JMB

Inter-domain Cross-linking and Molecular Modelling of the Hairpin Ribozyme

David J. Earnshaw[1], Benoît Masquida[2], Sabine Müller[1]
Snorri Th. Sigurdsson[3], Fritz Eckstein[3], Eric Westhof[2]
and Michael J. Gait[1]*

[1]*Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK*

[2]*IBMC-CNRS, 15 rue René Descartes, F-67084 Strasbourg France*

[3]*Max-Planck-Institut für Experimentelle Medizin Hermann-Rein-Strasse 3 D-37075, Göttingen, Germany*

The hairpin ribozyme is a small catalytic RNA composed of two helical domains containing a small and a large internal loop and, thus, constitutes a valuable paradigm for the study of RNA structure and catalysis. We have carried out molecular modelling of the hairpin ribozyme to learn how the two domains (A and B) might fold and approach each other. To help distinguish alternative inter-domain orientations, we have chemically synthesized hairpin ribozymes containing 2'-2' disulphide linkages of known spacing (12 or 16 Å) between defined ribose residues in the internal loop regions of each domain. The abilities of cross-linked ribozymes to carry out RNA cleavage under single turnover conditions were compared to the corresponding disulphide-reduced, untethered ribozymes. Ribozymes were classed in three categories according to whether their cleavage rates were marginally, moderately, or strongly affected by cross-linking. This rank order of activity guided the docking of the two domains in the molecular modelling process. The proposed three-dimensional model of the hairpin ribozyme incorporates three different crystallographically determined structural motifs: in domain A, the 5'-GAR-3'-motif of the hammerhead ribozyme, in domain B, the J4/5 motif of group I ribozymes, and connecting the two domains, a "ribose zipper", another group I ribozyme feature, formed between the hydroxyl groups of residues A_{10}, G_{11} of domain A and C_{25}, A_{24} of domain B. This latter feature might be key to the selection and precise orientation of the inter-domain docking necessary for the specific phosphodiester cleavage. The model provides an important basis for further studies of hairpin ribozyme structure and function.

© 1997 Academic Press Limited

Keywords: hairpin ribozyme; RNA folding; RNA modelling; RNA-RNA cross-linking; ribose zipper

*Corresponding author

Introduction

Small catalytic RNAs have been discovered in a number of types of viroid, virusoid and satellite RNAs and are fascinating paradigms of RNA structure and function (Long & Uhlenbeck, 1993; Pyle, 1993; Symons, 1992). The hairpin ribozyme, which is derived from the catalytic centre of the negative strand of the satellite RNA of tobacco ringspot virus, is the second smallest catalytic RNA (reviewed in Burke (1994) and Burke (1996)), but so far there is no crystallographic and very little other structural data available for this ribozyme.

The chemical outcome of hairpin ribozyme action is phosphodiester cleavage at a defined location to generate 5'-hydroxyl and 2',3'-cyclic phosphate moieties on the resultant ends, which is identical to the outcome of hammerhead ribozyme cleavage. Yet the methods of achieving such cleavage appear to be very different. For the hammerhead ribozyme there is strong evidence for participation of magnesium ion both in the deprotonation of the proximal 2'-hydroxyl to generate

the attacking oxyanion (Dahm *et al.*, 1993) and in the lowering of the activation energy of the transition state by binding the *pro*-Rp oxygen of the scissile phosphodiester (Dahm & Unlenbeck, 1991; Koizumi & Ohtsuka, 1991; Slim & Gait, 1991). The direct *pro*-Rp chelation has now been observed directly by an elegant "freeze-trapping" crystallographic technique (Scott *et al.*, 1996). By contrast for the hairpin ribozyme, the metal ion dependency seems to be quite different (Chowrira *et al.*, 1993a) and direct magnesium interaction with a non-bridging oxygen atom of the scissile phosphate appears less likely (Young *et al.*, 1997; D.J.E. & M.J.G., unpublished results).

A second difference is structural. Whereas the hammerhead catalytic core is formed by folding of a 3-way junction connecting RNA duplexes, two of which are coaxially stacked (Pley *et al.*, 1994; Scott *et al.*, 1995), the hairpin attains its catalytically competent state by docking of two apparently separable domains, each of which consists of a pair of helices interposed by a region of internal loop (Figure 1). In the *trans*-cleaving mode, domain A is formed by annealing of the substrate strand to one section of the ribozyme strand to form helices 1 and 2. Domain B is formed by folding back of a second part of the ribozyme strand onto itself. This generates helices 3 and 4 as well as a terminal three-base loop. Domain B can also be formed

from two annealed strands of RNA if helix 4 is extended by three base-pairs and the terminal loop is omitted (Chowrira *et al.*, 1993b; Grasby *et al.*, 1995).

From studies where the two domains are constrained by linkers of variable length between residue −5 on the substrate strand and residue 50 on the ribozyme strand (Feldstein & Bruening, 1993; Komatsu *et al.*, 1994), it appears that these domains are hinged at the junction of helixes 2 and 3 (between residues 14 and 15) such that the ribozyme can bend and the two internal loop regions can approach each other (Figure 1). By linking the two domains in a completely different "reverse" configuration (i.e. internucleotide bonds 14-15 and 30-31 in the ribozyme strand are broken and residue 1 linked to residue 30 *via* a number of cytidine residues) it was also found possible to maintain ribozyme activity (Komatsu *et al.*, 1995). More recently it has been shown that the two domains can be completely separated into independent sections and ribozyme activity reconstituted by addition of one to the other, albeit at high magnesium ion concentration (Butcher *et al.*, 1995; Komatsu *et al.*, 1996; Shin *et al.*, 1996). Whereas separation of the domains results in a 10^4-fold increase in K_M, there is remarkably little effect on k_{cat} (Butcher *et al.*, 1995). Attachment of a complementary pairing arm to an end of each domain, such that base-pairing brings the two domains into contact, allows ribozyme activity under more standard magnesium ion concentration (12 mM) (Komatsu *et al.*, 1996). All these results suggest that tertiary interactions are required between the two domains to achieve catalysis.

