doi:10.1006/jmbi.2001.4472 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 307, 51–65

JMB

# The Hairpin Ribozyme Substrate Binding-domain: A Highly Constrained D-shaped Conformation

**Robert Pinard[1], Dominic Lambert[2], Joyce E. Heckman[1] José A. Esteban[3], C. William Gundlach IV[4], Ken J. Hampel[1] Gary D.Glick[4], Nils G. Walter[4], François Major[2] and John M. Burke[1]***

[1]*Markey Center for Molecular Genetics, Department of Microbiology and Molecular Genetics, The University of Vermont, Burlington VT 05405, USA*

[2]*Département d'Informatique et Recherches Opérationnelles Université de Montréal Montréal, Québec, H3C 3J7 Canada*

[3]*Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA*

[4]*Department of Chemistry University of Michigan, Ann Harbor, MI 48109, USA*

**Corresponding author*

The two domains of the hairpin ribozyme-substrate complex, usually depicted as straight structural elements, must interact with one another in order to form an active conformation. Little is known about the internal geometry of the individual domains in an active docked complex. Using various crosslinking and structural approaches in conjunction with molecular modeling (constraint-satisfaction program MC-SYM), we have investigated the conformation of the substrate-binding domain in the context of the active docked ribozyme-substrate complex. The model generated by MC-SYM showed that the domain is not straight but adopts a bent conformation (D-shaped) in the docked state of the ribozyme, indicating that the two helices bounding the internal loop are closer than was previously assumed. This arrangement rationalizes the observed ability of hairpin ribozymes with a circularized substrate-binding strand to cleave a circular substrate, and provides essential information concerning the organization of the substrate in the active conformation. The internal geometry of the substrate-binding strand places G8 of the substrate-binding strand near the cleavage site, which has allowed us to predict the crucial role played by this nucleotide in the reaction chemistry.



*Keywords:* hairpin; ribozyme; modeling; crosslinking; domain A

## Introduction

Determination of the three-dimensional structure or even the global shape of catalytically active RNA molecules remains a very difficult task. To date, only three ribozyme structures have been solved to atomic level: the hammerhead ribozyme (Pley *et al*., 1994; Scott *et al*., 1995, 1996), the P4-P6 domain of the *Tetrahymena thermophila* group I intron (Cate *et al*., 1996a,b) and the hepatitis delta virus (Ferré-D'Amaré & Doudna, 2000). These works have been crucial to understanding how catalytic RNA molecules can self-assemble through a small number of structural motifs.

One intrinsic limitation of high-resolution techniques, e.g. NMR or X-ray crystallography, is their inability to define flexible regions. This problem is especially pertinent for the resolution of ribozyme structures, since catalytically relevant conformations are expected to be rather flexible and/or strained. Several studies have shown that the active site of the hairpin ribozyme is composed of multiple conserved groups that are located in the two domains of the RNA and brought together in the three-dimensional structure *via* a flexible hinge region (reviewed by Walter & Burke, 1998). There is no X-ray crystal or complete NMR structure for the hairpin ribozyme (Cai & Tinoco Jr., 1996; Butcher *et al*., 1999) and few studies have investigated the global structure of the ribozyme-substrate complex (Pörschke *et al*., 1999). Three-dimensional models for the isolated domains of the hairpin ribozyme have been recently reported based on NMR spectroscopy (Cai & Tinoco, 1996; Butcher *et al*., 1999). However, as recognized by the authors, it is not clear whether it represents a catalytically competent structure, since the structure of the loops may change upon interaction with one another.

Lower-resolution techniques such as chemical modification (Butcher & Burke, 1994b), hydroxyl

Abbreviations used: FRET, fluorescence resonance energy transfer.

E-mail address of the corresponding author: john.burke@uvm.edu

0022-2836/01/010051–15 $35.00/0 

radical footprinting (Hampel *et al.*, 1998) site-specific crosslinking (Sigurdsson *et al.*, 1995; Earnshaw *et al.*, 1997; Pinard *et al.*, 1999a), fluorescence resonance energy transfer (FRET) (Walter *et al.*, 1998) and transient electric birefringence (Pörschke *et al.*, 1999) have been used to obtain partial insights into the spatial arrangement of RNA molecules. Although none of these techniques is able to provide atomic details, they can generate very useful constraints for tertiary structure prediction and modeling (Earnshaw *et al.*, 1997; Pinard *et al.*, 1999a,b). Until recently, very few topographical constraints were available to guide modeling efforts of the hairpin ribozyme. Hydroxyl radical footprinting studies and site-specific photoaffinity crosslinking experiments performed in our laboratory have shed some light on the relative alignment of the two domains in the active docked state and provided distance constraints between specific residues in the molecule (Hampel *et al.*, 1998; Pinard *et al.*, 1999a). Using these data in conjunction with molecular modeling (constraint-satisfaction program MC-SYM), we have predicted and experimentally demonstrated the existence of a crucial interdomain base-pair between the conserved guanosine residue immediately downstream of the cleavage/ligation site and C25 in the ribozyme large domain (domain B).

For reasons of simplicity, the global shape of the individual domains is usually depicted as straight structural elements. Here, we have used various crosslinking approaches and specifically engineered constructs to study the three-dimensional proximity of the two ends of the substrate-binding domain and to generate more topographical constraints essential for three-dimensional modeling. Using MC-SYM and the structural constraints described here, we have generated a three-dimensional model for the internal geometry of the substrate-binding domain in a catalytically active complex. This model depicts how the area surrounding the cleavage site is organized in the docked state of the ribozyme and suggests that the substrate-binding domain probably adopts a bent conformation during the course of the cleavage reaction The D-shape arrangement of the domain places G8 of the substrate-binding strand in proximity to the cleavage site and suggests that it could play a crucial role in the self-cleaving mechanism.

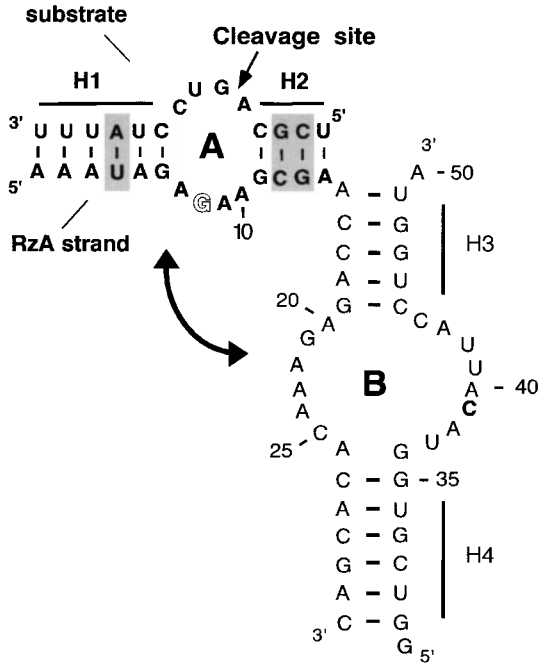


**Figure 1.** The hairpin-ribozyme substrate complex. Secondary structure of the hairpin ribozyme substrate complex used in this study. Three base-pairs (one in H1 and two in H2) of the naturally occurring substrate have been changed to minimize self-complementarity (Esteban *et al.*, 1997). A rate-enhancing U39C mutation was also introduced. Modifications are indicated by a gray shadow. The substrate-binding domain modeled by MC-SYM is in bold. The cleavage site between residues A − 1 and G + 1 is indicated with an arrow. G8, a crucial nucleotide in the reaction chemistry, is indicated in white.

## Results

### Crosslinking approaches

We have used various crosslinking approaches to explore the global shape of the substrate binding domain in the docked configuration and to generate more topographical constraints essential for computer-assisted modeling (MC-SYM). As depicted in Figure 2(a)-(c), three different approaches were used to probe the conformation of the domain. In the first method, the ribozyme-substrate complex was assembled in the presence of cobalt hexammine and irradiated with ultraviolet light (312 nm). In a typical crosslinking experiment, one of the strands was $^{32}$P end-labeled. The crosslinked species generated were isolated on denaturing polyacrylamide gels and were mapped by partial ribonuclease digestion under denaturing conditions or by partial alkaline hydrolysis. Crosslinking reactions performed in the presence of various concentrations of cations or the absence of one or more strands of the ribozyme-substrate complex indicated that the crosslinked species obtained is dependent upon the formation of properly assembled and folded complex (data not shown). In the second approach, an azidophenacyl crosslinking agent was coupled to a sulfur atom that replaced a non-bridging oxygen atom at a specific site in the substrate-binding domain as described (Pinard *et al.*, 1999a). With this method, one of the strands of the ribozyme substrate complex carries the crosslinking agent and one is an end-labeled target strand (see Figure 2(b)). The reassembled complex is irradiated with 312 nm light, isolated and sequenced as described above. For each set of crosslinking reactions, several control experiments

were performed. We have verified that all the purified crosslinked species were occurring exclusively *via* the azidophenacyl group, by performing the crosslinking reactions in the presence of uncoupled phosphorothioate-containing oligonucleotide. We have also verified that the azidophenacyl group does not perturb cleavage activity (data not shown). Finally, we have probed the conformation of the substrate-binding domain using engineered disulfide crosslinks. Alkylthiol groups were introduced on the 2′-hydroxyl as illustrated in Figure 2(c). Ten disulfide combinations were tested. The alkylthiol groups were introduced into the ribozyme substrate-binding strand at positions 7-11 (a single position at a time) and at position 25 in the bottom part of the large loop B or at position 36 in the opposite strand of loop B.

