Nucleic Acids Research, 1993, Vol. 21, No. 2 201–207

Folding of DNA substrate-hairpin ribozyme domains: use of deoxy 4-thiouridine as an intrinsic photolabel

Djalma Vitorino Dos Santos, Anna-Luiza Vianna, Jean-Louis Fourrey[1] and Alain Favre*
Groupe de Photobiologie Moléculaire, Institut Jacques Monod, CNRS Université Paris VII, 2 Place Jussieu, 75251 Paris cedex 05 and [1]Institut de Chimie des Substances Naturelles, CNRS, 91198 Gif Sur Yvette, France

Received November 2, 1992; Revised and Accepted December 14, 1992

ABSTRACT

Hairpin ribozymes derived from (−)sTRSV RNA exhibit substantial cleavage activity when wobble GU base pairs are introduced in place of the AU pairs normally involved in helices I and II between substrate and ribozyme. This finding prompted us to synthesize by *in vitro* transcription a new hairpin ribozyme, active against a 14-mer substrate derived from a conserved HIV sequence. Interactions of the canonical and anti-HIV hairpin ribozymes with non cleavable DNA substrate analogues containing the photoaffinity probe deoxy-4-thiouridine (ds^{4}U) at a single site were investigated. Upon near-UV light irradiation (365 nm), all these substrate analogues were covalently attached to ribozyme via single or multiple crosslinks. In contrast, no crosslinks were detected using either a DNA substrate analogue lacking ds^{4}U or a ds^{4}U containing oligomer unrelated to the substrate sequence. As expected, if the dissociation constant is in the range of 5–15 μM, the yield of crosslinked ribozyme increased markedly with increasing the substrate analogue concentration. The ribozyme residues involved in the crosslinks were determined by RNA sequencing. The pattern of crosslinks obtained with the two ribozyme systems provides additional evidence in support of the consensus secondary structure proposed for the hairpin domain. Minor alternative conformations were detected in the case of the (−)sTRSV system.

INTRODUCTION

Ribozymes catalyse site specific RNA cleavage and ligation reactions and, unlike most protein ribonucleases, exhibit a high degree of sequence specificity. Most of the small self-cleaving RNAs from plant viral satellites, viroids and virusoids can be folded into a 'hammerhead' secondary structure (1–5). Among the few exceptions known so far is the catalytic center of the 359-bases minus strand of the satellite RNA of Tobacco Ringspot Virus {(−)sTRSV} which undergoes both cleavage and ligation (6–7). Cleavage yields a 5′ fragment terminated with a 2′,3′ cyclic phosphate together with a 3′ fragment having a 5′-OH (6), while by means of the reverse reaction a normal 3′-5′ ApG bond is formed (7).

Efficient *trans* cleavage of the RNA substrate by a ribozyme derived from (−)sTRSV RNA has been demonstrated (8). Such a ribozyme can be reduced to a core of 50 nucleotides that behaves as a Michaelian enzyme with respect to its 14-mer substrate. Cleavage occurs at the expected ApG bond with a temperature optimum near 37°C (9). A two-dimensional 'hairpin' model has been proposed for the substrate–ribozyme complex on the basis of computer assisted RNA-folding, phylogenetic data and analysis of compensatory mutations (9–11). Of the four minihelices described in this model, two are formed between substrate and ribozyme (helices I and II), the cleavage site $A_{-1}pG_{+1}$ being present within a four nucleotide loop (Fig.1A). The ribozyme is also able to cleave a DNA substrate analogue containing a single ribonucleotide at A_{-1}, while the full DNA analogue behaves as a poor inhibitor (12). Mutagenesis and *in vitro* selection strategies have shown that substrate G_{+1} is absolutely required for cleavage and ligation to occur. This guanosine, located immediately 3′ to the cleavage-ligation site, is not needed for substrate binding, but its 2-amino group is an essential component of the active site (13).

The secondary structure of the sustrate–ribozyme complex as well as the absolute requirement for the substrate residue G_{+1} suggest that the conformation of the bulge between helices I and II plays a key role in the catalytic process. In order to explore the substrate–ribozyme interactions, we have designed a number of substrate analogues containing at selected positions a probe which minimally perturbs the conformation of the molecule and is able to form covalent bonds with nucleotide residues which are in permanent or transient contact with the probe in the folded structure. Here we have taken advantage of the high photocrosslinking ability of the uridine analogue 4-thiouridine (reviewed in 14) which can be incorporated *in vivo* into the RNA of growing cells to an appreciable level without major alterations in the cellular metabolism (15–17). 4-Thiouridine, which selectively absorbs light in the 330–370 nm range, is able to form photocrosslinks with all major nucleosides by a collisional mechanism (17). It is known to efficiently induce long-range

* To whom correspondence should be addressed

crosslinks both in *cis*, (8−13 link in *E.coli* tRNA, see Ref.14, or multiple crosslinks in a randomly substituted rRNA fragment, see Ref.18) and in *trans* between mRNA analogues and the RNA components of the ribosome (19−20).

Recently, we have developed a reliable method for insertion of deoxy-4-thiouridine (ds^{4}U) residues into synthetic oligodeoxynucleotides (21). Here we report our study on the interactions of these oligodeoxynucleotides with both the standard ribozyme derived from (−)sTRSV and with a new hairpin ribozyme, designed to cleave a conserved sequence of HIV RNA. We show that the substrate analogue interacts with the ribozymes as predicted by the consensus model. Some minor alternative conformations were detected in the case of the (−)sTRSV system.