A second conclusion that can be drawn from these experiments is that each domain must undergo an initial folding process before the two domains approach. These "ground state" structures subsequently may then become altered as the two domains approach to attain a catalytically competent state. A solution structure for isolated domain A has been proposed based on NMR spectroscopic measurements (Cai & Tinoco, 1996). Here the sequence of the domain within the looped region is similar to the natural hairpin except for an A to C alteration at position −1 in the substrate, the identity of which is non-essential (Hampel *et al.*, 1990; Joseph *et al.*, 1993). The NMR data are consistent with a sheared base-pair between G_{+1} and A_9 and a wobble pair between a protonated A_{10} and C_{-1}. In addition, the bulging of U_{+2} allows the stacking of A_9 upon C_{+3} and hydrogen-bonding between G_8 and the ribose rings of G_{+1} and U_{+2} (Figure 1).

Although no NMR data is yet available for the structure of domain B, other structure-probing techniques have been utilized. A detailed study of the accessibilities of the ribozyme to chemical modifying reagents under different conditions suggested that domain B folds independently of substrate binding (Butcher & Burke, 1994b). Another informative finding was that a stable UV
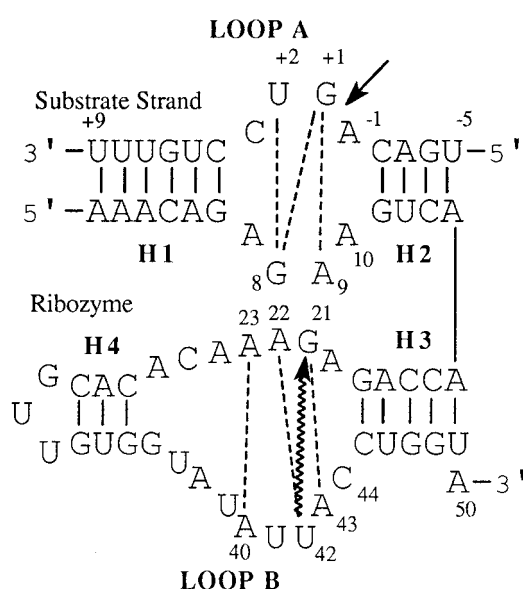
Figure 1. Secondary structure of the catalytic core of the (−)sTRSV hairpin ribozyme. The four helixes (H1 to H4) are represented with the two internal loops (A and B). Postulated base-pairs within loops A and B are depicted by broken lines (Butcher & Burke, 1994b; Cai & Tinoco, 1996). Wavy arrow represents a UV cross-link observed between positions U_{42} and G_{21} (Butcher & Burke, 1994a). Plain arrow indicates the cleavage site. Note that residues on the substrate strand are denoted with plus or minus signs on either side of the cleavage site, whereas residues on the ribozyme strand are numbered without sign.

cross-link could be formed between G_{21} and U_{41} (Butcher & Burke, 1994a), later reassigned to U_{42} (Burke, 1996). The cross-linked domain was reduced about 100-fold in ribozyme activity but the results suggested nevertheless a close positioning of G_{21} and U_{42}. The nucleotide sequence in this region was found to be very similar to that of other UV sensitive domains such as those found in the loop E of eukaryotic 5 S RNA and in the conserved C domain of viroids. From these studies Butcher and Burke proposed a partial secondary structure model of this section of loop B involving non-canonical base-pairs between $G_{21}:A_{43}$, $A_{22}:U_{42}$ and $A_{23}:A_{40}$ (Figure 1; Butcher & Burke, 1994a,b).

We now wished to obtain a preliminary molecular model of how the two domains of the hairpin ribozyme might fold and orient towards each other as they approach the active configuration. In order to distinguish alternative possibilities for docking of the two domains, we have utilized a cross-linking procedure to tether particular nucleoside residues in domain A to other nucleoside residues in domain B. Because the chemistry of such cross-linking is defined precisely, we hoped that the tethers could act as a "molecular ruler" to help distinguish those residues which are severely distorted by cross-linking and which therefore have difficulty in attaining an active configuration from those residues where cross-linking results in little or no distortion. The cross-linking data might thus help distinguish alternative docked conformations.

Results

3D Modelling of the hairpin ribozyme

The hairpin ribozyme is composed of two domains each consisting of an internal loop (A or B) flanked by two regular A-helixes (H1, H2 and H3, H4, respectively). The strategy for modelling was divided into two steps: (i) individual modelling of domains A and B for the 3-stranded ribozyme used for cross-linking (Figure 2); followed by (ii) docking of the two domains. The first step could be carried out with the help of existing data in the literature, whereas the second step of docking required further data to help distinguish alternative possibilities for interdomain orientation. These additional distance constraints were obtained from cross-linking experiments (see below).

Individual modelling of loops A and B

Loops A and B were modelled starting from NMR structural models (Cai & Tinoco, 1996; Wimberly et al., 1993) and then integrated between H1-H2 and H3-H4, respectively. The modelling uncertainties resulting from sequence divergences between our system and those studied by NMR could be circumvented by considering the available probing data (Butcher & Burke, 1994b). These data provide guidelines for checking the validity of
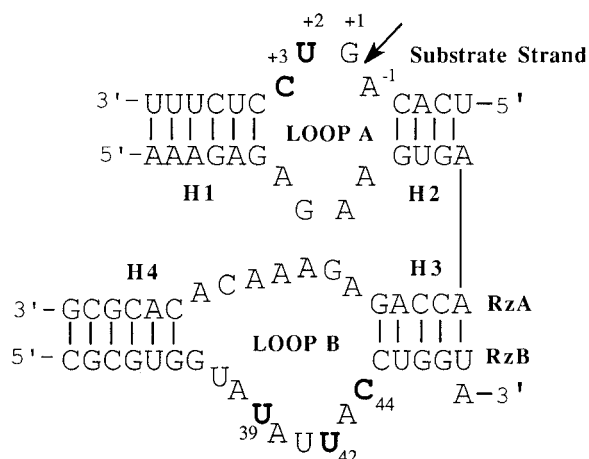

