists of a hairpin loop (loop III in Figure 1A) with seven nucleotides which are all important for activity. Strikingly, the genomic model contains the same loop sequence, with the exception of one extra U, which is not required for cleavage (Kawakami *et al.*, 1993). *In vitro* selection experiments on a pool of randomized sequences of the genomic ribozyme yielded a molecule with a similar secondary structure and a central loop that was practically identical to loop III in the antigenomic ribozyme (Nishikawa *et al.*, 1996).

The composition of this loop is unlike other RNA loops for which structural data are available. RNA tetraloops, which are the most extensively studied of all RNA loops, very often achieve their structural stability through the formation of GU or GA base pairs (Woese *et al.*, 1990). This is also true for larger hairpin loops which generally have a considerable content of purines (Hoffman and White, 1995; Fountain *et al.*, 1996; Huang *et al.*, 1996). Alternatively, the HDV central loop contains six pyrimidines and only one purine. Mutational analyses have shown no indication of canonical or non-canonical base pairs within the loop. Moreover, it was reported that base pairing of any type between loop residues and single-stranded regions elsewhere in the molecule is very unlikely (Tanner *et al.*, 1994). The energy contribution of other possible external interactions concerning the loop are expected to be insufficient to alter its conformation significantly.

These considerations led us to believe that the central loop of the HDV ribozyme is structurally autonomous and that the isolated loop III will adopt a conformation that may be maintained in the ribozyme molecule. We have investigated the structure of a 19 nucleotide RNA hairpin (Figure 1B) containing the above mentioned loop sequence by means of high-resolution NMR spectroscopy, in order to investigate these hypotheses and to find out what

Fig. 1. (A) Secondary structure model for the antigenomic HDV ribozyme as proposed by Perotta and Been (1991). The site of cleavage is marked by an arrow; stem-and-junction regions are numbered by roman numerals. Invariant regions, essential for activity, are boxed. (B) The SLIII hairpin. The shaded region indicates the nucleotides that are identical in the ribozyme.

possible interactions stabilize the structure of the pyrimidine-rich loop. The results indicate a relatively well-defined structure for the RNA hairpin with some novel structural features. Furthermore, the loop structure matches the chemical modification and mutational studies conducted on the ribozyme, supporting the notion of a structurally autonomous loop III.

Results

Because the transcription reaction of the RNA yielded two distinct bands of about the same desired molecular weight, care was taken to ensure that the correct fragment was used for the NMR experiments. RNA sequencing of the selected band confirmed the correct sequence of 15 out of 19 residues. The identity of the last four residues at the 3'-end could not be determined by this method. However, a ^{13}C - ^1H HMQC NMR experiment showed that the selected fragment had the correct number of adenine H2s, purine H8s and uridine and cytidine H5s, so that the possibility that the incorrect fragment had been chosen can be excluded.

In the 3D model of the HDV ribozyme proposed by Westhof and co-workers (Tanner *et al.*, 1994) the central helix III is co-axially stacked upon helix II (Figure 1). We mimicked this colinear system in the 19-mer hairpin, hereafter referred to as stem-and-loop III (SLIII), so as to account for possible effects of the stem on the conformation of the loop.

Assignments

Assignment of the stem region was done using standard NMR methods (Wijmenga *et al.*, 1993). Sequential con-

nectivities were established for the entire hairpin following an anomeric to aromatic proton walk, except for the correlations between C9 and U10, and C11 and G12, but these residues could be interconnected via other sugar-base proton NOEs. The rest of the sugar spin systems could then be determined from TOCSY and NOESY ladders connected to the assigned H1's. All loop residues had observable TOCSY crosspeaks from H1' to H2', and for C9–13 even H1' to H3' and H4' correlations were observable. The sugar proton resonances that could not be assigned in the TOCSY spectra were identified by their NOE intensities to already assigned sugar protons—generally the well-dispersed H1's—and to base protons. This procedure resulted in the assignment of all sugar proton resonances, except for a few cases in which the resonance identification was hampered by overlap or line broadening (marked as 'tentative assignment' in Table I).

Stem imino proton resonances could be identified via imino-imino and imino-adenine H2 NOEs. The imino resonance of G1 is less intense than the others due to fraying at the helix end. Intrabase-pair crosspeaks connecting guanine-imino to H5 and H6 resonances of the cytidines, caused by spin diffusion via the C-amino group, agreed with the aforementioned assignments.