MATERIALS AND METHODS

Oligonucleotide synthesis

RNAs used in this study (figure 1) were synthesized by *in vitro* transcription with T7 RNA polymerase of synthetic single-strand DNA templates that are double stranded at the promoter site (22). The products of transcription were isolated by running on 20% PAGE-8 M urea (substrate RNAs) or 15% PAGE-8 M urea (catalytic RNAs). The bands of interest were cut out, quantified by Cerenkov counting, eluted by shaking overnight in 3 volumes of 150 mM NaCl, ethanol precipitated and redissolved in H_2O-DEPC (Diethylpyrocarbonate). Single-strand DNA sequences to be used for templates were synthesized on a Gene Assembler Plus DNA Synthesizer (Pharmacia) using the phosphoramidite method. Oligodeoxynucleotides containing a single ds^{4}U residue at defined position were synthesized similarly, following the protocol which uses 4-S-pivaloyloxymethylthiodeoxyuridine phosphoramidite (21). After purification according to size on 20% PAGE-8 M urea, the thiolated and non thiolated oligonucleotides were separated by affinity electrophoresis (18, 23).

Dephosphorylation and labelling

RNA transcripts were dephosphorylated using Calf Intestine Alkaline Phosphatase (CIAP) (24) and subsequently 5′-labelled with [γ^{32}P]ATP according to (25). 3′ end labelling was achieved with T4 RNA Ligase and [α^{32}P]pCp essentially as in (26). Labelled RNAs were isolated and purified as described above.

Cleavage reaction

All RNAs to be used were pre-heated at 65°C for 2 min and cooled in ice. The cleavage reaction was carried out by incubating catalytic RNA with substrate RNA at 37°C in 50 mM Tris−HCl pH 7.5, 10 mM $MgCl_2$, 0.5 mM EDTA at different incubation times and terminated by the addition of 63 mM EDTA. The products were run on 20% PAGE-8 M urea, detected by autoradiography, excised and quantified by Cerenkov counting.

Irradiation and separation of the oligonucleotide-ribozyme crosslinks

The ds^{4}U containing oligodeoxynucleotides (Tables 2 and 3) at various final concentrations were mixed with 0.4 μM end-labelled ribozyme in 50 mM Tris−HCl pH 7.5, 10 mM $MgCl_2$, 0.5 mM EDTA at 4°C. The mixture (5 μl) was introduced in a siliconized glass capillary and placed 2 cm from the exit slit of a Baush and Lomb monochromator equiped with a HBO 200 W superpressure mercury lamp and a Schott WG 345 filter (14). The samples thermostated at 10°C were irradiated at 365 nm during 30 min and purified on 15% PAGE-8 M urea. Crosslinked and non crosslinked species were detected by autoradiography, excised, quantified by Cerenkov counting, eluted and precipitated.

Identification of ribozyme crosslinked residues

Limited alkaline hydrolysis of the end-labelled intact or crosslinked ribozyme was achieved in 50 mM carbonate buffer pH 9.5 in the presence of 0.5 μg/μl of tRNA (27). Each sample, 10 μl, was heated at 95°C for 2 to 3 min, cooled and loaded onto a sequencing gel. The gel (400×200×0.4 mm), 15% polyacrylamide (acrylamide to bis acrylamide weight ratio 19:1) containing 8 M urea was prepared in the electrophoresis buffer: 0.089 M Tris-borate pH 8.3, 2.5 mM Na_2 EDTA. During preelectrophoresis (30 min, 60 W, 60°C) a saline gradient was established by making the lower buffer 1 M sodium acetate. The gels were run at 60 W for 1.5 h. Correspondence between the individual bands of the ladder obtained with intact ribozyme and its sequence was established using limited RNAs digestions and in particular RNase T_1, RNase ϕM and RNase U_2 (27).

RESULTS AND DISCUSSION

The interactions between hairpin ribozyme and substrate related oligonucleotides were investigated here by an improved photoaffinity method, using two different systems. The first one is the ribozyme R1 (Fig.1A), directly derived from the (−)sTRSV RNA (9). The second system is the ribozyme R′ (Fig.1B), designed to match a conserved sequence found in the long terminal repeat (LTR) of HIV RNA at the beginning of region U_5.

Cleavage activity and GU base-pairs

The ribozymes and their cognate substrates were obtained by *in vitro* transcription with T7 RNA polymerase (see Methods). The yield of transcription, however, depends critically upon the first hexanucleotide from the start of transcription (22). Accordingly, it could be increased ten fold when three G residues were substituted for the three A residues at the 5′ end of R1(R2 ribozyme, see Fig.1A). R2 behaved as a Michaelian enzyme in the presence of substrate S (Table 1), retaining substantial cleavage activity. Furthermore, substrate variants containing U either at position −1 (S1) or at position −2 (S2) were cleaved by R2 with bimolecular rate constants $Kcat/K_M$ close to that for the substrate S (6 $min^{-1} \cdot \mu M^{-1}$, see Table 1).

It is well established that the nature of the nucleotide at position −1 does not affect cleavage (11, 28). However, the introduction of a single mismatch within helix I or II inevitably resulted in abolition of catalytic activity. The effect of GU 'mismatches' has not been studied in the hairpin system so far and we demonstrate here that they allow cleavage activity even when present simultaneously in both helix I and helix II, as is the case with the R2-S2 system (Fig. 1A). G-U base pairing has been observed in model bihelical structures and, accordingly, it is considered in the current models for RNA secondary structure in spite of the lower stability of GU as compared to the standard AU pairs (29), they have been introduced in the current models of RNA secondary structure. Evidence for a GU base pair inserted into a standard A RNA helix without significant ribose-phosphate perturbations was indeed obtained from the crystallographic analysis of yeast $tRNA^{Phe}$ (30). It is reasonable to expect, therefore, that in the ribozyme-substrate complex, G-U pairing should not cause significant distortions of minihelices I and II.