Figure 2. Secondary structure of the 3-stranded hairpin ribozyme in *trans*-cleaving mode used as the basis for molecular modelling and cross-linking. The positions highlighted, two on the substrate strand (U_{+2} and C_{+3}) and three on the ribozyme strand B (U_{39}, U_{42} and C_{44}), are the sites for the specific incorporation of corresponding 2′-amino-2′-deoxy pyrimidine nucleosides. These positions (one on the substrate strand and one on the RzB strand) are cross-linked *via* an aryl disulphide or alkyl disulphide linkage. The arrow indicates the site of cleavage.

each model in a whole ribozyme context, since the sequences of loops A and B studied by probing match perfectly those of our system. However, the modelling of internal loops with non-canonical pairings is fraught with pitfalls and, by necessity, requires decisions based on specific structural hypothesis. As a general rule, we have tried to incorporate previously identified RNA motifs. The models of each domain are shown in Figures 3 and 4.
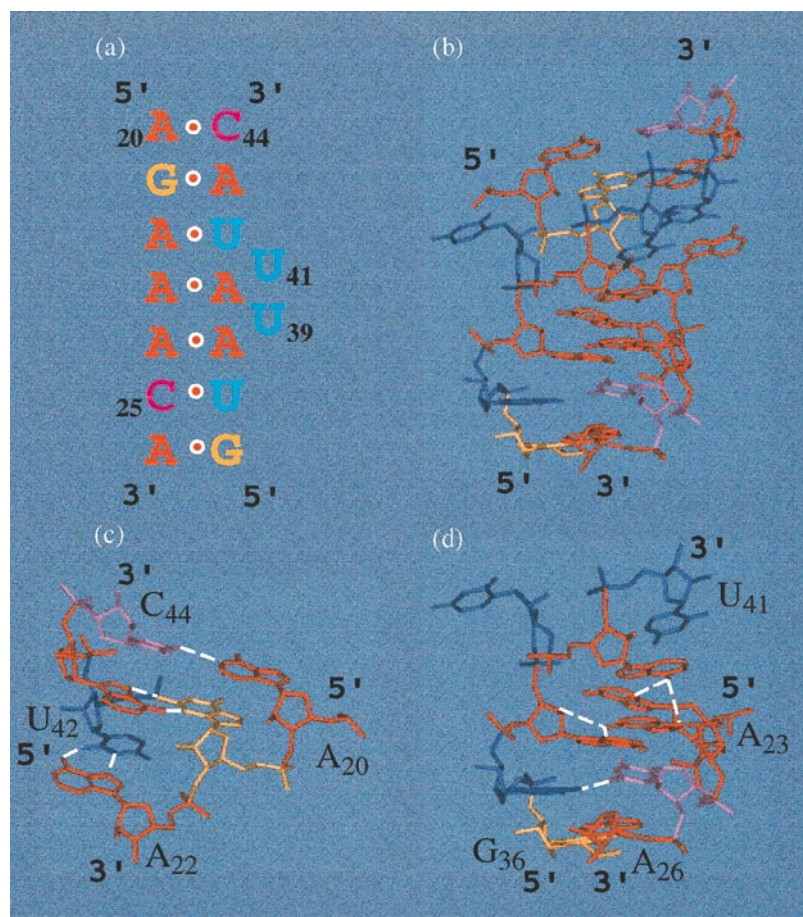
Loop A

Although the sequence of loop A in the present work differs from that studied by NMR (Cai & Tinoco, 1996) at position "−1", where an A instead of a C is present, the probing experiments yield a protection pattern compatible with the stacking of A_{10} between the "sheared" base-pair A_9 (N6, N7)·G_{+1} (N3, N2) and the H2-closing Watson−Crick base-pair $C_{-2}·G_{11}$. Although positions 10 and −1 do not covary (Chowrira & Burke, 1991), A_{-1} and A_{10} could interact in a non Watson−Crick manner, in agreement with NMR data which suggest stacking between these positions and the adjacent canonical pair (Cai & Tinoco, 1996). The resulting pair involves the N6 atom of A_{10} with the N3 and O2′ atoms of A_{-1}, a sheared A·A pair as seen recently by crystallography (Cate et al., 1996). Interestingly, it is almost isosteric to an A·C with H-bond contacts between $A_{10}(N6)$ and $C_{-1}(O2, O2′)$. The geometry of the $A_{10}·A_{-1}$ base-pair optimizes stacking interactions between A_{10} and A_9 on the ribozyme strand as
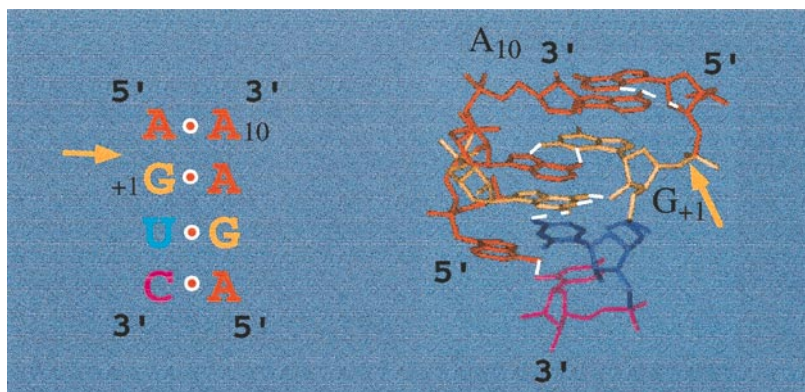
well as between A_{-1} and G_{+1} on the substrate strand. At the same time, it buries A_{10} (N7), in agreement with the observation that this residue is the least sensitive to modifying agents in loop A (Butcher & Burke, 1994b). It is worth noting that the small distance between atoms G_{+1} (O3′) and A_9 (P) prevents residues U_{+2} and G_8, both adjacent to the sheared pair, from interacting in a canonical way (Gautheret *et al.*, 1994). The formation of the