## Crosslink species and mapping of the crosslinking sites

When the ribozyme-substrate complex was incubated in the presence of cobalt hexammine, which stabilizes the formation of the docked complex, and irradiated with UV light, a specific crosslink was observed consistently. The efficiency of formation of this crosslink varied from 4-8 %. In order to determine which strands of the ribozyme-substrate complex were involved in that contact, a set of crosslinking reactions, where the 5′end-labeled strand varied, were performed. Gel analysis revealed that the crosslink observed is an intra-domain crosslink involving both strands of the substrate binding domain The gel-purified crosslinked RNA species was mapped by partial alkaline hydrolysis of 5′end-labeled material (Figure 2). Sequence analysis was performed on both strands to determine each extremity of the crosslink. The result indicates that nucleotides A − 1 in the substrate and G8 in the substrate-binding strand are involved in that crosslink. The similarity between this interstrand crosslink and the interstrand crosslink observed for the well characterized 5 S rRNA loop E-like structures indicates that stacking between A − 1 and G8 is most likely responsible for the formation of the UV-induced linkage observed here (Wimberly *et al.*, 1993; Butcher *et al.*, 1999). The only other interstrand UV-induced crosslinks for which the exact chemical structures are known involve cytidine and uridine bases, which form cyclobutane dimers and require base stacking (see Behlen *et al.*, 1992) The locations of the crosslink sites are shown in Figure 2 and summarized in Figure 1(a). In order to place the formation of the G8-A − 1 stack within the general folding pathway of the ribozyme-substrate complex we monitored the ability of these sites to become crosslinked as a function of time after adding $Co(NH_3)_6^{3+}$. A 5′end-labeled construct of the minimal ribozyme-substrate complex, shown in Figure 3, was incubated with $Co(NH_3)_6^{3+}$ for the indicated time and then UV irradiated (312 nm) for ten seconds. The rate of stacking in $Co(NH_3)_6^{3+}$ monitored by this method was 2.03(±2.2) $min^{-1}$, identical, within error, with the rate of docking determined by time-resolved hydroxyl-radical footprinting under identical conditions (Hampel & Burke, 2001). Non-docking controls did not give rise to detectable levels of G8-A − 1 crosslinks.

Using the azidophenacyl photocrosslinking system, one crosslinked species was observed for each individual position of the photoagent in the ribozyme ribose-phosphate backbone (Figures 2(b). The crosslinked species were gel-purified and submitted to partial ribonuclease digestion under denaturing conditions and by partial alkaline hydrolysis of 5′ end strands. A fraction (1-5 %) of the crosslinked strand is released from the linked species during the purification. This release is not unprecedented and has been described (Wower *et al.*, 1989; Burgin & Pace, 1990; Pinard *et al.*, 1999a). Introduction of the photoagent between positions 7 and 8 in the substrate strand results in a crosslink to position 13 in the substrate-binding strand. The two photoagent insertions in the substrate-binding strand between positions 2 and 3, and 4 and 5, respectively, formed crosslinked species that exhibit similar mobility on gel electrophoresis and yet map to the same position, A38 in the right strand of the B domain.

Finally, disulfide crosslink formation was attempted with all ten dithiol combinations (Figure 2(c)); only the G8-C25 pair reacted to form a disulfide bond. Treatment of the crosslinking species with DTT results in products that co-migrate with the corresponding precursor modified RNA. The yield of crosslinked ribozyme was 5-10 %, which is comparable to the overall yield of disulfide crosslinks obtained with the hammerhead ribozyme (Sigurdsson *et al.*, 1995). The relatively low crosslink yield suggests the disulfide bond may be strained and/or the alignment of the alkylthiol groups for these two positions is suboptimal. Nevertheless, this active inter-domain crosslink clearly indicates that the ribose 2′-hydroxyl groups of these residues are in close proximity and represents an important constraint for the positioning of the G8 residue in the active site.

## Assembly and catalytic activity of crosslinked species

Each individual gel-purified crosslink was examined for ability to mediate cleavage (Figure 5). The crosslinked species were isolated from gels, and reassembly of an active ribozyme-substrate complex was attempted by incubating the purified species with an excess of the missing strand: the 3′ strand of domain B or the 5′ strand of the ribozyme, depending on the species studied.

Three of the five crosslinked species described here and used for modeling retain catalytic activity in normal cleavage conditions. Between 10 and 15 % of the intra-domain crosslinked substrate (A − 1 linked to G8) was cleaved after one hour at 37 °C (Figure 5(a)). The azidophenacyl-dependent

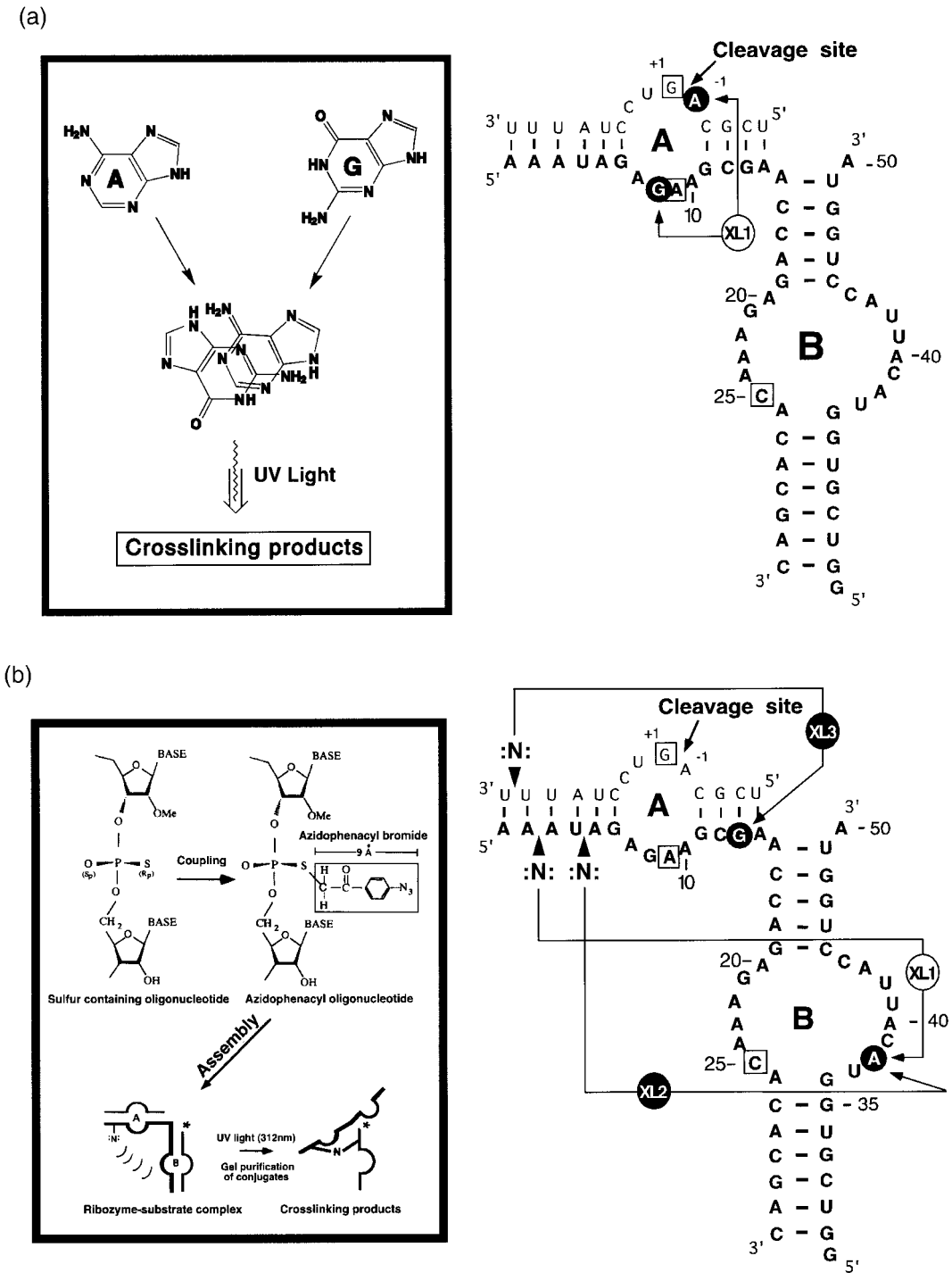


**Figure 2** (*legend opposite*)

interdomain crosslink between positions 2 and 3 in the substrate (photoagent insertion site) and position 38 in the 3′ strand of domain B, retain catalytic activity (1-5 % of cleavage observed), whereas the crosslink between positions 4 and 5 and 38 is inactive (see Figure 5(b)). As mentioned, some of the crosslinked species were contaminated with spontaneously dissociated material during the purification. To determine the contribution of this small contamination in the cleavage assay, various concentrations of the corresponding non-crosslinked strand and a tenfold excess of the missing strand were used as controls for the cleavage assay. The cleavage expected from the contamination is shown in the control lanes on the right of Figure 5(b). Our results clearly indicate that the