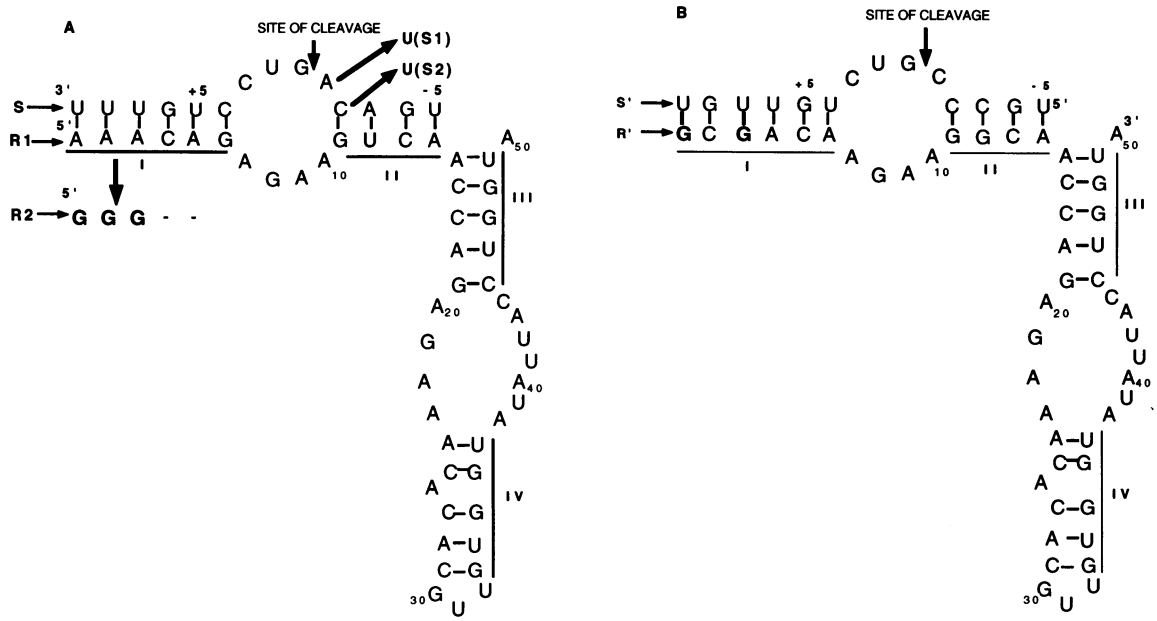

Figure 1. Substrate-ribozyme complexes studied here. Both substrate and ribozyme contain an additional G residue at the 5′ end, added as pppG by T7 RNA polymerase (22). **A**: R1-S is the canonical catalytic domain derived from (−)sTRSV RNA with the current designation of bihelical regions I, II, III, IV (11). R2 is a variant ribozyme modified at the 5′ end to improve transcription. Two variant substrates S1 ($A_{-1}\rightarrow U_{-1}$) and S2 ($C_{-2}\rightarrow U_{-2}$) were also studied. **B**: R′ is an hairpin ribozyme designed to cleave a conserved sequence (S′) found in the Long Terminal Repeat (LTR) region of HIV RNA between positions 107 and 121, numbered following Muesing et al. (39).

Sequence requirements for cleavage of heterologous RNA by an engineered hairpin ribozyme have also been determined. A substrate molecule can be cleaved 5′ of a GUC sequence with flanking regions of variable sequence (11) although only G_{+1} which is proximal to the cleavage site appears absolutely required for activity (13). The selected HIV sequence S′ contains a 5′-CGUC-3′ sequence in appropriate position and ribozyme R′ was designed so as to anneal the flanking regions of the substrate (Fig.1B). In order to improve the transcription yield, the 5′ end of R′ was modifed by introducing 2G in place of 2A residues thus forming two GU base pairs in helix I (Fig.1B). Again, cleavage occurs although with a significantly lower efficiency (Table 1). Hence, in two different hairpin ribozyme systems introduction of GU wobble base pairs allowed substantial cleavage activity. This may have important consequences *in vivo* as a down regulator of gene expression. If regions I and II of the appropriately designed ribozyme contain G and U residues, it could potentially cleave not only the designed target but also related RNA sequences, capable of forming GU base pairs. Minimizing the number of G and U residues within these ribozyme regions is expected to remove degeneracy and increase the specificity of recognition.

Crosslinking formation and identification

Efficient methods have recently been developed for the introduction of sulfur containing bases at selected positions into synthetic oligodeoxynucleotides using phosphoramidite chemistry (31, 32). Following the procedure of Clivio et al. (21), the 4-S-pivaloyloxymethylthiodeoxyuridine phosphoramidite was prepared in three steps from 4-thiodeoxyuridine and used to insert the latter as a single residue at position i of the substrate, originally

Table 1. Kinetic parameters for engineered ribozymes and substrates.

Ribozyme	Substrate	K_M (μM)	k_{cat} (min^{-1})
R1	S	0.030*	2.1*
		0.040	1.0
R2	S	0.035	0.20
R′	S′	0.10	0.02

R1 is the canonical (−)sTRSV ribozyme and S is its cognate substrate. R2 is a variant where the AAA sequence at the 5′ end of R1 (Fig.1A) has been substituted by GGG. R′ is the anti-HIV hairpin ribozyme and S′ is the corresponding substrate (Fig.1B). The asterisks indicates the previously reported data of Hampel and Tritz (9) obtained in the presence of 2 mM spermidine.