sheared pair thus forces U_{+2} to bulge out while G_8 remains stacked within the helix (Figure 3), providing an explanation for the cross-strand contacts of G_8 with G_{+1} and U_{+2} and for the kink in the helix, which both generate intricate NMR contacts between C_{+3} and A_9 (Cai & Tinoco, 1996). Moreover, the introduction of a bulge into the structure confers on loop A a dynamic motion in which a flip of the bulge could decrease or increase the

kink. This could play some role during the docking of loop A with loop B. The subsequent insertion of loop A between H1 and H2 follows naturally from the geometry of the first and terminal base-pairs of the loop which places the connecting atoms 5′-P and O3′ within a distance of 18 Å, compatible with standard helical RNA.

Loop B

The base-pairing schemes of internal loop B were settled by taking advantage of the NMR structural models of the loop E of eukaryotic 5 S rRNA (Wimberly *et al.*, 1993) and of the sarcin/ricin loop from the 23 S rRNA (Szewczak *et al.*, 1993) following the sequence similarities in the common UV cross-link loops (Butcher & Burke, 1994a). The UV cross-link motif was modelled accordingly to form a sheared base-pair between residues G_{21} and A_{43} and a Hoogsteen-*trans* $A_{22} \cdot U_{42}$ (Figure 4). This choice is corroborated by probing results (Butcher & Burke, 1994b). Furthermore, the lower reactivity of the Watson–Crick sites of A_{20} compared to the N7 modification would argue for a Watson–Crick pair with C_{44}, involving atoms N6 and N1 (in the protonated state) atoms of A_{20} towards atoms N3 and O2 atoms of C_{44}, since these bases are inserted between two helical base-pairs ($G_{19} \cdot C_{45}$ and $G_{21} \cdot A_{43}$). But any type of H-bonded, neutral *cis* base-pair, such as the $A_{20}(N1) \cdot C_{44}(N4)$ present in the model, could fit without raising geometrical problems, in agreement with recently published results (Siwkowski *et al.*, 1997).

For the remaining part of loop B, we exploited the effects of nucleotide analogues on hairpin catalysis (Grasby *et al.*, 1995; Schmidt *et al.*, 1996a) as well as structure probing results (Butcher & Burke, 1994b). These data provide a basis for the proposal of an original interaction scheme which allows for the formation of tandem sheared $A \cdot A$ pairs, respectively for $A_{38} \cdot A_{24}$ and $A_{40} \cdot A_{23}$, in a geometry similar to that of the J4/5 junction recently seen in the crystal structure of a portion of a group I ribozyme (Cate *et al.*, 1996). The unpaired base in the consensus of J4/5 (A_{115} in the *Tetrahymena thermophila* group I intron) corresponds to the residue proposed to bulge within the narrow groove of the helix in both NMR structural models of the UV-crosslink loops (Szewczak *et al.*, 1993; Wimberly *et al.*, 1993), thus giving further support for unpairing this residue. U_{39} is the candidate in the hairpin ribozyme loop B for bulging out, since it is strongly reactive to the carbodiimide reagent CMCT (Butcher & Burke, 1994) and can be replaced by a propyl linker without affecting the catalytic process (Schmidt *et al.*, 1996a). On the side of helix H4, the proposed interaction between A_{26} and G_{36} can be mediated theoretically both in canonical Watson–Crick or in sheared geometries (Gautheret *et al.*, 1994). Nonetheless, the protection patterns of these residues are more consistent with a Watson–Crick rather than a sheared geome-

try, since both the N1 and N7 atoms of these residues are protected in the native ribozyme state (Butcher & Burke, 1994b). An indirect argument results from the geometry of the next base-pair, C_{25} (N4)$\cdot U_{37}$ (O4), which contains only one H-bond and which exposes to the solvent the N3 atoms of both residues, as seen by chemical probing. We can conclude from the individual modelling of loops A and B that, once sandwiched between their helical partners, they constitute extended duplexes with local backbone deformations resulting from their unusual pairing schemes.

Comparison of the theoretical and experimental accessibilities of loops A and B models

As a check of the models of loops A and B in the hairpin ribozyme, we have investigated the correlation between the reactivities in solution of the bases in the loops (Butcher & Burke, 1994b) and the theoretical accessibilities (Figure 5). The Watson–Crick data correlate very well for the unmodified or slightly modified residues as well as for the bulging residue U_{39}, which shows one of the strongest modification rates (Figure 5A). Furthermore, the other theoretically accessible N1 atoms all belong to adenines which are involved in non Watson–Crick pairings in our model (A_{22}, A_{23}, A_{24} and A_{43}), even though the variations in the theoretical accessibilities do not strictly follow the distribution of the experimental data. Nevertheless, the important criterion is that these N1 atoms are not involved in contact, which gives the possibility for molecular motions to modify the distribution of their accessibilities (Brunel, 1992).

For the Hoogsteen data (Figure 5B), the average theoretical accessibilities are low as in the experimental data (except for residue A_{23}, for which the associated surface of 13.6 Å seems to be greater than the corresponding experimental reactivity). Nonetheless, the pattern of protection of adenines 23, 24, 38, and 40 is compatible with the motif presented in our model which involves these nucleotides, i.e. A_{23} is more reactive in solution than are either A_{38} or A_{40}, whilst A_{24} is at background level. A noteworthy remark is that the sheared pair $A_{23} \cdot A_{40}$ is capped by bulging residues, the dynamics of which would modify the shielding of A_{23} and A_{40}.