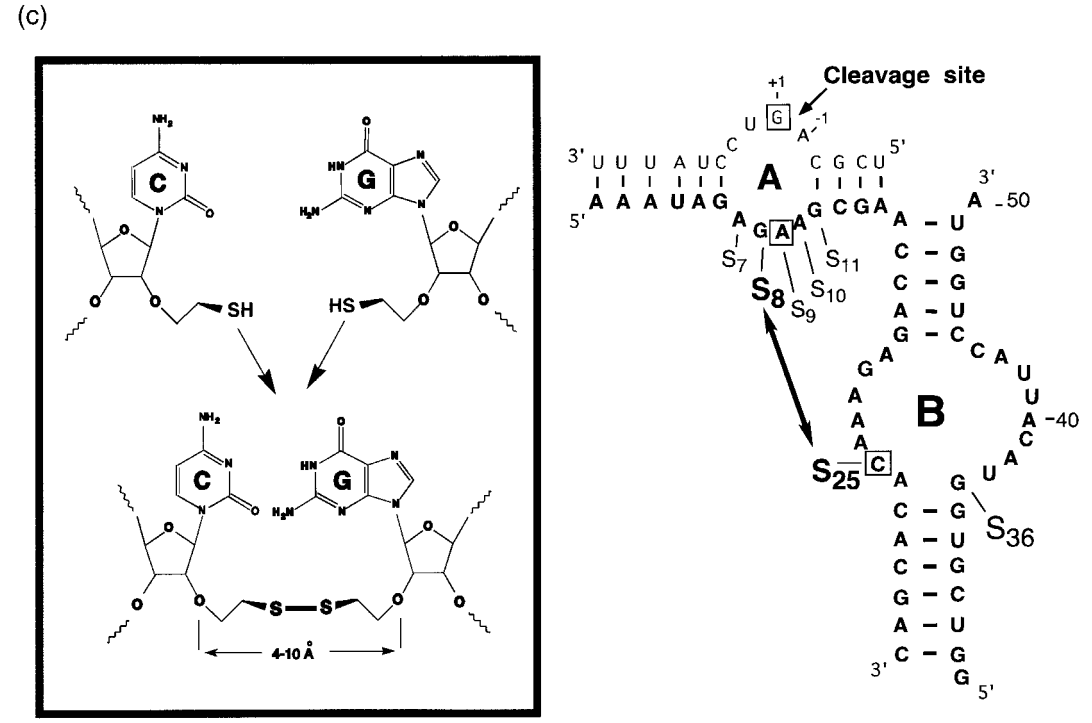


**Figure 2.** Crosslinking approaches. Secondary structure of the complex between the hairpin ribozyme and its cognate substrate is indicated on the right of each panel. The bases involved in the proposed triplet interactions are boxed. The crosslinked sites are indicated with black circles. Ribozyme nucleotides are numbered from 1 to 50. Substrate nucleotides are numbered with negative numbers 5′ to the cleavage site, and *vice versa*. A short arrow indicates the cleavage site. (a) Photocrosslinking *via* base stacking in the presence of cobalt hexammine. The assembled complex is irradiated with ultraviolet light and the crosslinking products are analyzed through 20 % polyacrylamide gels. (b) The photoagent-containing strand is pre-assembled with the other strands and irradiated with 312 nm light. The target strand indicated by an asterix permits the identification and analysis of the cross-strand crosslink. Three crosslinks have been identified. The active species are labeled with a white circle. (c) Alkythiol-containing RNAs were used in reassembly of the ribozyme-substrate complex. Ten dithiol combinations were tested: 7, 25 or 36; 8, 25 or 36; 9, 25 or 36; 10, 25 or 36; and 11, 25 or 36. The active disulfide crosslink between G8 and C25 is indicated by a double-headed arrow. The disulfide crosslink formation was conducted in air at room temperature as described (Results & Materials and Methods). The active disulfide crosslink species between G8 and C25 is indicated by a double-headed arrow.

small concentration of the dissociation products is responsible for a very small proportion of the cleavage observed but the efficiency of cleavage of the active species is much higher compared to the controls. Finally, the disulfide crosslinked ribozymes were analyzed by cleavage assays of 5′ end-labeled substrate. We were able to purify enough material to perform multiple turnover reactions and determine a cleavage rate for the crosslinked species of $3.6 \times 10^{-4}$ min$^{-1}$ (Figure 5(c)). We have used as a control a similar concentration of non-crosslinked thioalkyl-modified oligomers and obtained a rate slightly faster ($7.2 \times 10^{-4}$ min$^{-1}$). These results indicate that the rate and the extent of cleavage observed for the crosslinked complexes cleavage are comparable to the non-crosslinked complexes with thioalkyl-modified RNA strands. However, in both cases, the rates obtained are slower than the unmodified ribozymes in the same conditions (0.11 min$^{-1}$).

On non-denaturing polyacrylamide gels, the inactive crosslinked species are capable of reconstituting a ribozyme-substrate complex and still do not show cleavage activity (not shown). These results suggest that catalytically important regions or local tertiary structure are probably perturbed by the crosslink formation. However, we cannot rule out the possibility that these constraints also may depict an active conformation of the molecule. This is the case for the crosslink involving position 7-8 in H1 and position 13 in H2, which indicates that the two helices could be in close proximity.

### Cleavage activity of circularized domain A

To evaluate the possibility that the substrate-binding domain adopts a bent conformation during the course of the cleavage reaction, presumably during docking of the two domains, we have tried to mimic the bent conformation by locking the domain in a constrained shape, first by circu-

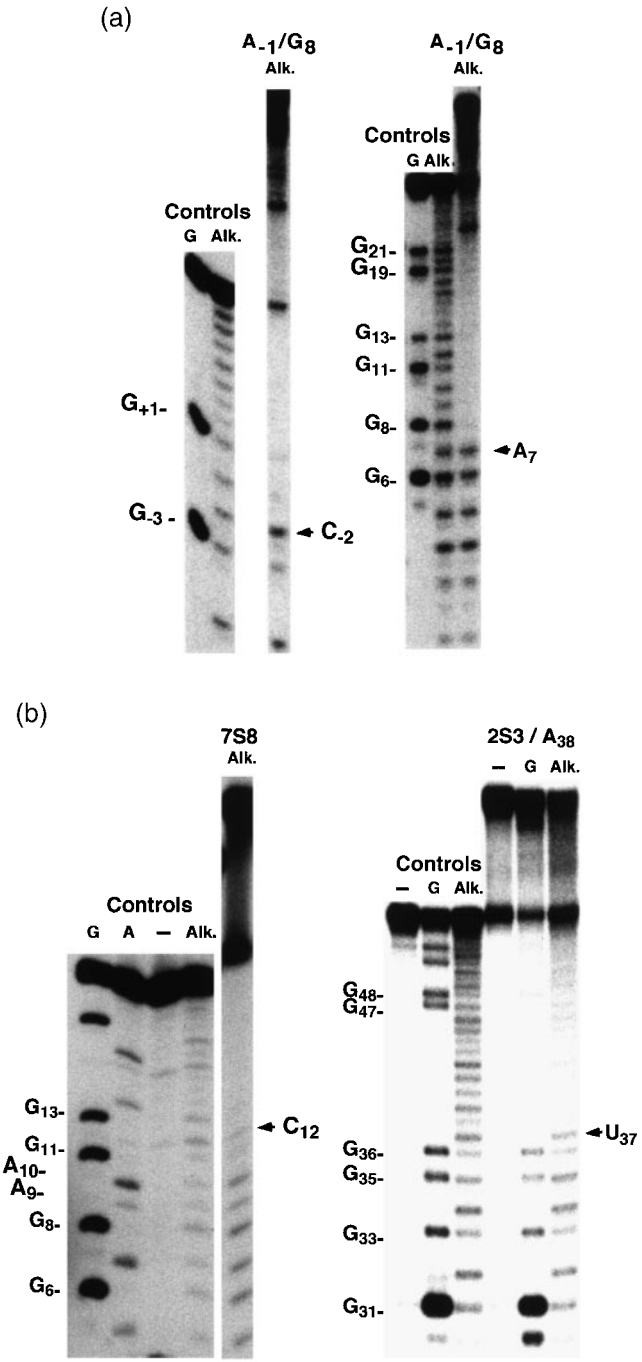


**Figure 3.** Mapping of crosslinked sites. The crosslinked sites were mapped by limited alkaline hydrolysis and/or partial digestion with RNASE $T_1$ (G). The interruptions in the patterns are indicated by arrows. Cleavage occurring 3′ of the conjugated site will create a fragment with a slower mobility and create a gap allowing the deduction of the crosslinked sites. (a) Analysis of the UV-induced crosslink between A − 1 and G8. The (5′-$^{32}$P) end-labeled substrate (left) or substrate-binding strand (right) was used. (b) Photoaffinity crosslinking analysis. The photoagent was introduced between positions 7 and 8 in the substrate (left) or positions 2 and 3 (right) in the substrate-binding strand or 4 and 5 (not shown), see also Figure 2(b).