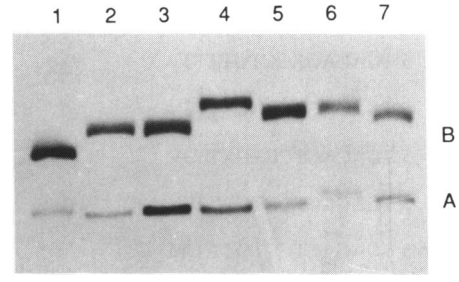

Figure 2. Separation on affinity gel of non thiolated (bands A) and thiolated (bands B) ds^4U containing substrates analogues of (−)sTRSV ribozyme. Lines 1−7 represent dS(−5), dS(+2), dS(+5), dS(+7), dS(+8), dS(+9) and dC, respectively; dC is a 14-mer, unrelated to S.

occupied by the parent nucleotide U. Thus, a series of non cleavable analogues dSi of the substrate S were obtained (Table 2). An advantage of our procedure is that removal of both the normal base label and 4-thiouracil protecting groups is accomplished in a single step. After purification according to size by PAGE, the thiolated oligonucleotides were further purified by affinity electrophoresis (18, 23). The mercury atoms embedded into the affinity gel interact with the SH groups of the thiolated oligodeoxynucleotides and slow down their migration. As seen in Fig.2, the thiolated fraction in the oligonucleotide preparations varied between 50 and 80%.

Each purified analogue dSi directed against the (−)sTRSV derived ribozyme R1 was combined with 5′-^{32}P (in some cases 3′-^{32}P) labelled R1 (0.4 μM) in the cleavage buffer at 4°C and the mixture irradiated at 365 nm with fluences ~30 kJ/m^2, known to induce maximum crosslinking. The mixtures were then subjected to electrophoresis on a denaturing gel to separate crosslinked from unreacted R1. All tested substrate analogues dSi yielded single or multiple crosslinks (Figs.3A and 3B) in a strictly light-dependent process. In contrast, no retarded bands could be detected when R1 was irradiated alone (not shown) or

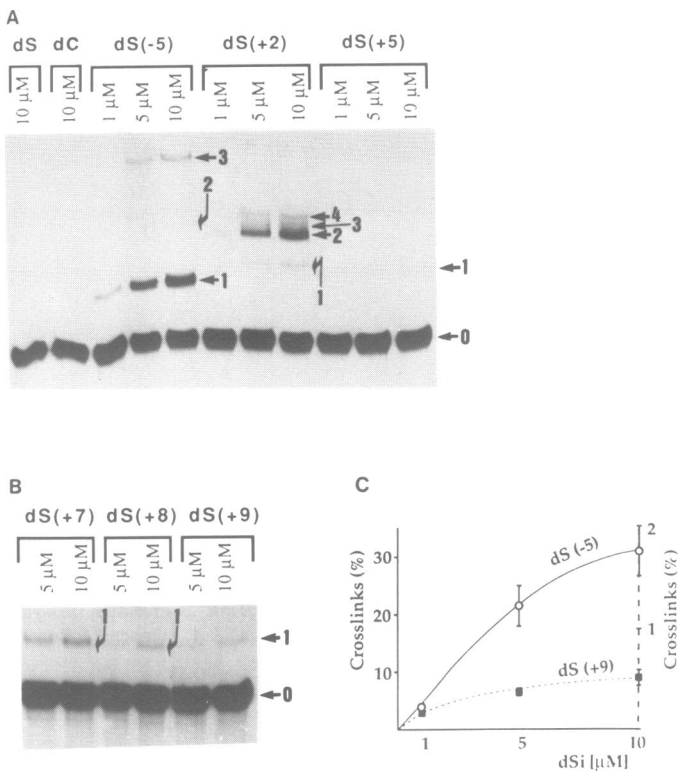

Figure 3. Ribozyme−substrate crosslinks generated at different substrate concentrations. **A** and **B**: autoradiogram of a denaturing 15% acrylamide gel showing the unreacted ribozyme (bands 0) and the different crosslinks (bands 1−4) obtained with the ds^{4}U containing substrate analogues substituted at position **i**. dS does not contain ds^{4}U. dC is a ds^{4}U containing 14-mer, unrelated to the substrate. **C**: changes of the yield of crosslinking, i.e. the sum of crosslinked R1 /unreacted R1 + crosslinked R1 as a function of the oligodeoxynucleotide concentration.

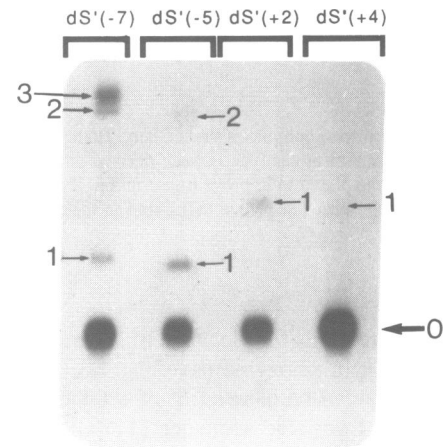

Figure 4. Formation of crosslinks between 5′-^{32}P-labelled anti-HIV ribozyme R′ and its substrate analogues containing a ds^{4}U residue at position **i** (dS′**i**) as assayed by a denaturing 15% polyacrylamide gel. Crosslinks are numbered according to Figs.3A and 3B.

Table 2. Efficiency of crosslink formation and localization of crosslinks obtained with the (−) sTRSV derived ribozyme R_1 and its substrate analogues.