Inter-domain cross-linking of the hairpin ribozyme

In order to provide data to help in the second step of molecular modelling, docking of the two domains, inter-domain cross-linking was carried out between the two internal loops A and B. Since very many functional group modifications in the heterocyclic bases, especially in loops A and B, were found to be highly deleterious to ribozyme activity (Chowrira & Burke, 1991; Grasby *et al.*,
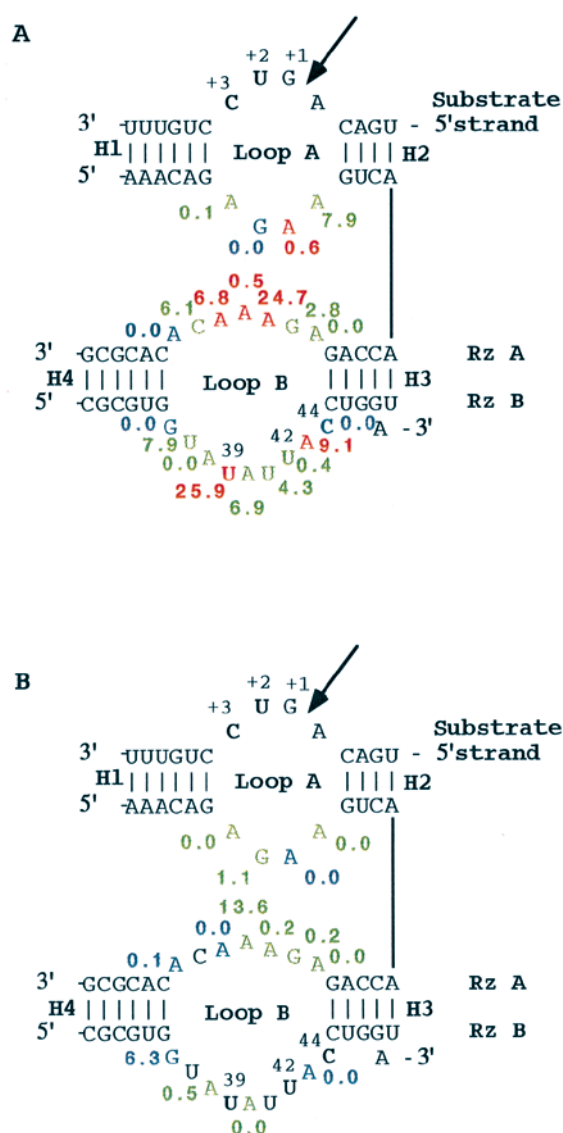
Figure 5. Secondary structure of the hairpin ribozyme showing the correspondence between the theoretical accessibilities (expressed in $Å^2$) and the chemical accessibilities at (A) Watson–Crick positions or at (B) Hoogsteen positions. The colour code represents the strengths of the hits obtained in chemical probing experiments (Butcher & Burke, 1994). Strong hits: red; moderate hits: green; marginal hits: blue. For two residues (A_{10} and U_{39}), docking of the two domains results in an accessibility decrease at the Watson–Crick sites, respectively, from 7.9 to 1.8 and 25.9 to 16.1 $Å^2$.

domain cross-linking. In a previous study it was found that in only four cases did single 2′-deoxynucleoside substitutions in the ribozyme strand of the hairpin result in serious loss of catalytic efficiency (Chowrira *et al.*, 1993b; Schmidt *et al.*, 1996a). A new procedure has been described for inter-helical cross-linking of the hammerhead ribozyme in which the 2′-positions of two specific pyrimidine nucleosides were each substituted by 2′-amino-2′-deoxynucleoside residues and then modified by reaction with an aryl isothiocyanate containing a 2-pyridyl disulphide. Reduction and subsequent oxidation allowed the formation of aryl disulphide linkages between the 2′-positions. These cross-links, which are expected to span a distance of some 16 Å when fully extended, were useful in distinguishing alternative models for the folding of the hammerhead ribozyme (Sigurdsson *et al.*, 1995a,b). We decided to apply this cross-linking strategy to a three-stranded hairpin ribozyme that we have used in previous studies (Grasby *et al.*, 1995; Schmidt *et al.*, 1996a). Note that recently a different disulphide cross-linking technique has been described to link the 2′-hydroxyl positions of uridine residues inter-helically in yeast tRNA[Phe] (Goodwin *et al.*, 1996).

We chose for cross-linking three pyrimidine nucleosides in ribozyme strand B (U_{39}, U_{42} and C_{44}) which form part of the catalytically essential loop B (Figure 2). These are all sites where we had previously found that 2′-deoxynucleoside substitution had no significant effect on cleavage (Schmidt *et al.*, 1996a). Most of the pyrimidine nucleosides in domain A are located in the substrate strand, but since no data is available about the sugar requirements in this strand, we chose two pyrimidine nucleosides in the catalytically important loop A (U_{+2} and C_{+3}). In each case ribozyme-directed cleavage would result in release of the identical unconstrained pentanucleotide, which would be expected to be able to dissociate from the ribozyme strand in the same way as for a non-cross-linked substrate.

Three-stranded hairpin ribozymes were prepared by chemical synthesis and subsequent annealing in which two of the strands (substrate and ribozyme B) each carried single 2′-amino-2′-deoxynucleoside substitutions. The catalytic parameters of six such doubly functionalized pre-annealed hairpin ribozymes were measured under single turnover conditions and compared with those of the unmodified ribozyme (Table 1). The results showed that the K'_M values were in each case within threefold of the umodified ribozyme. Four of the six catalytic rates (k'_{cat}) were within threefold of the unmodified ribozyme. Only two (U_{+2}-U_{42} and U_{+2}-C_{44}) were slightly reduced (fivefold). The results suggested that doubly 2′-amino-functionalized hairpin ribozymes were not greatly affected in cleavage rates compared to the unmodified hairpin.