Most ^{31}P -assignments could be derived from a 2D ^{31}P - ^1H hetero-TOCSY-NOESY spectrum. During the TOCSY mixing period, coherence is transferred from the phosphorus atom of residue (*i*) to its own H5'/H5'' protons and to the H3' proton of residue (*i*-1). In the subsequent NOE mixing time there is an NOE buildup from the (*i*-1) H3' to the H1' proton of the same sugar. This pathway resulted in P(*i*)-H1'(*i*-1) connectivities for all residues but one. No crosspeaks were observed to H1'(*i*) protons, since these are located at a distance of ~5 Å away from the H5'/H5''(*i*) protons. The unambiguity of the correlations with the well-dispersed H1's makes this experiment a very useful tool in assigning ^{31}P -resonances, even when these are overlapping. For relatively isolated ^{31}P -resonances, assignments could be confirmed by connection to H3' and H5'/H5'' protons in the 2D HETCOR and hetero-TOCSY spectra.

Magnesium titration

As mentioned previously, HDV catalytic RNA has an absolute requirement for Mg^{2+} , but it is not known whether the cation is only required for the cleavage step or is also involved specifically in the stabilization of RNA tertiary interactions. To test the influence of Mg^{2+} on the loop conformation, we titrated MgCl_2 in 2.5 mM steps to the RNA sample and monitored the behaviour of the H5-H6 and H1'-H2' resonances in 100 ms TOCSY spectra (data not shown). Only very modest changes in chemical shifts were observed in the range 0–10 mM MgCl_2 , but all resonances broadened considerably. This indicates that the loop structure is quite independent of the presence of magnesium.

Structure determination

A total number of 192 intra- and 136 inter-residue NOEs were collected, of which 58 and 51 relate to loop residues, respectively. As expected, residues 1–6 and residues 14–19 form normal Watson-Crick base pairs, and for these regions all NMR data are in agreement with a normal

Table I. Chemical shifts of proton and phosphorus resonances relative to TSP in the SLIII hairpin

	H8/H6	H2/H5	H1'	H2'	H3'	H4'	H5'/H5''	im/am	³¹ P	J _{H1'H2'}
G1	8.11	n.a. ^a	5.78	4.92	4.73	4.51	4.25/4.12	13.01	−0.01	2
G2	7.62	n.a.	5.88	4.52	4.60	4.25	4.51/4.25	13.38	−1.88	<1
C3	7.71	5.29	5.50	4.48	4.56	4.43	4.12/4.11 ^b	8.61/6.96	−2.39	<1
A4	8.06	7.41	5.94	4.54	4.66	4.48	4.13/4.02		−2.14	<1
C5	7.48	5.23	5.37	4.21	4.31	4.37	4.46/4.03	8.45/7.04	−2.39	<1
C6	7.60	5.40	5.51	4.30	4.35	4.37	4.03/4.01	8.48/6.87	−2.48	<1
U7	7.67	5.46	5.57	4.33	4.34	4.44	4.06/4.04		−2.49	2
C8	7.72	5.66	5.56	4.18	4.43	4.31	4.06/3.99		−2.23	3
C9	7.86	6.00	5.76	4.28	4.27	4.48	4.16/4.00 ^b		−2.39	5
U10	7.69	5.80	5.76	4.23	4.47	4.29 ^b	3.98/3.94		−1.75	9
C11	7.61	5.82	5.64	4.24	4.56	4.21	3.91/3.89 ^b		−2.05	5
G12	7.92	n.a.	5.65	4.25	4.82	4.48	4.17/4.10 ^b		−2.06	4
C13	7.62	5.42	5.68	4.56	4.31	4.48	4.34/4.33 ^b		−1.50	2
G14	7.56	n.a.	5.65	4.47	4.32 ^b	4.34 ^b	4.14/4.13	12.75	−1.72	<1
G15	7.33	n.a.	5.80	4.46	4.54	4.07	4.06/4.05	13.31	−2.42	<1
U16	7.72	5.13	5.56	4.64	4.55	4.43	4.12/4.08	13.70	−2.59	<1
G17	7.72	n.a.	5.80	4.49	4.55	4.12 ^b	4.08/4.05	12.66	−2.10	<1
C18	7.66	5.23	5.45	4.22	4.41	4.36	4.06/4.06 ^b	8.61/6.96	−2.65	<1
C19	7.64	5.50	5.73	3.98	4.15	4.12	4.45/3.97	8.61/7.13 ^b	−2.34	<1

^aNot applicable.^bTentative assignment.The ³¹P-chemical shift was calibrated as described by Ikura *et al.* (1991), by multiplying the ¹H TSP frequency by 0.40480793.