Substrate analogues		**Ei**		Relative yields (%) and position of crosslinks							
Name	Sequence	1.4μM	10μM								
dS	5′TGACAGTCCTGTTT3′		0								
dS(−5)	5′**X**GACAGTCCTGTTT3′	5.6	32	1(**83**) A50		2 (**4**) A7				3 (**13**) A14	
dS(+2)	5′TGACAG**X**CCTGTTT3′	2.3	16	1 (**14**) A9	2 (**55**) G6			3 (**13**) A5		4 (**18**) A7	
dS(+5)	5′TGACAGTCC**X**GTTT3′	0.4	1.3			1(**100**)* A1					
dS(+7)	5′TGACAGTCCTG**X**TT3′	0.4	1.5			1 (**100**) A7					
dS(+8)	5′TGACAGTCCTGT**X**T3′	0.18	0.8			1 (**100**) A1					
dS(+9)	5′TGACAGTCCTGTT**X**A3′	0.2	0.5			1 (**100**) A1					
dC	5′GTTCTAA**X**AATAGC3′	0									

dS(±**i**) are substrate S analogues containing deoxy-4-thiouridine (marked by **X**) at position **i** numbered as shown in Fig.1A. dS is a non thiolated substrate analogue. dC is a control ds^{4}U containing 14-mer, unrelated to S. Note that dS(+9) has an additional nucleotide in 3′ position required for its chemical synthesis. Total crosslinking efficiency (**Ei**) is defined as % 5′^{32}P-labelled ribozyme R1 crosslinked to a given dSi relative to the total amount of R1. 1.4 μM and 10 μM indicate the concentrations of the substrate analogues. The relative yield of given crosslinks (marked 1−4, following their migrations on the gel as shown in Fig.3A and 3B) is determined as described in the legend to Fig.3. *In this case no clear-cut window on the gel has been observed (see Results and Discussion).

in the presence of dS, a substrate analogue lacking ds^{4}U (Fig.3A). Photoactivated ds^{4}U, therefore, is absolutely required for crosslinking. Furthermore, no crosslinks were detected in the absence of Mg^{2+} (not shown), which is expected to strongly stabilise the R1-dSi complex, or upon using the probe dC, a ds^{4}U containing analogue unrelated to S (Fig.3A).

Each crosslinked species was then purified and sequenced following a limited alkaline hydrolysis. Intact 5′-^{32}P labelled R1 yielded a control ladder where the positions of individual nucleotides were identified using limited RNases digestions (Fig.5). When R1 is crosslinked at nucleotide Y, all cuts on the 5′ side of the bridge generate a partial ladder that parallels the control one. On the other hand, all cuts 3′ to the bridge generate 5′-^{32}P labelled oligonucleotides attached to the alkaline resistant deoxyoligomer. Their migration is retarded, leading to the formation of a clear cut window upon the gel, which allows an unambigous identification of Y (Fig.5). The pattern of observed crosslinks is summarized in Table 2. With the exception of the minor crosslink R-dS(+5) which never produced a clear-cut window on the gel, all listed crosslinks were highly reproducible and, in some cases, confirmed by using 3′ labelled ribozyme. The unusual behaviour of the mentioned minor crosslink might