The chemistry of cross-linking is summarized in Figure 6. Essentially, the two strands containing 2′-

1995; Schmidt *et al.*, 1996a), it was decided not to use cross-linking methods that would involve the heterocyclic bases, for example photochemical cross-linking initiated by introduction of 4-thio-U, 6-thio-G or 6-thio-I residues (Favre & Fourrey, 1995; Vitorino Dos Santos *et al.*, 1993) or interhelical disulphide linkages between uridine (Goodwin & Glick, 1994) or cytidine (Allerson & Verdine, 1995) residues.

By contrast, the 2′-positions of nucleoside sugars seemed particularly appropriate sites for inter-

Table 1. Kinetic data for the cleavage of the unmodified and the doubly functionalized 2'-amino-2'-deoxy ribozymes hairpin ribozyme under single turnover conditions

Ribozyme	K'_M [μM]	k'_{cat} [min^{-1}]	k'_{cat}/K'_M
Unmodified	0.094 ± 0.006	0.16 ± 0.01	1.7
2'-amino U$_{+2}$/U$_{39}$	0.046 ± 0.009	0.096 ± 0.008	2.1
U$_{+2}$/U$_{42}$	0.033 ± 0.007	0.028 ± 0.002	0.85
U$_{+2}$/C$_{44}$	0.12 ± 0.04	0.027 ± 0.005	0.23
C$_{+3}$/U$_{39}$	0.055 ± 0.017	0.12 ± 0.02	2.1
C$_{+3}$/U$_{42}$	0.15 ± 0.05	0.099 ± 0.018	0.66
C$_{+3}$/C$_{44}$	0.17 ± 0.067	0.067 ± 0.011	0.39

amino-2'-deoxynucleosides (the substrate strand being ^{32}P-labelled) were first reacted with 3-isothiocyanatobenzyl 2-pyridyl disulphide (Linking agent A) to give the corresponding thiourea derivatives. After reduction with dithiothreitol (DTT) and subsequent air oxidation, intermolecular aryl disulphide bonds were formed. The hetero-linked substrate-ribozyme B was separated from homo-linked strands and unreacted strands by denaturing polyacrylamide gel electrophoresis, in a similar manner to the separation of cross-linked hammerhead RNAs (Sigurdsson & Eckstein, 1996a). In a preliminary communication describing our first cross-linking attempts, cross-linking was carried out in the presence of an inactivated ribozyme A strand (Schmidt *et al.*, 1997). Here we found that there was no need to form an annealed structure for cross-linking. If an initial ratio of 1:5 of substrate to ribozyme B strand was used, acceptable conversions into hetero-linked substrate-ribozyme B could be obtained (6 to 10%), which were isolated by denaturing polyacrylamide gel electrophoresis in 1 to 5% overall yields.

Kinetic Parameters of disulphide cross-linked and reduced hairpin ribozymes

Aryl disulphide cross-linked substrate-ribozyme B complexes were prepared from all six possible combinations of the 2'-amino-2'-deoxynucleoside substituted strands. For each combination, experiments were then carried out in parallel to measure the kinetic parameters under single turnover conditions. In each case ribozyme A strand was carefully pre-annealed in the absence of magnesium ions to either (a) the cross-linked substrate-ribozyme B complex, or (b) substrate-ribozyme B complex treated with DTT to open the disulphide linkage (Figure 7). Ribozyme cleavage was then effected by addition of magnesium ions. Since neither the folding of domain B nor formation of the substrate-ribozyme complex is reported to be significantly affected by magnesium ions (Butcher & Burke, 1994b; Cai & Tinoco, 1996), the foldings and hence the reaction pathways of the cross-linked and the DTT-treated hairpins should be identical.
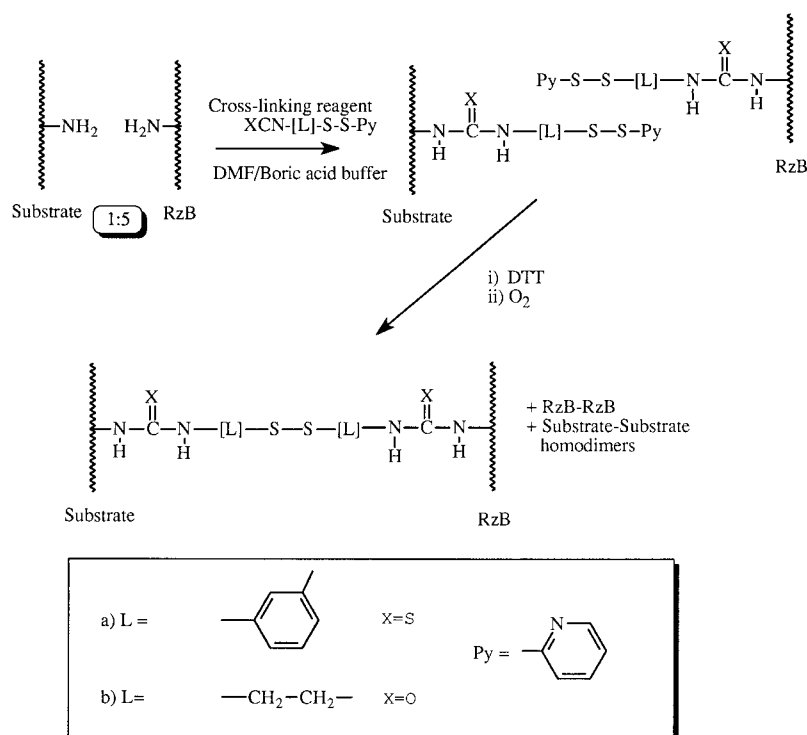
Figure 6. Scheme for the cross-linking of substrate strand and RzB strand each substituted by a 2'-amino-2'-deoxynucleoside.
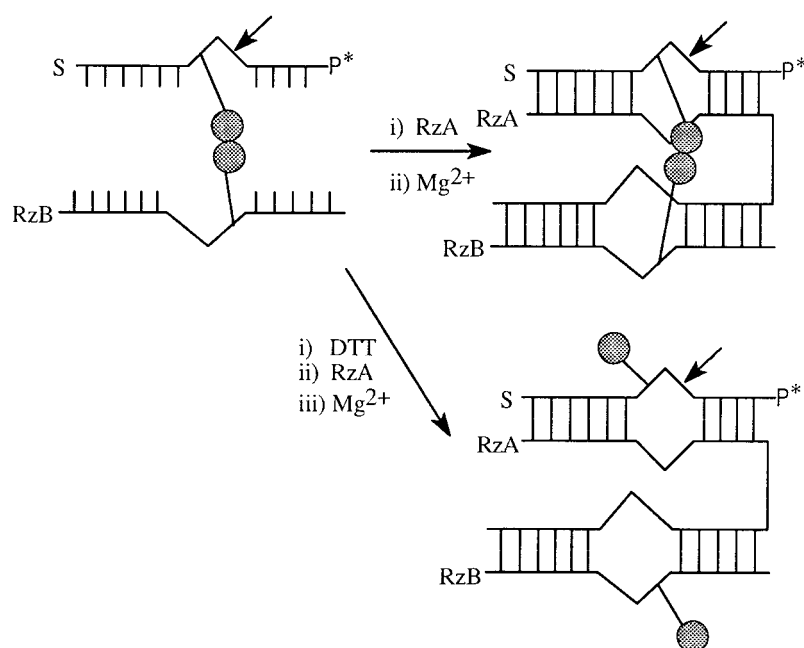
Figure 7. Annealing of the cross-linked substrate-RzB strands with RzA strand in the absence (top right) or in the presence (bottom right) of dithiothreitol (DTT) to break the constraint of the cross-link. Each type of hairpin (cross-linked and uncross-linked) is then incubated with magnesium ion in order to effect ribozyme cleavage. P* denotes $5'-^{32}$P-radiolabelling.