RNA A-helix. Imino proton resonances are at the expected positions and imino–imino contacts are observed throughout the stem region. All G-imino to C-amino contacts are also indicative of regular Watson–Crick base pairs. Sequential H1', H2' and H3' to base connectivities as well as low-intensity base–base contacts can be seen for all stem residues. With the exception of G1, all stem residues have an N-type sugar pucker. Therefore we kept the stem residues 1–5 and 15–19 in a fixed helical conformation during all molecular dynamics (MD) simulations that were necessary for calculating the correct loop conformation. The base pair preceding the loop, C6G14, was allowed to move within the boundaries of the standard A-helical distance restraints to accommodate possible distortions imposed by the loop conformation. The final restrained molecular dynamics (RMD) run, in which all experimental data were used as restraints on a standard A-helix, yielded no violating structures, indicating that all stem NOEs agree with this conformation.

The loop region. Despite the availability of 51 inter-residue constraints for the loop region alone, the first RMD run, which started from randomized loop conformations, yielded an ensemble of 50 structures with 20 or more violations. A set of 10 structures with the lowest energies and the lowest number of violations was selected from this ensemble and subjected to the same simulated annealing (SA) protocol. This resulted in a family of structures with only 5–10 violations. Following yet another round of selection, no structures were obtained with violations >0.5 Å. It is interesting to see that RMD turned out to be unable to search conformational space efficiently enough to come up with structures fulfilling all experimental data using a limited amount of distance restraints in a single MD run of ~30 ps.

Figure 2 shows an overlay of a family of structures fulfilling all experimental restraints. Residues U7–C9 at the 5'-end of the loop are structurally best defined, but U10 and C13 also have little conformational freedom.

These five residues have a local root-mean-square deviation (r.m.s.d.) of 0.83 Å (with respect to the mean structure), but the overall r.m.s.d. for the loop is 1.02 Å as a result of the much poorer structural definition for residues C11 and G12. The ³¹P-chemical shifts for the first three residues at the 5'-side of the loop have A-helical values, and a considerable number of sequential base–base and sugar–base NOEs forces residues C6–C9 into a stacked orientation (see Figure 3). All regular A-helical connectivities are observable for this region, albeit with different intensities. In particular, the measured distances between H2'(i) and H5(i+1) for C6, U7 and C8 are ~4.2 Å, which is ~1 Å less than in a standard A-helix. This leads to an orientation in which the bases are positioned directly above each other, thereby reducing the helical twist for these residues practically to zero. This pattern is disrupted between C9 and U10, as is indicated by a large value for J_{H1'H2'}, meaning that the uridine sugar adopts an S-conformation, and the position of the phosphorus resonance of U10. A relatively strong NOE connecting the C9 H4' and the H6 of U10 is observed, while an inter-residue H1'–H6 crosspeak is absent for these residues. This forces the uridine base into an orientation that is different from the preceding residues (discussed below).

The NOEs at the 3'-side of the loop display a more or less regular stacking between C13 and G14 (in the stem). According to the H1'–H2' coupling constant, the sugar pucker of the cytidine is approximately N-type. Alternatively, the preceding residues C11 and G12 have sugar puckers that are in conformational equilibrium between N- and S-type puckers. G12 has only few sugar–base contacts with C13, and shows some irregular NOEs with the sugar of C11. Its sugar proton resonances are broadened at 750 MHz, and the same broadening is seen for residue C11. Interestingly, the lines sharpen up at 500 Hz and the J_{H1'H2'}s for C11 and G12, which were not or hardly resolved at the higher field strength, could be easily

Fig. 2. Overlay of the 10 best out of the final ensemble of 50 structures, superimposed on all residues (A), and on loop residues U7, C8, C9, U10 and C13 (B). Only the loop residues and the closing base pair of the stem are shown.

Fig. 3. Overview of the inter-residue NOE distance restraints used in the structure calculations of the loop residues of the SLIII hairpin. Connectivities between protons are indicated by lines.

measured. This indicates a conformational exchange for these residues with estimated lifetimes of tens of milliseconds. Such an internal mobility is also suggested by the two unidentified imino resonances that are observed in the H_2O spectra, one at 12.1 p.p.m. and another very broad peak at 11.1 p.p.m. They are clearly visible at 400 MHz, but are broadened almost beyond detection at higher field strengths. Unfortunately, no NOEs could be observed to these resonances. Since G12 is a rather unlikely candidate for a visible imino proton considering its orientation in the 3D structure (*vide infra*), these two resonances can be attributed to U7 and U10.

No NOEs from G12 towards the 5'-side of the loop are observed, and in most of the calculated structures its base is not pointing in that direction, which makes it very unlikely to be involved in some kind of base pairing to a residue in the opposite strand. Unfortunately, it is difficult to assess the conformation around the angle χ because of the mobility of this residue mentioned earlier. All NOEs concerning this residue are rather weak because of a possible conformational averaging. Additionally, the G12

$H2'$ - and $H3'$ -resonances are broadened and partly overlapping with other resonances, hampering an accurate determination of χ through the H8 to $H2'/H3'$ NOE intensities. However, as will be discussed in the following section, the data suggest that the guanine base moves within the limits of two conformations, and therefore the χ angle was constrained to the corresponding boundaries of 50–100°.