larizing the substrate, then the substrate-binding strand, and finally both strands simultaneously. We have circularized a substrate extended at the 3′ end with two guanosine residues (Figure 6(a)) and performed cleavage reactions under single turnover conditions, (see Materials and Methods). Note that the circular form of the substrate migrates faster than the linear form on denaturing polyacrylamide gels. Circular molecules were cleaved , as shown by the appearance of the corresponding linear form as a function of time, with a rate (0.066 $min^{-1}$) similar to the rate observed with the linear counterpart (0.11 $min^{-1}$) under our conditions (Figure 6(b)). RNA sequencing reactions

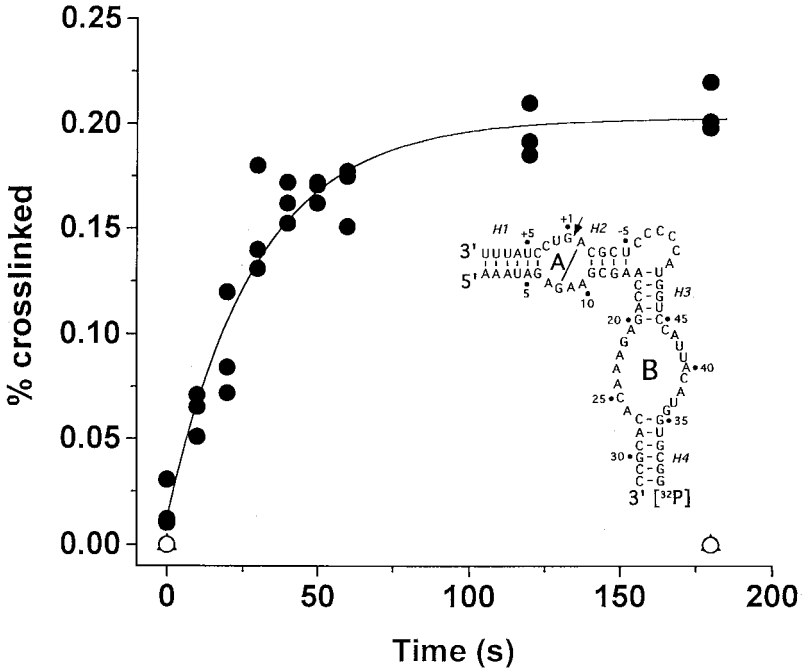


**Figure 4.** Acquisition of a G8-A − 1 stack during docking can be followed by time-resolved crosslinking. A minimal ribozyme-substrate complex was allowed to fold at 25 °C in the presence of 2 mM $Co(NH_3)_6^{3+}$ for the specific times indicated on the *x* axis and then subjected to $UV_{312nm}$ irradiation for ten seconds (three replicate experiments, filled circles). Open circles and triangles represent control experiments where the construct was incubated in the absence of cation or the labeled strand contained a G + 1 A mutation, respectively. The resulting cross-linked and non-cross-linked RNAs were separated on an 8 M urea/10 % polyacrylamide gel and quantified by phosphorimager analysis. The curve fit to a single exponential equation yielded a rate of 2.03(±0.22) $min^{-1}$ and a *y*-intercept of 0.013.

indicated that the cleavage site was at the same position as the normal cleavage (not shown). Although they are cleaved at a rate and to an extent similar to linear substrates, circular substrates showed a low affinity for the ribozyme. Using non-denaturing polyacrylamide gels we have determined a $K_d$ value of 100 nM, which is about 50 times lower than the value measured for the linear substrate. We have measured the cleavage activity with three saturating concentrations of ribozymes (5, 10 and 15 μM) and similar rates were observed. We have tested whether the crucial interdomain base-pair between the conserved guanosine immediately downstream of the cleavage and ligation site (+1 position) and C25 in the ribozyme large domain was still necessary for the cleavage of circularized substrate. To test this hypothesis, we used a G-to-A mutation at the +1 position that does not significantly affect substrate binding but impairs docking. Cleavage was observed only with a ribozyme that contained a C-to-U mutation at the pairing partner position (C25), although with a much slower rate (see Figure 6(b)).

We have then tested the cleavage activity with a ribozyme where the 5′ end of the substrate-binding strand is brought into close proximity of the 3′ end of H3. This construct is the counterpart of the circularized substrate and can be seen as a circularization of the substrate-binding domain (Figure 6(a)). Surprisingly, this construct can cleave both linear (0.037 $min^{-1}$) and circular (0.062 $min^{-1}$) substrates at a cleavage rate only two- to threefold slower than with the non-constrained ribozyme (Figure 6(c) and (d)). The affinity of a linear substrate for this ribozyme is low with a $K_d$ of 266 nM determined by gel-shift assays. We were not able to determine a precise $K_d$ for the binding of the circular substrate but rate measurements carried out at various concentrations indicated that the value is higher than 5 μM. The same rates were obtained with three concentrations of ribozymes (10, 15 and 20 μM), indicating that the rate measurements were carried out at saturating conditions.

## Modeling of the substrate-binding domain using MC-SYM

A model of the substrate-binding domain in the context of a docked configuration was generated using the constraint satisfaction program MC-SYM (see Materials and Methods). The average distance for the disulfide and azidophenacyl crosslinks was limited to 7.5-15 Å. When all the constraints were provided to MC-SYM, the generated model, subsequently submitted to energy minimization, showed that the substrate-binding domain adopts a D-shape, resulting in a sharp bend within the substrate.

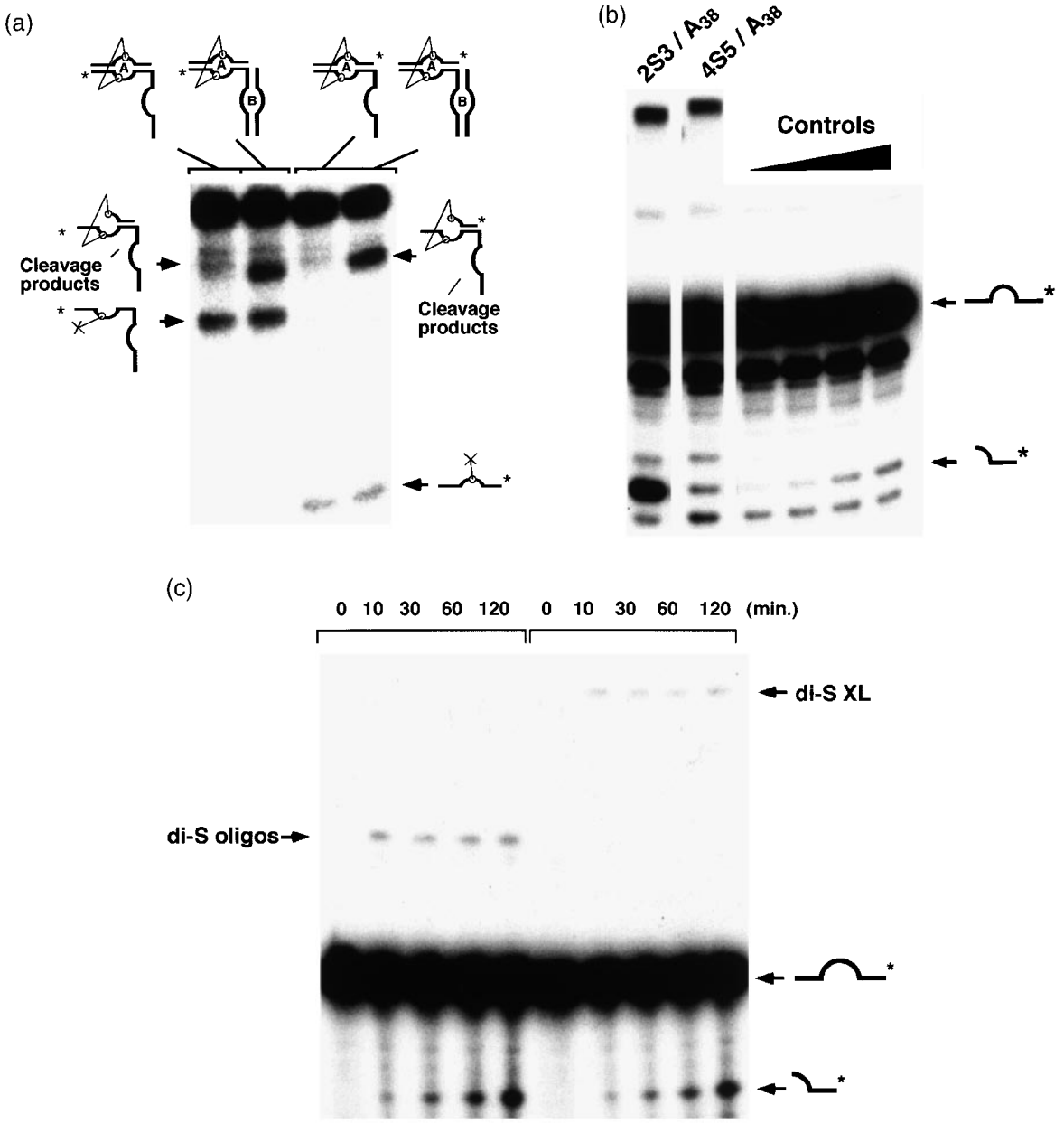


**Figure 5.** Cleavage activity of crosslinked ribozyme. (a) Cleavage assays were performed by providing the right strand of the ribozyme large domain to the crosslinked substrate-binding domain. Cleaved, uncleaved and reversal products are indicated by arrows. (b) Crosslinked products containing the two strands of the ribozyme covalently linked were reassembled in the presence of $MgCl_2$ and assayed for cleavage activity against 5′ end-labeled substrate. Increasing amounts (from 1-10 pM) of uncrosslinked materials were used as a control. Cleaved and uncleaved products are indicated on the right. (c) Time-courses for the cleavage reaction of the 5′ end-labeled substrate with uncrosslinked hairpin ribozyme possessing thio-modified oligomers (positions 8 and 25) (left) and with the active species bridging G8 and C25 (right). The hairpin ribozyme complexes were reassembled in the presence of $MgCl_2$ and the reaction was started by adding the right (independent) strand of the large domain.