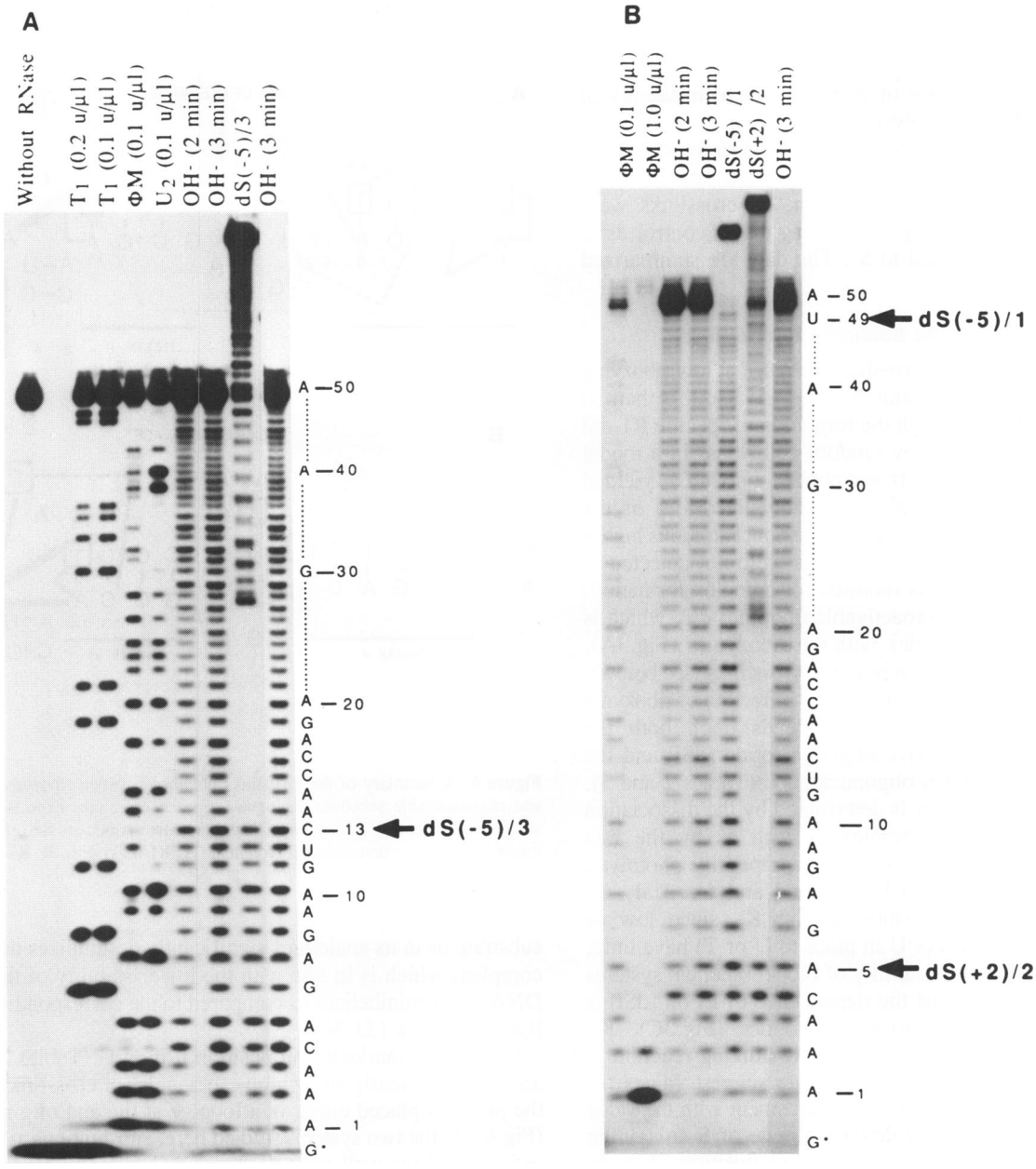

Figure 5. Identification of the ribozyme residues involved in crosslinks 1–3 (See Figs.3A and 3B). Autoradiograms of the sequencing gel of 5′-^{32}P-labelled (–)sTRSV ribozyme R1 after limited digestion with specific RNases or after limited alkaline hydrolysis (OH^-) of R1 and the purified crosslinked complexes. The arrows indicate the ribozyme nucleotide immediately 5′ to the crosslinked residue. The asterisks designates the additional G residue added at the 5′ end by the T7 RNA polymerase. **A**: A14 is the ribozyme residue involved in the crosslinks 3 {R1- dS(–5)}. **B**: A50 and G6 are the ribozyme residues involved in the crosslinks 1 {R1- dS(–5)} and 2 {R1- dS(+2)}, respectively.

Table 3. Efficiency of crosslink formation and localisation of crosslinks between anti-HIV ribozyme and its substrate analogues

Substrate analogues		Ei	Relative yields (%) and position of crosslinks			
Name	Sequence					
dS′	5′TGCCCGTCTGTTGT3′	0				
dS′(−7)	5′XATGCCCGTCTGTTGT3′	9.5	1(22) A50		2(20) C13	3(58) A14
dS′(−5)	5′XGCCCGTCTGTTGT3′	3.5		1(73)	2(27)	
dS′(+2)	5′TGCCCGXCTGTTGT3′	2			1(100) A6	
dS′(+4)	5′TGCCCGTCXGTTGT3′	1.1			1(100) A6	
dC	5′GTTCTAAXAATAGC3′	0				

All designations are as in Table 2, except that position i is numbered as in Fig.1B and the crosslinked species are marked following their migration in the gel shown in Fig.4. In this case a single substrate concentration was used [1.4 μM]. Note that dS′(−7) has two additional nucleotides in 5′ position.

be due to unresolved mixture of crosslinks or to instability of the bridge during OH^-hydrolysis.

The R′-dS′i system was studied at single concentration of dS′ (1.4 μM). The separation of the crosslinked species from unreacted R′ is shown in Fig.4. Again, no crosslinks were detected in the absence of Mg^{2+}, or using dC, a control ds^{4}U containing oligomer unrelated to S′. The data are summarized in Table 3.

Hairpin ribozyme-substrate analogue interactions.

Our data demonstrate that the crosslinks obtained at low ribozyme and substrate analogue concentrations (0.4 μM and less than 10 μM, respectively) occur through the formation of specific R1-dSi (R′-dS′i) complexes and not by random collisions. In a model oligonucleotide system, the trinucleotide d(s^{4}UCC) yielded detectable 'collisional' crosslinks with ribonucleosides or the corresponding homodecanucleotides only at concentrations higher than 1 mM (17). Furthermore, crosslinks were not detected in the absence of Mg^{2+}, which is required for complex formation, or in the presence of the photoactivable oligomer dC, which is unable to form a stable complex with the ribozymes (Fig.3A). This is further proved by the nature of the crosslinks (see below).

The efficiency of crosslinking (Ei), i.e. the fraction of ribozyme that has been crosslinked to dSi, depends upon both the concentration of the photoactive oligonucleotide [dSi] and the position i of ds^{4}U within this oligonucleotide (Tables 2 and 3). More precisely, the efficiency is determined by the dissociation constant K_D of the R-dSi complex as well as by the rate constants for crosslink formation and for competitive photolytic processes which take place in both bound and free dSi. All analogues are expected to exhibit similar K_D since low or moderate substitution levels (s^{4}U in place to U or T) have little, if any, effect on the thermal stability of model bihelical systems (14, 31). Having in mind that the dependence of Ei on [dSi] is similar for the different substrate analogues (Fig.3C), K_D appears to be a basic parameter of the crosslinking reaction in this system (See Tables 2 and 3). Our data suggest values for K_D between 5 and 15 μM which are in agreement with the value of K_M (10 μM) reported for a deoxyanalogue of S containing a single ribonucleotide rA at the cleavage position (12). In contrast, the K_M for the substrate is much lower than 10 μM (Table I) and the inhibition constant K_I for a non cleavable substrate analogue, containing dA at position −1 was found to be 4 μM. Therefore, the introduction of deoxyresidues in the