Discussion

Description of the structure

A representation of the average structure is given in Figure 4. The two bases at the stem-loop junction, U7 and C13, form a hydrogen bond between the O2 of U7 and the amino group of C13 (Figure 5). The geometry of this non-Watson-Crick UC base pair is different from the one observed in the crystal structure of an RNA duplex (Holbrook *et al.*, 1991), where a hydrogen bond is formed between the O4 of the uridine and the cytidine amino group. This difference in configuration is related to the fact that in our case the base of U7 is positioned directly above C6. As a consequence, U7 is turned towards the major groove with respect to C13 and offers its O2 rather than its O4 as a hydrogen bond acceptor to the amino group of C13. The same base-base overlap is observed for residues C8 and C9, which also have their bases turned towards the major groove, rather than towards the helical axis. The orientation of these three residues makes them accessible for possible long-range tertiary interactions in the HDV ribozyme.

The stacked region at the 5'-side of the loop is followed by a turn in the backbone between C9 and U10. This is clearly manifested by the opposite sugar orientation of these residues (Figure 4). This sudden inversion of the direction of the sugar-phosphate backbone is accompanied by an S-puckered conformation of the sugar of U10, and

Fig. 4. Representation [drawn with MOLMOL (Koradi *et al.*, 1996)] of the average structure of the ensemble depicted in Figure 2. Only the loop residues and the closing base pair of the stem are depicted. The current view is towards the major groove side of the loop, and the base of G12 is pointing out of the plane. No sugar protons are shown, except for the 2'-hydroxyls. Oxygen atoms are marked in red, phosphorus in yellow, nitrogen in blue, base carbons in green and hydrogens in white. Putative hydrogen bonds (based on the proximity of the donors and acceptors in the calculated structures) are marked as blue broken lines.

Fig. 5. Geometry of the UC base pair formed between the first (U7) and the last (C13) nucleotide in the loop. The present data suggest a hydrogen bond between H41 of the cytosine and O2 of the uridine.

causes its base to be positioned directly above the phosphate of C9. The stabilizing role of such a base-phosphate stacking interaction is complemented by two hydrogen bonds, i.e. between the hydroxyls of C9 and U10 and the phosphate oxygens of U10 and C9, respectively. This local network of interactions keeps the participating residues in place and explains the non-standard conformational properties such as the S-pucker of U10. It also separates the rigid 5' region of the loop from the more flexible 3' region, in particular residues C11 and G12. The dynamic character of these two residues, which was suggested by the NMR data, is also apparent from the family of structures in Figure 2. However, on average, the base of C11 seems to be roughly positioned above its own phosphate, thus enabling base-phosphate stacking.

The flexibility in this part of the loop is even more evident for residue G12. The N1 of the base is pointing towards the major groove (which is towards the viewer

in Figure 4), but there is a considerable degree of motional freedom. The H1'-H2' scalar coupling constant of 5 Hz reveals a conformational equilibrium between N- and S-puckering of the sugar. Structure calculations in which the sugar pucker was constrained to either one of these conformations resulted in two different families in which the amino group of the guanine is close to hydrogen bonding distance to the phosphate of either G12 or C13, respectively. In the absence of sugar puckering restraints, the orientation of the base is, on average, somewhere in between these positions. The sugar of G12 remains more or less in place in these different situations, which is reflected by the fact that its hydroxyl proton is always within or near hydrogen bonding distance to the phosphate of C13.

Novel structural features

The SLIII loop contains a number of structural features not observed earlier. The new configuration of the UC base pair has already been mentioned in the preceding section. The other aspects can best be discussed by comparing the present results with the well-studied structure of the tRNA-anticodon loop, which is also a seven-membered loop. In the anticodon loop the structure is dominated by extensive base-base stacking interactions proceeding from the 3'-end of the loop through the fifth nucleotide, after which a sharp turn between the fifth and the sixth nucleotide changes the direction of the sugar-phosphate backbone (see Figure 6A). The remaining gap is closed by the last two residues, i.e. at the most 5'-side of the loop [for a review of the tRNA^{Phe} crystal structure

Fig. 6. Schematic overview of the π -turn in the yeast-tRNA^{Phe} anticodon loop (A) and the 'reversed π -turn' found in the SLIII loop (B). The direction of the chain is indicated by small arrows next to the sugars (depicted as ellipses). Other symbols are: plates (bases), spheres (phosphates), dashed arrows (base-base stacking) and curved arrows (base-phosphate stacking).

see for instance Quigley and Rich (1976)]. In the geometrical model developed for loop folding (Haasnoot *et al.*, 1986; Hilbers *et al.*, 1994) it has been indicated how such an A-type stacking configuration reduces the distance between the fifth residue and the 3'-end of the helix to such an extent that the remaining gap can be closed by the remaining two residues. As a result of this stacking, the bases involved are in a suitable orientation for engagement in anticodon-codon interactions.