## Discussion

The aim of the present study was to investigate the global conformation of the substrate-binding domain of the hairpin ribozyme in an active complex and refine our understanding of the hairpin ribozyme three-dimensional structure by generating more topographical constraints essential for computer-assisted modeling (MC-SYM). Using various approaches, we have identified crosslinked species capable of reassembly and catalysis of the cleavage reaction. These results indicate that the constraints obtained most likely reflect the actual structure of an active ribozyme-substrate complex. In the case of the hairpin ribozyme, it is important to consider that this active conformation is stabilized by specific interactions between loop A and loop B, and it is essential to study the conformation

of the ribozyme in the docked state when the two domains are in intimate contacts. One way to achieve this is to improve the ability of the ribozyme to carry out a correct and more stable inter-domain docking. In the present study we have used a version of the hairpin ribozyme where three base-pairs (one in H1 and two in H2) of the naturally occurring substrate have been changed to minimize self-complementarity and to increase the stability of the docked conformation (Esteban *et al*., 1998; see also Figure 1).

We have used in the molecular modeling process all the active crosslinks presented here in addition to data published previously (Pinard *et al*., 1999a). The substrate-binding domain shown here represents a portion of a complete model, currently being built and refined in our laboratories, where inter-domain as well as intra-domain interactions influence the global shape of the molecule. When all our constraints were provided simultaneously to the constraint-satisfaction program MC-SYM, a kink within the substrate, near the cleavage site, was observed, in agreement with the ability of the hairpin ribozyme to cleave short circular substrates. Interestingly, a preliminary 3D model for the hairpin ribozyme also proposed a dynamic kink in loop A induced by residue U + 2 of the substrate bulging out of the helix (Earnshaw *et al*., 1997). The substrate-binding domain of the hairpin ribozyme comprises two short helices that flank a small symmetrical loop. A kink within this type of arrangement is not unprecedented and bends in symmetrical loops have been described (Tang & Draper, 1994; Zacharias & Hagerman, 1996).

Further support for the model shown in Figure 7 is provided by the experiments carried out in this and other laboratories connecting the 5′ end of the substrate-binding strand (H1) with the distal end of H4 through linkers of different lengths (Komatsu *et al*., 1995; J.A.E. & J.M.B., unpublished results). The H1-H4 connection yielded active ribozymes, although their catalytic efficiency was severely reduced and in the former case, very long linkers were required. This suggests that although they are closer in the docked than the extended conformation, as shown by FRET assays, the ends of H1 and H4 cannot be linked together easily. Interestingly, the kink described here in the model generated by MC-SYM indicates that the ends of H1 and H2 are in closer proximity than was assumed previously. This result is in agreement with the ability to create this bend artificially by connecting the 5′ end of the substrate-binding strand (H1) with the 3′ end of H3 without affecting cleavage significantly. Moreover, the striking observation that this constrained ribozyme is able to cleave circularized substrate provides additional evidence for the presence of a bend in the substrate-binding domain where the ends of H1 and H2 are close to one another. It is possible that the hairpin-substrate complex in the docked state is much more compact then previously assumed, which could explain the strong energy transfer measured by FRET even when fluorophores were attached to the extremity of H1 and H4 (Walter *et al*., 1998). Interestingly, it also has been reported that cross-links between position U + 7 in the substrate (middle of H1) and U42 and C44 (proximity of helix 3) in the ribozyme catalytic domain did not drastically reduce the cleavage activity and also support a model where the substrate-binding domain adopts a bent conformation (Earnshaw *et al*., 1997).

The actual conformation of the substrate-binding domain described in our model is imposed by a series of topographical constraints. Their structural interpretation is relevant for the understanding of the steps involved in the docking event and the reaction mechanism taking place. The formation of all the crosslinked species described here was dependent upon the formation of properly assembled and folded ribozyme-substrate complexes. Various lines of evidence suggest that the two domains interact with considerable adjustments of their structures, principally in each internal loop (Grasby *et al*., 1995; Schmidt *et al*., 1996; Walter & Burke, 1997). The D-shape conformation of domain A might represent an induced-fit stabilized by the interaction with domain B. It is possible that the substrate-binding domain initially adopts an extended conformation that allows the formation of loop A and base-pairing in H1 and H2, which is followed, during docking of the two domains, by a conformational rearrangement bringing the ends of H1 and H2 closer to one another. This hypothesis is being investigated in our laboratory.

At this point, it is not well understood why the small symmetrical loop A would adopt a sharply bent conformation when assembled in a catalytically active complex. In the docked state, there is formation of an essential inter-domain Watson-Crick base-pair between the guanosine residue adjacent to the cleavage site and the cytosine residue at position 25 in the left strand of the ribozyme catalytic domain. According to our model, one role for the kink in the substrate strand could be to facilitate the base-pair formation by favoring an anti-parallel arrangement of the two nucleotides involved. The sharp bend could have a role in positioning essential nucleotides at or near the cleavage site in the docked complex. Transphosphoesterification reactions catalyzed by the hairpin require the deprotonation of the 2′ hydroxyl group immediately 5′ of the cleavage site phosphate group to promote the nucleophilic attack. In the hairpin ribozyme, it has been demonstrated that the catalysis does not require metal-bound water or direct coordination of metal cations to phosphate oxygen atoms, suggesting that the RNA itself might play the role of the general base or acid in the reaction. However, the p$K_a$ values of the side-chains of nucleotides (3.5-4.5) seem too low to provide efficient acid-base catalysis at neutral pH. It is possible that these values are shifted toward neutrality in certain conformations of an RNA mol-

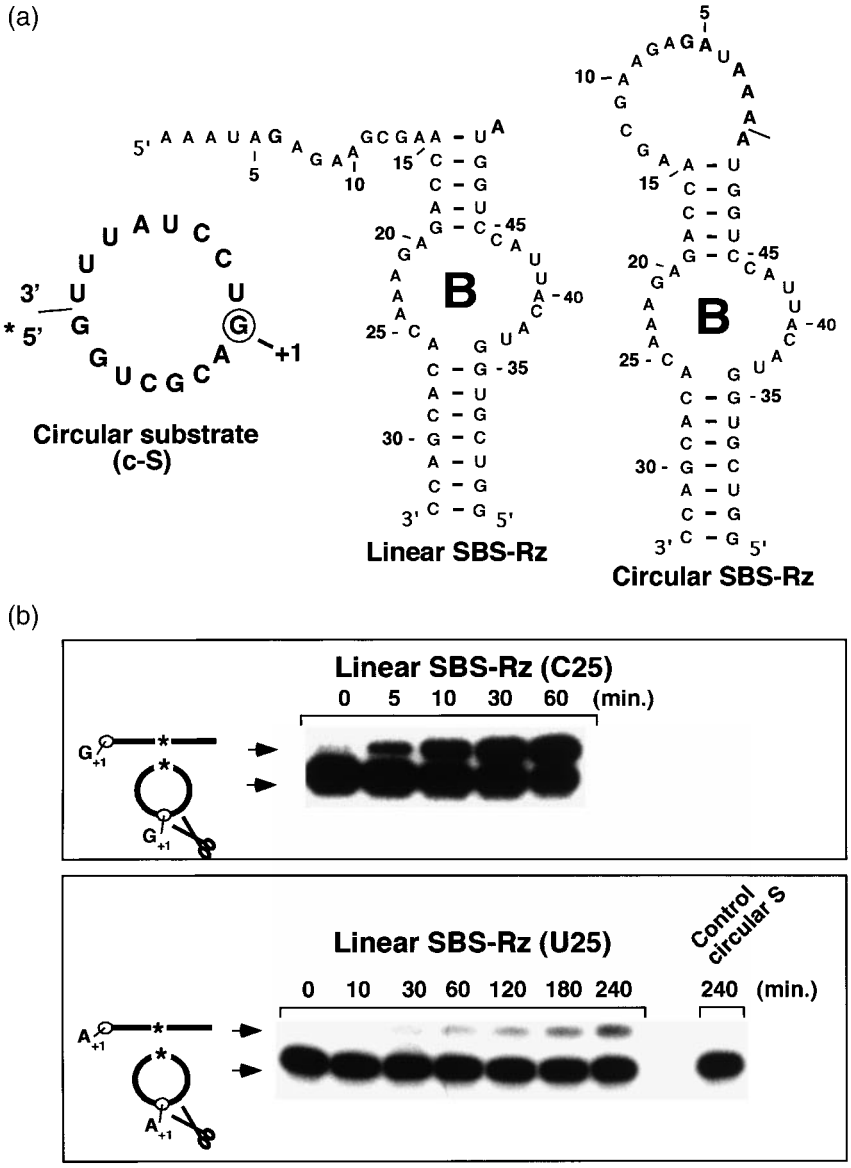


**Figure 6** (*legend opposite*)

ecule and that a specific nucleotide may play the role of the general acid and/or base in the reaction. Remarkably, in the D-shaped model of the substrate-binding domain generated by MC-SYM and subjected to energy minimization, the cleavable linkage is aligned for an in-line attack, ready for a nucleophilic substitution (SN2 mechanism). If the hairpin ribozyme exploits a general acid-base catalytic mechanism in which the RNA is involved, the kink observed in the substrate-binding domain might be required to position the important functional groups. In our model, residue G8 , which is stacked under $A-1$ when the two domains are docked, is located in close proximity to the cleavage site and represents an excellent candidate to participate in the cleavage mechanism. Using G8 variants in combination with structural and biochemical analysis, we have obtained strong evidence supporting the participation of this nucleotide in the active-site chemistry.