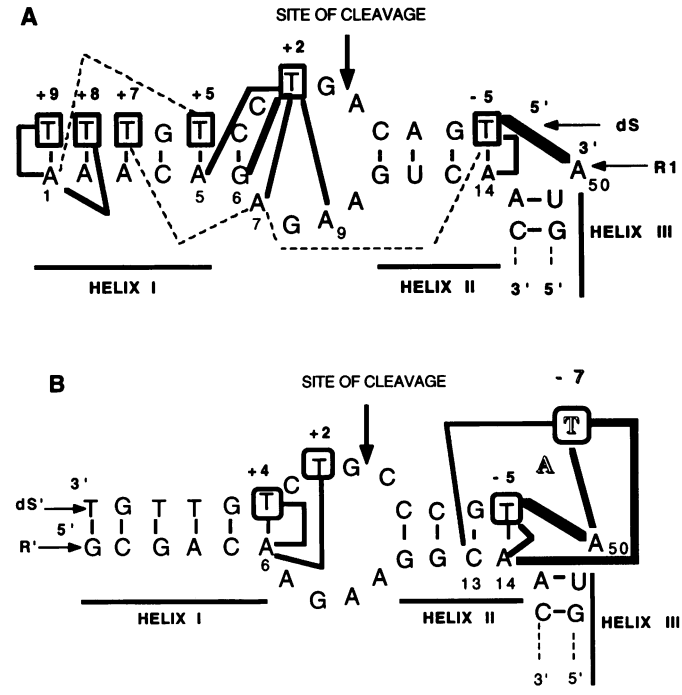

Figure 6. A summary of the crosslinks obtained between ribozyme R1 or R′ and photoactivable substrate analogues. ——— crosslinks expected from the secondary structure; the width of the line corresponds to the efficiency of crosslinking; - - - - non-expected crosslinks. **A**: R1-dS system. **B**: R′-dS′ system.

substrate or in its analogues significantly destabilizes the R1-dSi complex, which is in line with the lower stability of the hybrid DNA-RNA minihelices as compared to the corresponding RNA-RNA duplexes (33,34).

Ei depends markedly on position i of ds^{4}U (Tables 2 and 3), and our data clearly show high efficiencies of crosslinking when the probe is placed either in a loop or at the end of a minihelix (Fig.6). In the two systems studied here, substitutions at positions −5 and +2 (as well as at positions −7 and, to a lesser extent, +4 in R′-dS′i only) resulted in moderate or high crosslink yields, in contrast to the results obtained when the ds^{4}U residues were placed within a helix (positions +5, +7 and +8 in R1-dSi). The same rule was observed in the study of a self-organized ribosomal

fragment with random s^4U substitutions (18). It is well established the two end pairs of a minihelix largely fluctuate (35). The reactive residues, therefore, are located at positions that allow sufficient mobility, necessary for an optimal matching between excited ds^4U and its partner nucleotide. An apparent exception is dS(+9) which reacts poorly with the 5′ end of R1. A possible explanation is that in the R1-dS(+9) complex ds^4U reacts much more efficiently with its neighbour 5′ residue, thymidine being the most reactive acceptor nucleoside (17,36–38) than in *trans* with the 5′ end of R1.

The majority of the crosslinks are in agreement with the consensus structure (Fig.6) although the observation of some minor crosslinks suggest the existence of alternative conformers. In both R1-dS and R′-dS′ systems position −5 reacts predominantly with A50 followed by A14, whereas no crosslinks with A15-U49 base pair were detected, suggesting that a transient formation of continuous helix II + helix III structure is unlikely to occur. Owing to its higher mobility, position −7 in the extended analogue dS′(−7) reacts with the same residues as well as with C13. Similarly, positions +9 and +8 in R1-dS and position +4 in R′-dS′ yielded crosslinks in agreement with helix I. An interesting finding are the crosslinks involving position +2 of the substrate, which in both systems reacts with position 6 of ribozyme. However, while in R′-dS′ position +2 reacts exclusively with position 6, in the R1-dS system, besides the major +2 to 6 crosslink, a few additional minor crosslinks were detected (Figs. 5 and 6). These multiple crosslinks could be due to enhanced flexibility of the bulge, but more likely reflect the lower intrinsic reactivity of the acceptor G residues as compared to the other bases (17). In any case, the data suggest a preferential folding of the bulge with residue T_{+2} in close proximity to the purine at position 6 of ribozyme. In conclusion, all results and considerations about the role of K_D, K_M and position of substitution in the efficiency of crosslinking claim that crosslinking is accomplished via complex formation between ribozyme and substrate.

In the R1-dS system, however, a few minor long-range crosslinks were found that definitively could not be accomodated in the consensus structure. Such is the case with the bridges between positions +7, +5 and −5 of the substrate analogue and, respectively, residues A7, A1 and A7 of R1. Their formation necessarily requires the existence of minor conformers. For example, the generation of the bridge dS(−5)- A7 could be explained if one assumes that following the formation of the standard helix I, the remaining of the substrate analogue is transiently folded on itself with A_{-3} G_{-4} bases paired to T_{+2} C_{+3} respectively, thus bringing T_{-5} in close proximity to A7. A similar explanation can be suggested for the bridge linking T_{+7} to residue A7 of R1: the formation of helix II is accompanied by a transient base pairing of the substrate analogue T_{+9},T_{+8} and T_{+7} to R1 residues A5, G6 and A7, respectively.