The sharp change in the direction of the backbone is brought about by the turning phosphate, i.e. the torsion angles in the phosphate adopt a so-called π 3-turn (ζ^- , α'), between U33 and G34 (numbering of yeast-tRNA^{Phe}, see Figure 6). The sharp turn is stabilized by the stacking of the uridine base (U33) on the stacking phosphate G34-p-A35, which is located at the 3'-side of the turning phosphate (see Figure 6). In the SLIII loop the situation is reversed, i.e. the positions of the turning and the stacking phosphates have been interchanged. Thus, the phosphate in the SLIII loop corresponding to the turning phosphate in the anticodon loop has become the stacking phosphate and vice versa: the stacking phosphate (C8-p-C9) precedes the turning phosphate (C9-p-U10). Also in contrast to the anticodon loop, in the SLIII loop the uridine (U10) involved in phosphate stacking now directly follows the turning phosphate instead of preceding it. This means that

it has shifted two positions in the 3'-direction compared with U33 in the anticodon loop; in other words, in the 3-nucleotide sequence making up the U-turn the uridine is now at the end and not at the start and therefore we designate this configuration as a reversed U-turn. Analogously to the U-turn found in the anticodon loop in tRNA, it is expected that this turn may also be found in other molecules.

The orientation around the reversed U-turn has interesting and important consequences for the loop conformation. In the anticodon loop, the base of U33 stacks on the stacking phosphate by turning inwards into the loop forming, through its imino proton, a hydrogen bond with the phosphate of A36. On the other hand, in the SLIII loop, stacking of the base of U10 on the stacking phosphate turns it outwards towards the solvent (see Figure 4) and it is not possible to form a hydrogen bond by its imino proton and acceptor groups in the molecule. However, the turn is stabilized by a hydrogen bond of the hydroxyl group of the sugar of U10 and the stacking phosphate (Figure 4). One might wonder whether the same configuration could prevail if U10 were substituted by a cytidine. At this point, we see no problem with this; mutational data from Kawakami *et al.* (1993) indicated that a U10C substitution in the HDV genomic ribozyme did not abolish the catalytic activity.

Is the structure of the central loop in the antigenomic ribozyme preformed in the SLIII loop?

The structural and functional features of the central loop in the HDV ribozyme have been the subject of a number of biochemical studies. A comparison of these results with the structure obtained in the present paper for the isolated hairpin, SLIII, suggests that the basic structural features of the latter are retained in the ribozyme. Mutational studies (Tanner *et al.*, 1994) have indicated that the guanine corresponding to G12 in the SLIII loop is not involved in Watson-Crick base pairing with the cytosine corresponding to C8 in the SLIII loop in the opposite strand. This conforms nicely to the SLIII loop structure. The hydrogen bond pattern of the hydroxyl groups and the stacking in the 5'-side of the loop pulls the base of C8 away from G12. Close inspection of G12 and its surroundings in a spacefilling model also reveals that an approach of the guanine base towards the cytosine base is hindered by the sugar ring of the latter. This leaves the guanine as an element of local mobility in an otherwise relatively rigid loop structure. The mutational studies have also indicated that the nucleotides of the 5'-half of the central loop in particular are very important for catalysis. In order to obtain a clearer picture of the role of these residues of the central loop, more data are needed describing the tertiary fold of the molecule. Recently, progress in this matter has been achieved by studying the formation of long range photocrosslinks within the genomic and antigenomic ribozyme molecules (Bravo *et al.*, 1996; Rosenstein and Been, 1996). Bravo *et al.* (1996) showed that, using photo-active thiouridine, crosslinks could be obtained between substrate positions -1 and -2 (defined with respect to the cleavage site) and the cytosine and guanine corresponding to G12 and C8 in the SLIII loop, respectively. Rosenstein and Been (1996) observed, by employing the photo-activatable azidophenacyl group, that