We believe that our results provide essential information concerning the organization of the substrate in the active conformation, allowing us to refine our model of the catalytic core. The topographical constraints generated have also permitted us to develop testable hypotheses concerning the global structure of the molecule that might lead to the identification of other essential tertiary contacts between the two domains. But more important, this D-shaped arrangement of the substrate-binding domain has allowed us to predict the importance of G8 and helped us to developed a testable mechanism where the positioning of the functional groups $O-6$ and $N-1$ of G8 as pre-

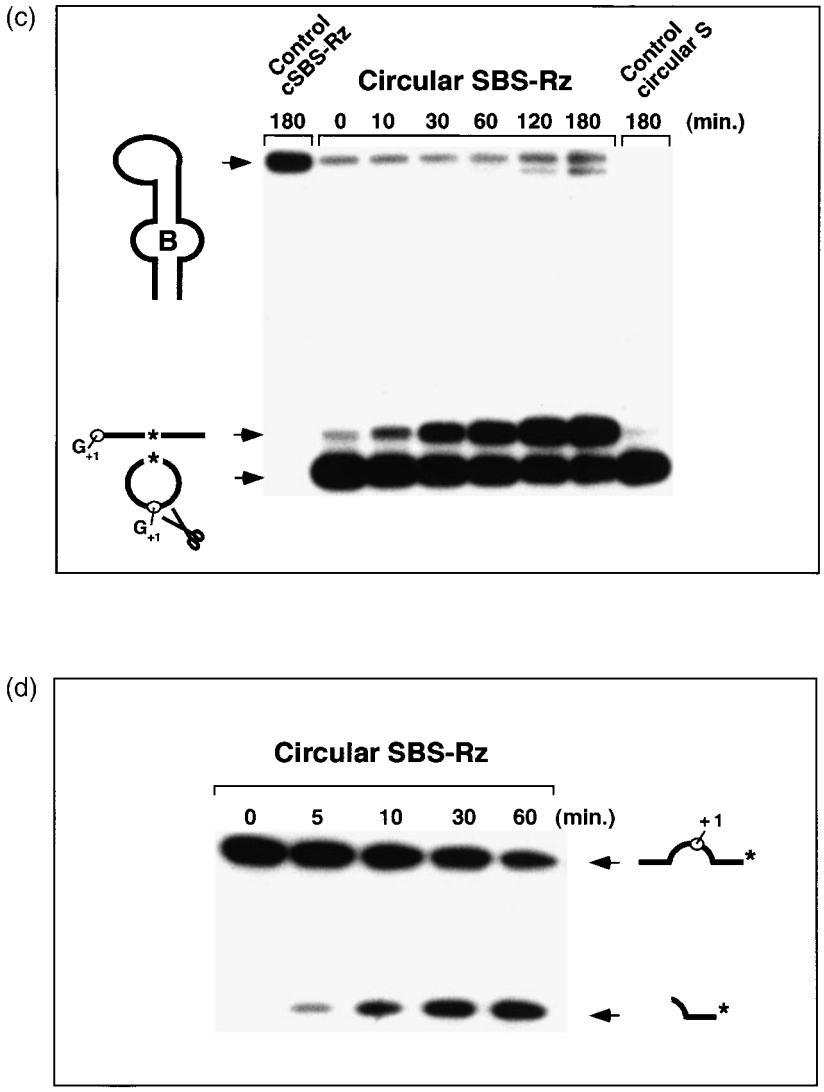


**Figure 6.** Hairpin ribozyme-catalyzed cleavage reaction of linear and circular substrate. Cleavage assays were performed as described (Materials and Methods). (a) Sequence of the 16-mer circularized substrate (c-S) containing two additional G residues to the 3′ end of the regular 14-mer substrate, the two-piece hairpin ribozyme construct with a linear substrate-binding strand (linear SBS-Rz) and a version of the hairpin ribozyme where the substrate-binding strand has been circularized (circular SBS-Rz). (b) Electrophoretic analysis of the cleavage reaction of circular substrate using ribozymes with linear substrate-binding strand (linear SBS-Rz). A version of this construct with a C-to-U mutation at position 25 also was used to cleave a circular substrate with a G-to-A mutation at the +1 position (bottom). The migration positions of the uncleaved circular and its linearized cleavage products are indicated with arrows. Time-course for the cleavage reaction performed with (c) circular substrate or (d) linear substrate using the ribozyme with circularized substrate-binding strand (circular SBS-Rz). Cleaved and uncleaved products as well as the internally labeled ribozyme are indicated with arrows.

dicted in the model presented here, is crucial (unpublished results).

## Materials and Methods

### RNA preparations

Fragments of the hairpin ribozyme and substrate RNAs were synthesized individually using standard phosphoramidite chemistry and PAC-protected β-cyanoethyl phosphoramidites (Glen Research Inc., Sterling VA and Amersham Pharmacia, Piscataway, NJ). Protected alkylthiol-modified phosphoramidites were synthesized as described (Goodwin *et al*., 1996; Maglott *et al*., 1998; Gundlach *et al*., 1997) or *via* modifications of previously reported synthesis strategies (Manoharan *et al*., 1993; C.W.G. & G.D.G., unpublished results). Synthesized RNAs were deprotected and purified by polyacrylamide gel electrophoresis and/or reversed-phase high pressure liquid chromatography as described previously

RNA phosphoramidites and 2′-*O*-methyl ribonucleotide phosphoramidites were purchased from Glen Research, Inc. The alkythiol-modified phosphoramidites

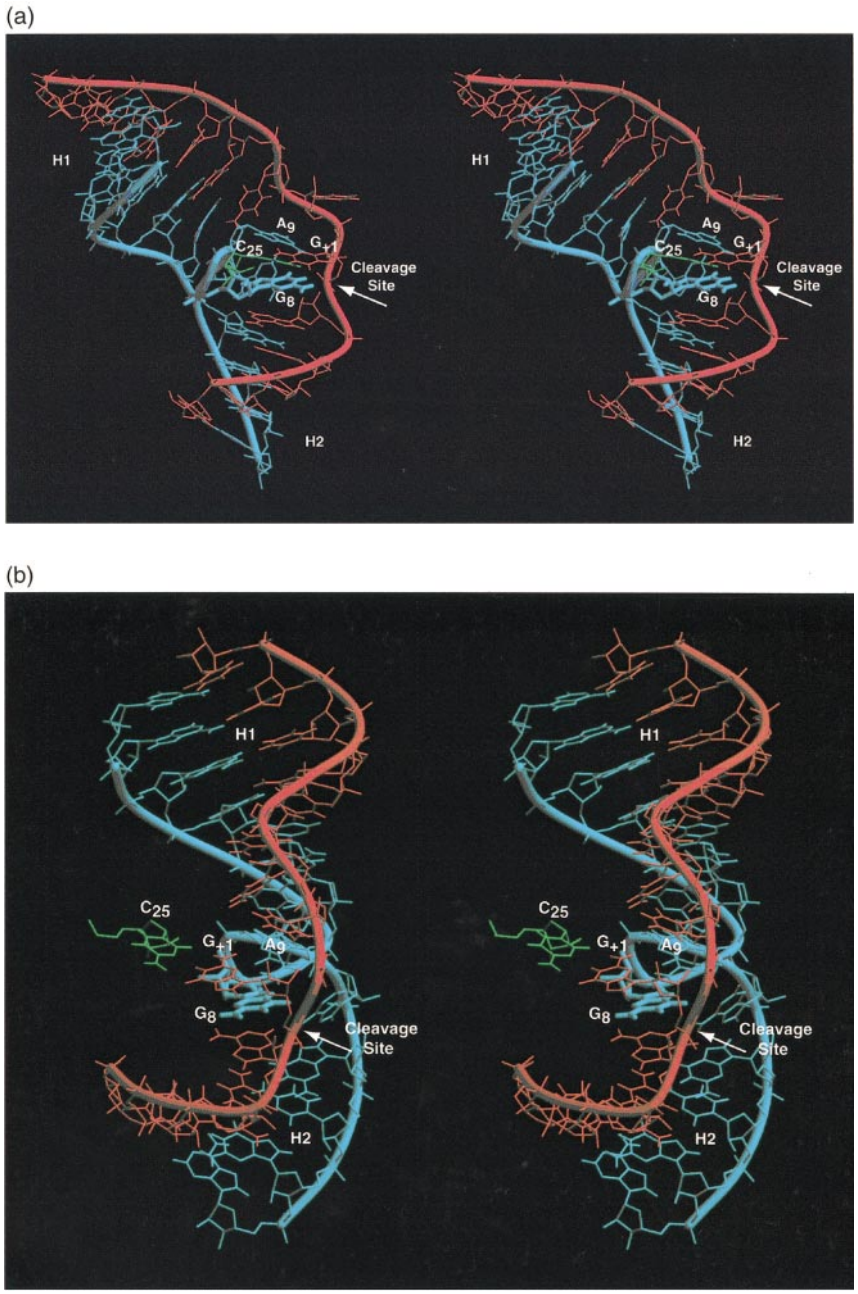


**Figure 7.** Three-dimensional model of the substrate-binding domain. Stereo view showing the proposed D-shape model. The substrate is depicted in red and the substrate-binding strand is in blue. The cleavage site is indicted by an arrow. The nucleotides C25 (in green) and G8 (thicker line) are also indicated. The model was generated by MC-SYM (see Materials and Methods).