The kinetic parameters for the aryl disulphide cross-linked and DTT-treated ribozymes are shown in Table 2A. First, for each set of paired data (cross-linked and DTT-treated), the K'_M values are within a twofold range. Secondly, the K'_M values in all cases (both cross-linked and DTT-treated) are within a factor of 3 of the completely unmodified three-stranded hairpin (Table 1). Thirdly, pairwise comparisons of the DTT-treated ribozymes (Table 2A) with their cor-

responding 2′-amino substituted ribozymes (Table 1) are also within a twofold range, except for C_{+3}-U_{42} which differ only by a factor of 3. All these results confirm that correct folding has taken place irrespective of the types of substitution or number of annealed strands, and that the reaction pathways of the cross-linked and DTT-treated ribozymes are identical.

Comparison of the catalytic rates (k'_{cat}) of the ribozyme pairs (cross-linked *versus* DTT-treated)

Table 2. Kinetic data for cleavage

Sub	RzB	DTT	K'_M [μM]	k'_{cat} [min^{-1}]	k'_{cat}/K'_M
A. Aryl disulphide cross-linked and reduced hairpin ribozymes					
U_{+2}	U_{39}	−	0.10 ± 0.04	0.00075 ± 0.00019	0.0074
		+	0.049 ± 0.001	0.0021 ± 0.0002	0.043
	U_{42}	−	0.097 ± 0.029	0.00010 ± 0.00001	0.0011
		+	0.047 ± 0.003	0.00021 ± 0.00003	0.0045
	C_{44}	−	0.076 ± 0.014	0.00045 ± 0.00004	0.0061
		+	0.15 ± 0.02	0.030 ± 0.003	0.21
C_{+3}	U_{39}	−	0.028 ± 0.004	0.00098 ± 0.00005	0.035
		+	0.051 ± 0.009	0.035 ± 0.004	0.69
	U_{42}	−	0.049 ± 0.003	0.00079 ± 0.00001	0.016
		+	0.035 ± 0.009	0.0039 ± 0.0004	0.11
	C_{44}	−	0.11 ± 0.01	0.0020 ± 0.0001	0.018
		+	0.12 ± 0.02	0.066 ± 0.004	0.57
B. Alkyl disulphide cross-linked and reduced hairpin ribozymes					
U_{+2}	U_{39}	−	0.051 ± 0.018	0.0023 ± 0.0004	0.045
		+	0.034 ± 0.005	0.0046 ± 0.0003	0.13
	U_{42}	−	0.088 ± 0.018	0.00011 ± 0.00001	0.0013
		+	0.13 ± 0.05	0.00049 ± 0.00012	0.0036
	C_{44}	−	0.051 ± 0.026	0.00010 ± 0.00002	0.0017
		+	0.024 ± 0.006	0.0088 ± 0.0005	0.36
C_{+3}	U_{39}	−	0.029 ± 0.006	0.0014 ± 0.0001	0.046
		+	0.034 ± 0.014	0.0072 ± 0.0012	0.21
	U_{42}	−	0.36 ± 0.04	0.00095 ± 0.00009	0.0026
		+	0.099 ± 0.014	0.012 ± 0.001	0.12
	C_{44}	−	0.091 ± 0.009	0.00010 ± 0.000004	0.00077
		+	0.12 ± 0.03	0.022 ± 0.003	0.18

Table 3. Ratios of k'_{cat} values of the cross-linked k'_{cat}/reduced ribozyme k'_{cat}

k'_{cat} rel	U_{39} ARYL	U_{39} ALKYL	U_{42} ARYL	U_{42} ALKYL	C_{44} ARYL	C_{44} ALKYL
U_{+2}	0.33	0.50	0.50	0.23	0.015	0.010
C_{+3}	0.028	0.19	0.20	0.081	0.031	0.0032

shows that two of the cross-links (U_{+2}-U_{39} and U_{+2}-U_{42}) show only modest reductions in catalytic rate, whereas U_{+2}-C_{44} and all three cross-links originating from C_{+3} in loop A resulted in significant or strong (5 to 70 fold) losses in catalytic activity. Another interesting comparison is between the k'_{cat} values for the DTT-treated ribozymes (Table 2) and the corresponding 2′-amino substituted ribozymes, which lack the thiourea pendant arms (Table 1) which might thus be expected to behave similarly. Two are insignificantly altered, two are slightly reduced, whereas two are substantially reduced (25 and 100-fold, respectively). These results suggested that the aromatic thiourea substituent might be exerting in certain cases some structural effects on the ability to reach the transition state. We therefore decided to construct and measure the kinetic parameters of a second series of cross-linked hairpin ribozymes using the same locations of 2′-amino-2′-deoxynucleoside tether points, but with a shorter and more flexible, non-aromatic, alkyl disulphide linker.