the phosphate at the cleavage site crosslinks to the 3'-side of the central loop, that is to and adjacent to the guanine residue corresponding to G12 in the SLIII loop. These results show that in the tertiary structure these loop residues are in close proximity to, or form part of the catalytic site. In the SLIII loop these residues are in positions which, if maintained in the central loop in the ribozyme, would be easily accessible to these photocrosslinkers. Unfortunately, the photocrosslinking data are not able to discern the orientation of the central loop with respect to the cleavage site. In the three-dimensional model proposed by Tanner *et al.* (1994), the central loop is oriented with its minor groove in the direction of the site of cleavage, while the major groove side of the loop is directed towards the solvent. Conversely, the present data show that some of the loop residues, in particular the important residue G12, are accessible from the major groove side of the loop. In view of the mentioned cross-linking data, this implies a somewhat different orientation of the central loop relative to the other important domains in the ribozyme. As mentioned earlier, in the 5'-side of the SLIII loop the residues are turned outwards to the solvent. In such an orientation the bases could easily form, together with the residues in the JI/IV and JII/IV regions (see Figure 1), a substrate binding pocket as suggested by Rosenstein and Been (1996).

Materials and methods

Sample preparation

The 5'-pppGGGACCUCUCCGCGGUGCC-3' RNA oligonucleotide was prepared from a partially duplex DNA template by *in vitro* transcription with T7 RNA polymerase (Milligan *et al.*, 1987). The RNA was purified by preparative gel electrophoresis. The total yield was ~3 mg of RNA out of a 100 ml transcription reaction. The sequence of the RNA transcript was confirmed using an RNA enzyme sequencing kit (Pharmacia).

NMR spectroscopy

NMR samples were prepared by dialyzing the purified RNA versus a 10 mM Na₂HPO₄/NaH₂PO₄ buffer at pH 6.55 and subsequent concentration to 500 µl (yielding a 1 mM sample) using a Centricon microconcentrator. NMR experiments were performed on Varian Unity+ 500 and 750, and on Bruker AM 400 and AMX 600 spectrometers. For samples in 90% H₂O/10% D₂O, spectra were recorded at 5°C, 10°C and 15°C. For these samples, 1D spectra and 2D NOESY spectra with a mixing time of 250 ms were obtained using a jump-and-return read pulse (Plateau and Guéron, 1982).

Unless stated otherwise, all experiments on D₂O samples were performed at 25°C. 2D NOESY (Jeener *et al.*, 1979) and TOCSY (Griesinger *et al.*, 1988) experiments were carried out at 5°C and 25°C with different mixing times achieving water suppression by presaturation of the residual HDO peak. All experiments were recorded using the TPPI method (Marion and Wüthrich, 1983) for quadrature detection, unless stated otherwise. ³¹P-¹H correlations were established using a ³¹P-¹H HETCOR (Sklénar *et al.*, 1996), a ³¹P-¹H hetero-homo-TOCSY (Kellogg, 1992) and a ³¹P-¹H hetero-TOCSY-NOESY (Kellogg and Sweizer, 1993) experiment. The latter two experiments were carried out with a hetero-TOCSY mixing time of 75 ms followed by a 50 ms homo-TOCSY or 250 ms NOESY mixing time, respectively. A GARP sequence (Shaka *et al.*, 1985) was used for the decoupling of the phosphorus spins. All these heteronuclear experiments were performed using the States-TPPI method (Marion *et al.*, 1989) for quadrature detection. A ¹³C-¹H HMQC experiment (Bax *et al.*, 1983) at natural abundance was carried out with SHR quadrature detection (States *et al.*, 1982).

Structure calculations

All structure calculations were performed with X-PLOR V3.1 (Brünger, 1992). Distance restraints were derived from 100, 200 and 300 ms NOESY spectra. The NOE cross-peak intensities were corrected for spin-diffusion using the relaxation matrix approach of the program

NO2DI (Van de Ven *et al.*, 1991). Upper and lower bounds were set to 120% and 70% of the calculated distance, respectively.

Sugar puckers were determined from J_{H1'/H2'} coupling constants (Wijmenga *et al.*, 1993). These couplings could be measured for residues C8 through G12 in 2D NOESY and TOCSY spectra by making use of the fine structure of the crosspeaks to the H1' resonances (for numbering see Figure 1). For residues G1, U7 and C13, clear H1'-H2' crosspeaks were observable in 50 ms TOCSY spectra, but the coupling constants could not be determined accurately since the average linewidth of 4.5 Hz does not permit the measurement of splittings <3 Hz. Therefore, J_{H1'/H2'} for these residues was set to an estimated value of 2 Hz. N- and S-type puckers were used as constraints in the case of very weak or very strong H1'-H2' coupling, respectively. No puckering restraint was used in other cases.