were synthesized as described previously. The sulfur groups replacing a non-bridging oxygen atom and used for azidophenacyl coupling, were incorporated during synthesis using 3H-1,2-benzodithiol-3-one, 1,1-dioxide (Glen Research, Inc.). A 2′-*O*-methyl group adjacent to the phosphorothioate was added, to avoid degradation of the synthesized RNA due to the formation of unstable triesters following alkylation of the phosphorothioate group. The azidophenacyl crosslinking agent was coupled to the incorporated sulfur atom as described (Pinard *et al.*, 1999a). Ribozymes with circularized substrate-binding strand were synthesized by transcribing partially duplex synthetic DNA with phage T7 RNA polymerase as described (Milligan & Uhlenbeck, 1989). Transcribed RNA were internally labeled using [$^{32}$P]CTP.

### Preparation of circular substrates

The 5′ end-labeled substrates were incubated with 0.2 unit of phage T4 RNA ligase (Amersham) in ligation buffer (50 mM Tris-HCl (pH 7.5), 10 mM Mg $Cl_2$, 3 mM DTT, 10 % (v/v) DMSO, 10 g/ml BSA and 0.1 mM ATP) for 15 minutes at 37 °C. Reactions were stopped by adding an equal volume of loading solution (97 % (v/v) formamide, 15 mM EDTA). Circular substrates were purified and separated on 7 M urea/20 % polyacrylamide gels.

### Assembly of the ribozyme-substrate complexes and formation of crosslinked species

The assembly of the ribozyme complex or the ribozyme-substrate complex was performed in 50 mM Tris-

HCl (pH 7.5), 12 mM $MgCl_2$. In a tube protected from light exposure, the azidophenacyl containing oligonucleotide (100-200 nM) was incubated with a five- to tenfold molar excess of the appropriate synthetic ribozyme RNA in the presence of cleavable or non-cleavable substrate (deoxy A − 1) for 15 minutes at 37 °C in the reaction buffer. In each reaction, one of the fragments contains the crosslinking agent and one is $^{32}P$ end-labeled. Solutions were allowed to equilibrate for ten minutes at room temperature. The assembled complexes were transferred into a well of a U-bottom flexible assay plate (Becton Dickson). The assay plate was covered with a polystyrene filter and irradiated with 312 nm ultraviolet light (hand-held model VL-6M, IBI Inc. or equivalent) for ten to 20 minutes at room temperature. The crosslinked species were then separated on a denaturing 20 % polyacrylamide gel. The crosslinked species were cut out of the gel and eluted overnight in 0.5 M ammonium acetate, 1 mM EDTA and 0.1 % (w/v) SDS. The crosslinked species were then precipitated with ethanol and passed through a small column to remove salts (CentriSep column, Princeton Separations Inc.). For the non-azido-dependent crosslink, the assembly of the ribozyme complex or the ribozyme-substrate complex was performed with non-modified oligonucleotide strands, essentially as described above using 1 mM $Co(NH_3)_6^{3+}$ instead of $MgCl_2$ in the assembly buffer.

### Time-resolved UV crosslinking

For each time-point, 4 µl of an RNA solution containing $2 \times 10^5$ dpm of end-labeled 3′LB-S, 0.5 µM RzA, 25 mM sodium cacodylate (pH 7), and 0.1 mM EDTA was pipetted into a single well of a polystyrene 96-well plate. Folding was initiated by adding an equal volume of 4 mM $Co(NH_3)_6^{3+}$, 25 mM sodium cacodylate (pH 7), and 0.1 mM EDTA to each well. All wells were irradiated simultaneously at 312 nm with a lamp held at a distance of 2 cm for ten seconds after the prescribed folding time. Thus, the zero time-point was allowed to fold in the presence of $Co(NH_3)_6^{3+}$ during the ten second irradiation, or dead time of the experiment. Samples were separated on 8 M urea/15 % (w/v) polyacrylamide gels and quantified by phosporimager analysis. Three replicate experiments were plotted and fitted simultaneously to a single exponential equation.

### Disulfide crosslink formation

The alkythiol-containing oligos were 5′ end-labeled with [$\gamma^{32}$P]ATP, passed through a CentriSep column (Princeton Separations Inc.) and lyophilized. The *tert*-butyl disulfide protecting groups were reduced with 20 mM DTT in Hepes buffer (20 mM at pH 8.3) for 10-12 hours. The solution was then adjusted to pH 6 and passed through two or three CentriSep columns. After each passage, the filtrate was monitored with *N*-(4-(7-diethylamino-4-methylcoumarin-3yl)phenyl) maleimide (Molecular Probes, Inc.) to determine when complete removal of DTT was achieved (Parvari *et al.*, 1983). The reduced oligos were ethanol-precipitated and the dried pellets resuspended in Tris-HCl (pH 7), 100 mM NaCl. The other strands of the ribozyme-substrate complex were added and the reassembled complex was incubated at 37 °C for 20 minutes. Magnesium was added to a final concentration of 20 mM and the pH was adjusted to 8.0 with sodium hydroxide. The samples were stirred vigorously in air at room temperature using a micro stir bar for eight hours. Crosslinking was conducted using a concentration of 200-400 nM of the reconstituted ribozyme-substrate complex to minimize potential inter-molecular disulfide bond formation. Crosslinked RNAs were ethanol-precipitated, then separated on denaturing 15 % polyacrylamide gel.

### Mapping of the UV crosslinked species

The gel-purified crosslinked RNA species obtained by UV irradiation, either of unmodified or azidophenacyl-modified complexes, were mapped by partial ribonuclease digestion under denaturing conditions and by partial alkaline hydrolysis of 5′ labeled material, as described (Pinard *et al.*, 1999a,b). Partially digested, crosslinked species were loaded onto denaturing 20 % polyacrylamide gels. Non-crosslinked, 5′ end-labeled RNA fragments were subjected to the same treatment and run in parallel to allow the identification of the crosslinked site.

### Reassembly of the ribozyme-substrate complex and cleavage assays

To reconstitute the ribozyme-substrate complex, the purified crosslinked strands were incubated in standard reaction buffer (50 mM Tris-HCl (pH 7.5), 12 mM $MgCl_2$) in the presence of the missing strand(s) (100 nM) for 20 minutes at 37 °C. When the crosslinks did not involve the substrate strand, cleavage reactions were initiated by addition of (5′-$^{32}$P)-end-labeled substrate RNA (100,000 dpm) and were allowed to proceed under standard conditions for one hour at 37 °C. When the crosslinks involved the substrate strand, reactions were initiated by adding the missing ribozyme strands. Aliquots of the reaction were removed and quenched with an equal volume of loading solution (90 % (v/v) formamide, 15 mM EDTA). Samples were loaded onto denaturing 20 % polyacrylamide gels. Radioactive bands were quantified using a Bio-Rad GS-525 radioimaging system.

### Ribozyme cleavage assays

All reactions were performed in 50 mM Tris-HCl (pH 7.5), 12 mM $MgCl_2$ at 25 °C. Ribozymes were preincubated at 37 °C for 15 minutes in reaction buffer. The solutions were equilibrated at 25 °C for ten minutes. Cleavage reactions were initiated by adding 1-2 nM 5′end-labeled substrates. 2 µl aliquots of the reaction were then taken and quenched with 18 µl of loading solution (97 % formamide, 15 mM EDTA). Samples were analyzed as described above. Cleavage reactions were carried out with three different concentration of ribozyme (saturating excess at least 100 times higher than the substrate concentration). Cleavage rates were estimated by non-linear regression using the Origin software (Microcal software, Inc.) The standard error for the fitted curves was typically less than 10 %.