The present data clearly demonstrate that crosslinking of ribozyme to its substrate analogues by means of the intrinsic photolabel s^4U is accomplished via complex formation between the two molecules. This property of the photocrosslinking procedure makes it a powerful approach for structural analysis of nucleic acids as illustrated here for the hairpin ribozyme domain. The use of photoactivable ribooligonucleotide substrate analogues with either the complete hairpin ribozyme or the region which is involved in helices I and II will help to more precisely define the active site of the hairpin domain.

ACKNOWLEDGEMENTS

We are grateful to our colleagues A.Woisard and I.Pashev for helpful discussions. This work was supported by Grant ACC 4 of ANRS (Agence Nationale de Recherches sur le SIDA) to A.F.

REFERENCES

1. Forster,A.C. and Symons, R.H. (1987) Cell **49**, 211–220.
2. Uhlenbeck, O.C. (1987) Nature **328**, 596–600.
3. Symons, R.H. (1989) TIBS **14**, 445–450.
4. Epstein, L.M. and Gall, J.G. (1987) Cell **48**, 535–543.
5. Rubino, L., Tousignant, M.E., Steger, G. and Kaper,J.M. (1990) J. Gen. Virol. **71**, 1897–1903.
6. Buzayan, J. M., Gerlach, W. L. and Bruening, G. (1986) Nature **323**, 349–353.
7. Buzayan, J. M., Hampel, A. and Bruening, G. (1986) Nucleic Acids Res. **14**, 9729–9743.
8. Feldstein, P. A., Buzayan, J. M. and Bruening, G. (1989) Gene **82**, 53–61.
9. Hampel, A. and Tritz, R. (1989) Biochemistry **28**, 4929–4933.
10. Haseloff, J. and Gerlach, W. L. (1989) Gene **82**, 43–52.
11. Hampel, A., Tritz, R., Hicks, M. and Cruz, P. (1990) Nucleic Acids Res. **18**, 299–304.
12. Chowrira, B. M. and Burke, J. M. (1991) Biochemistry **30**, 8518–8522.
13. Chowrira, B. M., Berzal-Herranz, A. and Burke, J.M. (1991) Nature **354**, 320–322.
14. Favre, A., (1990) Bioorganic Photochemistry: Photochemistry and the Nucleic Acids Vol. **1** H. Morrison Ed.; J. Wiley & Son, Inc., New York, pp. 379–425.
15. Favre, A., Bezerra, R., Hajnsdorf, E., Lemaigre-Dubreuil, Y. and Expert-Bezançon, A. (1986) Eur. J. Biochem. **160**, 441–449.
16. Bezerra, R. and Favre, A. (1990) Biochem. Biophys. Res. Comm. **166**, 29–37.
17. Favre, A., Lemaigre-Dubreuil, Y. and Fourrey, J. L. (1991) New J. Chem. **15**,593–599.
18. Lemaigre-Dubreuil, Y., Expert-Bezançon, A. and Favre, A. (1991) Nucleic Acids Res. **19**, 3653–3660.
19. Wollenzien, P., Expert-Bezançon, A. and Favre, A. (1991) Biochemistry **30**, 1788–1795.
20. Dontsova, O., Kopylov, A. and Brimacombe, R. (1991) EMBO J.**10**, 2613–2620.
21. Clivio, P., Fourrey, J.L., Gasche, J., Audic, A., Favre, A., Perrin, C. and Woisard, A. (1992) Tetrahedron Lett. **33**, 65–68.
22. Milligan, J. F., Groebe, D. R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Nucleic Acids Res. **15**, 8783–8798.
23. Igloi, G. L. (1988) Biochemistry **27**, 3842–3849.
24. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. Cold Spring Habor Laboratory Press **1.**
25. Silberklang, M., Gillum, A. M. and RajBhandary, U.L. (1979) Methods Enzymol. **59,** 58–109.
26. England, T.E. and Uhlenbeck, O.C. (1978) Nature **275**, 560–561.
27. D'Alessio, J.M. (1982) in Gel Electrophoresis of Nucleic Acids (Richmond, D. and Hames, B.D. Eds) pp. 173–197.
28. Berzal-Herranz, A., Joseph, S. and Burke, J.M. (1992) Genes Dev. **6,** 129–134
29. Borer. P.N., Dengler, B., Tinoco, I. Jr. and Uhlenbeck, O.C. (1974) J. Mol. Biol. **86**, 843–853.
30. Jack, A., Ladner, J.E. and Klug, A. (1976) J. Mol. Biol. **108**, 619–649.
31. Xu, Y-Z., Zheng, Q. and Swann, P.F. (1992) Tetrahedron **48**, 1729–1740.
32. Nikiforov, T.T. and Connolly, B.A. (1992) Nucleic Acids Res. **20**, 1209–1214.
33. Hall, K.B. and McLaughlin, L.W. (1991) Biochemistry **30**,10606–10613.
34. Bevilacqua, P.C. and Turner, D.H. (1991) Biochemistry **30**, 10632–10640.
35. Heus, H.A. and Pardi, A. (1991) J. Mol. Biol. **217**, 113–124.
36. Peak, M.J., Midden, W.R., Babasick, D.M. and Haugen, D.A. (1988) Photochem. Photobiol. **48**, 229–232.
37. Fourrey, J.L., Gasche, J., Fontaine, C., Guittet, E. and Favre, A. (1989) J. Chem. Soc. Chem. Commun. 1334–1336.
38. Clivio, P., Fourrey, J.L., Gasche, J. and Favre, A. (1991) J. Am. Chem. Soc. **113**, 5481–5483.
39. Muesing, M.A., Smith, D.H., Cabradilla, C.D., Benton, C.V., Lasky, L.A. and Capon, D.J. (1985) Nature **313**, 450–458.