Dihedral restraints for the backbone angles α, ζ, β and γ were set within a range of ±20° of the A-form values for residues with helical sequential NOE connectivities and phosphorus chemical shifts around -2.3 p.p.m. that are indicative of a regular helix. The angle ε was set to 225 ± 60° to prevent the occurrence of conformations in the *gauche*⁺ region which is stereochemically forbidden for this angle. χ angles were set to 202 ± 30° or 65 ± 30°, in the *anti* or *syn* range, respectively, depending on their H8/H6 to H1'/H2'/H3' NOEs. The simulated annealing (SA) protocols include an 18 ps MD run of 6000 steps at 1000 K, followed by gradual cooling during 9 ps to 300 K in 3000 steps. In the first SA run, used to obtain random orientations for the loop region and a regular A-helix for the stem, standard A-helix distance and dihedral restraints were imposed on a random extended structure for all residues that are helical according to the NMR data. Subsequently, 10 starting structures were selected that had reasonable torsion angles for the loop region. All NOE data as well as the experimental dihedral angles were then added as restraints and 50 structures were calculated. After two cycles of selection and calculation (discussed in the Results section) 10 structures that best matched the restraints were used as a set of starting structures in the last SA run, employing experimental data for the stem in addition to standard restraints needed to maintain A-helical conformation. Finally, the resulting ensemble was subjected to an energy minimization step.

Acknowledgements

The NMR experiments were performed at the SON HF-NMR facility (Nijmegen, The Netherlands). This research was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization of Scientific Research (NWO). H.A.H. was supported by a grant from the Royal Netherlands Academy of Arts and Sciences (KNAW).

References

- Bax, A., Griffith, R. and Hawkins, B.L. (1983) Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Resonance*, **55**, 301-315.
- Been, M.D., Perotta, A.T. and Rosenstein, S.P. (1992) Secondary structure of the self-cleaving RNA of hepatitis delta virus: applications to catalytic RNA design. *Biochemistry*, **31**, 11843-11852.
- Bravo, C., Lescure, F., Laugaa, P., Fourrey, J.-L. and Favre, A. (1996) Folding of the HDV antigenomic ribozyme pseudoknot structure deduced from long-range photocrosslinks. *Nucleic Acids Res.*, **24**, 1351-1359.
- Brünger, A.T. (1992) *X-PLOR. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
- Fountain, M.A., Serra, M.J., Krugh, T.R. and Turner, D.H. (1996) Structural features of a six-nucleotide RNA hairpin loop found in ribosomal RNA. *Biochemistry*, **35**, 6539-6548.
- Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) Clean TOCSY for ¹H spin system identification in macromolecules. *J. Am. Chem. Soc.*, **110**, 7870-7872.
- Haasnoot, C.A.G., Hilbers, C.W., Van der Marel, G.A., Van Boom, J.H., Singh, U.C., Pattabiraman, N. and Kollman, P.A. (1986) On loopfolding in nucleic acid hairpin-type structures. *J. Biomol. Struct. Dynam.*, **3**, 843-857.
- Hilbers, C.W., Heus, H.A., Van Dongen, M.J.P. and Wijmenga, S.S. (1994) The hairpin elements of nucleic acid structure: DNA and RNA folding. In Eckstein, F. and Lilley, D. (eds), *Nucleic Acids and Molecular Biology*. Vol 8. Springer-Verlag, Berlin, pp. 56-104.