### Molecular modeling using MC-SYM and energy minimization

Models of the ribozyme substrate-binding domain and the catalytic core were generated using the constraints-satisfaction program MC-SYM. The structural constraints to generate an MC-SYM script were derived from the crosslinks described here and the

proposed base-triple interaction and crosslinks already published (Pinard *et al.*, 1999a,b). The average distance was limited to 7.5-12 Å for the azidophenacyl crosslinks and to 5.0-8.0 Å for the disulfide bridge. An SN2-like angle was obtained with a combination of distance constraints. The models generated in this first step were then used as a library and nucleotides $U_{+2}$ and $C_{+3}$ were placed in those libraries. The *A*-RNA form was assumed for all helices and all possible conformations were tested for the pucker modes and glycosyl angles for all the nucleotides in loop A and C25. Similar generated solutions (2 Å difference or less) were combined by MC-SYM. The MC-SYM scripts are available at: http://www-lbit.iro.umontreal.ca/McSym_Repository/

MC-SYM structures were refined using molecular mechanics calculations performed by the molecular simulation program Sander, from the Amber 4.1 suite of programs (Pearlman *et al.*, 1995) using the Amber 94 forcefield. All 1-4 electrostatic interactions were set to a factor of 1.2 and the distance-dependant dielectric model, $\varepsilon = 4R_{ij}$, for the Coulombic representation of electrostatic interactions was used. As a first step, energy minimization has been performed using the steepest descent for 100 steps, then the conjugate gradient method was applied until the maximum derivative was less then 0.1 kcal $mol^{-1}$. The distances of the general acid and base, and the base-pairs assigned by MC-SYM (except the end of helix 2) were restrained and a cutoff value of 15 was used during the minimization.

## Acknowledgments

We thank Gulnar Pothiawala, Sebastien Lemieux, Patrick Gendron, Nancy Bourassa, for technical support, valuable comments and helpful discussions, David Pecchia for the oligonucleotide synthesis. This work is supported by grants from the National Institutes of Health (AI 44186 to J.M.B. and GM52168 to G.D.G.) and from the Medical Research council of Canada (MT 14504 to F.M.). R.P. was supported by a postdoctoral fellowship from the Medical Research Council of Canada.

## References

Behlen, L. S., Sampson, J. R. & Uhlenbeck, O. C. (1992). Light-induced crosslink in yeast tRNA Phe. *Nucl. Acids Res.* **20**, 4055-4059.

Butcher, S. E. & Burke, J. M. (1994a). A photo-cross-linkable tertiary structure motif found in functionally distinct RNA molecules is essential for catalytic function of the hairpin ribozyme. *Biochemistry,* **33**, 992-999.

Butcher, S. E. & Burke, J. M. (1994b). Structure-mapping of the hairpin ribozyme. Magnesium-dependent folding and evidence for tertiary interactions within the ribozyme-substrate complex. *J. Mol. Biol.* **244**, 52-63.

Butcher, S. E., Allain, F. H. & Feigon, J. (1999). Solution structure of the loop B domain from the hairpin ribozyme. *Nature Struct. Biol.* **6**, 212-216.

Burgin, A. B. & Pace, N. R. (1990). Mapping the active site of ribonuclease P RNA using a substrate containing a photoagent. *EMBO J.* **9**, 4111-4118.

Cai, Z. & Tinoco, I., Jr (1996). Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry,* **35**, 6026-6036.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996a). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science,* **273**, 1678-1685.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R. & Doudna, J. A. (1996b). RNA tertiary structure mediation by adenosine platforms. *Science,* **273**, 1696-1699.

Earnshaw, D. J., Masquida, B., Müller, S., Sigurdsson, S. Th., Eckstein, F., Westhof, E. & Gait, M. J. (1997). Inter-domain cross-linking and molecular modeling of the hairpin ribozyme. *J. Mol. Biol.* **274**, 197-212.

Esteban, J. A., Walter, N. G., Kotzorek, G., Heckman, J. E. & Burke, J. M. (1998). Structural basis for heterogeneous kinetics: reengineering the hairpin ribozymes. *Proc. Natl Acad. Sci. USA,* **95**, 6091-6069.

Ferré-d'Amaré, A. & Doudna, J. (2000). Crystalization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binidg protein U1A as a crystallization module. *J. Mol. Biol.* **295**, 541-556.

Goodwin, J. T., Osborne, S. E., Scholle, E. J. & Glick, G. D. J. (1996). Design synthesis and analysis of yeast $RNA^{phe}$ analogs possessing intra- and interhelical disulfide cross-links. *J. Am. Chem. Soc.* **118**, 5207-5214.

Grasby, J., Mersmann, K., Singh, M. & Gait, M. J. (1995). Purine functional groups in essential residues of the hairpin ribozyme required for catalytic cleavage of RNA. *Biochemistry,* **34**, 4068-4076.

Gundlach, C. W., IV, Ryder, T. R. & Glick, G. D. (1997). Synthesis of guanosine analogs bearing pendant alkylthiol tethers. *Tetrahedron Letters,* **38**, 4039-4042.

Hampel, K. J. & Burke, J. M. (2001). Time-resolved hydroxyl-radical footprinting of RNA using Fe(II)-EDTA. *Methods,* in the press.

Hampel, K. J., Walter, N. G. & Burke, J. M. (1998). The solvent protected core of the hairpin ribozyme-substrate complex. *Biochemistry,* **37**, 14672-14682.

Komatsu, Y., Kanzaki, I., Koizumi, M. & Ohtsuka, E. (1995). Modification of primary structure of hairpin ribozymes for probing active conformations. *J. Mol. Biol.* **252**, 296-304.

Komatsu, Y., Kanzaki, I. & Ohtsuka, E. (1996). Enhanced folding of hairpin ribozymes with replaced domains. *Biochemistry,* **35**, 9815-9820.

Maglott, E. J. & Glick, G. D. (1998). Probing structural elements in RNA using engineered disulfide cross-links. *Nucl. Acids Res.* **26**, 1301-1308.

Manoharan, M., Johnson, L. K., Tivel, K. L., Springer, R. H. & Cook, P. D. (1993). Introduction of a lipophilic thioether in the minor groove of nucleic acids for antisense applications. *Bioorg. Med. Chem. Letters,* **3**, 2765-2770.

Milligan, J. F. & Uhlenbeck, O. C. (1989). Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180**, 51-62.

Parvari, D. R., Pecht, I. & Soreq, H. (1983). A microfluorometric assay for cholinesterases suitable for multiple kinetic determination of picomoles of released thiocholine. *Anal. Biochem.* **133**, 450-456.

Pearlman, D., Case, D., Caldwell, J., Ross, W., Cheatham, T., Ferguson, D., Seibel, G., Singh, U., Weiner, P. & Kollman, P. (1995). Amber, a package

of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic proporties of molecules. In *Computer Physics Communications 91*, pp. 1-41, University of California, San Francisco.

Pinard, R., Heckman, J. E. & Burke, J. M. (1999a). Alignment of the two domains of the hairpin ribozyme-substrate complex defined by interdomain photoaffinity crosslinking. *J. Mol. Biol.* **287**, 239-251.

Pinard, R., Lambert, D., Heckman, J. E., Walter, N. G., Major, F. & Burke, J. M. (1999b). Structural basis for the guanosine requirement of the hairpin ribozyme. *Biochemistry,* **38**, 16035-16039.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature,* **372**, 68-74.

Pörschke, D., Burke, J. M. & Walter, N. G. (1994). Global structure and flexibility of hairpin ribozymes with extended terminal helices. *J. Mol. Biol.* **289**, 799-813.

Schmidt, S., Beigelman, L., Karpeisky, A., Usman, N., Sorensen, U. S. & Gait, M. J. (1996). Base and sugar requirements for RNA cleavage of essential nucleoside residues in internal loop B in the hairpin ribozyme: implication for secondary structure. *Nucl. Acids Res.* **24**, 573-581.

Scott, W. G., Finch, J. T. & Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell,* **81**, 991-1002.

Scott, W. G., Murray, J. B., Arnold, J. R. P., Stoddard, B. L. & Klug, A. (1996). Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science,* **274**, 2065-2069.

Sigurdsson, S. T., Tuschl, T. & Eckstein, F. (1995). Probing RNA tertiary structure: interhelical crosslinking of the hammerhead ribozyme. *RNA,* **1**, 575-583.

Tang, R. S. & Draper, D. E. (1994). Bend and helical twist associated with a symmetrical internal loop from 5 S ribosomal RNA. *Biochemistry,* **33**, 10089-10093.

Walter, N. G. & Burke, J. M. (1997). Real-time monitoring of hairpin ribozyme kinetics throuh base-specific quenching of fluorescein labeled substrates. *RNA,* **3**, 392-404.

Walter, N. G. & Burke, J. M. (1998). The hairpin ribozyme: structure, assembly and catalysis. *Curr. Opin. Chem. Biol.* **2**, 24-30.

Walter, N. G., Hampel, K. J., Brown, K. M. & Burke, J. M. (1998). Tertiary formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer. *EMBO J.* **17**, 2378-2391.

Wimberly, B., Varani, G. & Tinoco, I. (1993). The conformation of loop E of Eucaryotic 5 S ribosomal RNA. *Biochemistry,* **32**, 1078-1087.

Wower, J., Hixson, S. S. & Zimmerman, R. A. (1989). Labeling the peptidyl transferase center of the *Escherichia coli* ribosome with photoreactive tRNAPhe derivatives containing azidoadenosine at the 3′ end of the acceptor arm: a model of the tRNA-ribosomes complex. *Proc. Natl Acad. Sci. USA,* **86**, 5232-5236.

Zacharias, M. & Hagerman, P. J. (1996). The influence of symmetrical internal loops on the flexibility of RNA. *J. Mol. Biol.* **257**, 276-289.

***Edited by J. A. Doudna***

*(Received 23 October 2000; received in revised form 12 January 2001; accepted 12 January 2001)*