- Hoffman,D.W. and White,S.W. (1995) NMR analysis of the *trans*-activation response (TAR) RNA element of equine infectious anemia virus. *Nucleic Acids Res.*, **23**, 4058–4065.
- Holbrook,S.R., Cheong,C., Tinoco,I., Jr and Kim,S.H. (1991) Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base pairs. *Nature*, **353**, 579–581.
- Huang,S., Wang,Y.-X. and Draper,D.E. (1996) Structure of a hexanucleotide RNA hairpin loop conserved in ribosomal RNAs. *J. Mol. Biol.*, **258**, 308–321.
- Ikura,M., Kay,L.E. and Bax,A. (1991) Triple-resonance multidimensional NMR study of calmodulin complexed with the binding domain of skeletal muscle myosin light-chain kinase: indication of a conformational change in the central helix. *Biochemistry*, **30**, 5498–5504.
- Jeener,J., Meier,B.H., Bachmann,P. and Ernst,R.R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.*, **71**, 4546–4553.
- Kawakami,J., Kumar,P.K.R., Suh,Y.-A., Nishikawa,F., Kawakami,K., Taira,K., Ohtsuka,E. and Nishikawa,S. (1993) Identification of important bases in a single-stranded region (SSrC) of the hepatitis delta (δ) virus ribozyme. *Eur. J. Biochem.*, **217**, 29–36.
- Kellogg,G.W. (1992) Proton-detected hetero-TOCSY Experiments with application to nucleic acids. *J. Magn. Resonance*, **98**, 176–182.
- Kellogg,G.W. and Schweizer,B.I. (1993) Two- and three-dimensional ³¹P-driven NMR procedures for complete assignment of backbone resonances in oligodeoxynucleotides. *J. Biomol. NMR*, **3**, 577–595.
- Koradi,R., Billeter,M. and Wüthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, **14**, 51–55.
- Kumar,P.K.R., Kazunari,T. and Nishikawa,S. (1994) Chemical probing studies of variants of the genomic hepatitis delta virus ribozyme by primer extension analysis. *Biochemistry*, **33**, 583–592.
- Kuo,M.Y., Goldberg,J., Coates,L., Mason,W., Gerin,J. and Taylor,J. (1988) Molecular cloning of hepatitis delta virus RNA from an infected woodchuck liver: sequence, structure, and applications. *J. Virol.*, **62**, 1855–1861.
- Macnaughton,T.B., Wang,Y.J. and Lai,M.M. (1993) Replication of hepatitis delta virus RNA: effect of mutations of the autocatalytic cleavage sites. *J. Virol.*, **67**, 2228–2234.
- Marion,D. and Wüthrich,K. (1983) Application of phase-sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ¹H-¹H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
- Marion,D., Ikura,M., Tschudin,R. and Bax,A. (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Resonance*, **85**, 393–399.
- Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.D. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.*, **15**, 8783–8798.
- Nishikawa,F., Kawakami,J., Chiba,A., Shirai,M., Kumar,P.K.R. and Nishikawa,S. (1996) Selection *in vitro* of *trans*-acting genomic human hepatitis delta virus (HDV) ribozymes. *Eur. J. Biochem.*, **237**, 712–718.
- Perotta,A.T. and Been,M.D. (1991) A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature*, **350**, 434–436.
- Plateau,P. and Guéron,M. (1982) Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences. *J. Am. Chem. Soc.*, **104**, 7310–7311.
- Quigley,G.J. and Rich,A. (1976) Structural domains of transfer RNA molecules. *Science*, **194**, 796–806.
- Rosenstein,S.P. and Been,M.D. (1991) Evidence that genomic and antigenomic RNA self-cleaving elements from hepatitis delta virus have similar secondary structures. *Nucleic Acids Res.*, **19**, 5409–5416.
- Rosenstein,S.P. and Been,M.D. (1996) Hepatitis delta virus ribozymes fold to generate a solvent-inaccessible core with essential nucleotides near the cleavage site phosphate. *Biochemistry*, **35**, 11403–11413.
- Shaka,A.J., Barker,P.B. and Freeman,R. (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Resonance*, **64**, 547–553.
- Sharmeen,L., Kuo,M.Y., Dinter-Gottlieb,G. and Taylor,J. (1988) Antigenomic RNA of human hepatitis delta virus can undergo self-cleavage. *J. Virol.*, **62**, 2674–2679.
- Sklenar,V., Miyashiro,H., Zon,G., Miles,H.T. and Bax,A. (1986) Assignment of the ³¹P and ¹H resonances in oligonucleotides by two-dimensional NMR spectroscopy. *FEBS Lett.*, **208**, 94–98.
- States,D.J., Haberkorn,R.A. and Ruben,D.J. (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson.*, **48**, 286–292.
- Tanner,N.K., Schaff,S., Thill,G., Petit-Koskas,E., Crain-Denoyelle,A.-M. and Westhof,E. (1994) A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. *Curr. Biol.*, **4**, 488–498.
- Van de Ven,F.J.M., Blommers,M.J.J., Schouten,R.E. and Hilbers,C.W. (1991) Calculation of interproton distances from NOE intensities. A relaxation matrix approach without requirement of a molecular model. *J. Magn. Resonance*, **94**, 140–151.
- Wijmenga,S.S., Mooren,M.W. and Hilbers,C.W. (1993) NMR of nucleic acids; from spectrum to structure. In Roberts,G. (ed.), *NMR of Macromolecules: A Practical Approach*. Oxford University Press, Oxford, pp. 217–288.
- Woese,C.R., Winker,S. and Gutell,R.R. (1990) Architecture of ribosomal RNA: constraints on the sequence of ‘tetra-loops’. *Proc. Natl Acad. Sci. USA*, **87**, 8467–8471.
- Wu,H.N., Lin,Y.J., Lin,F.P., Makino,S., Chang,M.F. and Lai,M.M. (1989) Human hepatitis delta virus RNA subfragments contain an autocleavage activity. *Proc. Natl Acad. Sci. USA*, **86**, 1831–1835.

Received on January 22, 1997; revised on February 14, 1997