It was recently shown that 2′-amino groups on oligoribonucleotides could be site-specifically reacted with a short aliphatic (two-carbon) isocyanate containing a protected thiol function to which reporter groups could be attached (Sigurdsson & Eckstein, 1996b). We thus prepared analogous but more flexible alkyl disulphide linkages between the two hairpin domains by reaction of 2′-amino groups with 2-isocyanatoethyl 2-pyridyl disulphide, reduction and subsequent oxidation (Figure 6, Linkage agent B). This alkyl disulphide linker would be expected to span about 12 Å, slightly shorter than the aryl disulphide linker. Cross-linking, purification of ribozyme B-substrate cross-links and annealing with ribozyme A strand in the presence or absence of DTT was carried out as before. Ribozyme cleavages were again effected by addition of magnesium ions and the kinetic parameters were measured under single turnover conditions (Table 2b).

Once again pairwise inspection showed that all of the combinations the K'_M values differ by twofold or less except for one (C_{+3}-U_{42}) which is increased threefold upon cross-linking. Further, all the K'_M values are within a threefold range of completely unsubstituted three-stranded ribozyme (Table 1). Finally, all six of the K'_M values measured for the DTT-treated ribozymes (Table 2B) and the 2′-amino substituted ribozymes (Table 1) are within a twofold range. A comparison of the k'_{cat} values of DTT-treated hairpins (Table 2B) with that of the 2′-amino substituted ribozymes (Table 1)

shows only small reductions of two- to eightfold except 20-fold for C_{+3}-U_{39}.

To assess the overall effects of inter-domain constraint on hairpin ribozyme cleavage, we have calculated the ratio of the catalytic rate (k'_{cat} value) of each of the cross-linked ribozymes (aryl and alkyl) to that of their unconstrained, DTT-treated counterparts (Table 3). High values of this ratio reflect marginal effects of cross-linking on cleavage rate whereas low values reflect strong effects. The results fall into three distinct categories. For two of these combinations (U_{+2}-U_{39}, U_{+2}-U_{42}), there are only marginal effects (two- to fourfold). In two cases (U_{+2}-C_{44} and C_{+3}-C_{44}) there are strong effects of cross-linking leading to drastic reductions in the ratio (up to 300-fold). For C_{+3}-U_{42} there is an intermediate value (moderate effect of five to 12-fold). In all these cases the aryl and alkyl data are in broad agreement, which adds to the confidence that the presence or absence of catalytic rate reductions is related to differential effects of cross-linking restraint depending on the locations of the respective 2′-positions. For C_{+3}-U_{39} there is some uncertainty as to whether this cross-link should be in the moderate or strong category. Since the shorter and more flexible alkyl linker led to the smaller effect on the ratio, it would be reasonable to place this cross-link in the moderate category.

Docking of loop A with loop B

The rank order of effects of the cross-links (marginal, moderate and strong) could now be used as a guide in the docking of the two domains. Specifically, as different docking alternatives were sampled, the complete set of 2′-2′ distances were measured for the residues that had been used in cross-linking and the process of sampling continued until a "best fit" had been obtained. However, there were also a number of other criteria which needed to be considered during the docking process.

The docking of loop A with loop B imposes a sharp turn in the sugar-phosphate backbone at the level of the hinge region between residues 14 and 15 (Komatsu et al., 1994). *In vitro* selection studies (Joseph et al., 1993) show a random distribution of residues 14 and " −5", and probing experiments (Butcher & Burke, 1994b) bring evidence about the stabilization of base-pair $A_{15} \cdot U_{49}$ and the destabilization of $A_{14} \cdot U_{-5}$ upon substrate binding. Therefore, the opening of the latter base-pair was assumed and we used the torsional angles in the sugar-phosphate backbone between nucleotides 13 and 15 to explore the allowed conformational
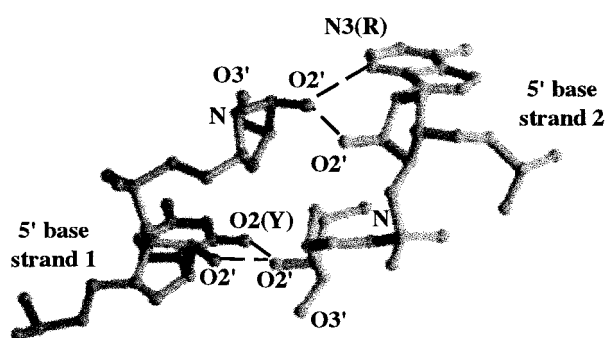
Figure 8. A schematic drawing of the ribose zipper as present in the modelled structure. The structure is very close to that seen in the crystal structure of the P4-P6 domain (Cate *et al.*, 1996). A ribose zipper is a structural motif which occurs between two consecutive sugar moieties of two antiparallel strands belonging to separate and adjacent helixes. The hydroxyl O2′ of the 3′-end sugar of one strand forms H-bonds with the hydroxyl O2′ and a base atom of the 5′-end residue of the other strand. The base atom involved is either O2 (in the case of a pyrimidine residue) or N3 (in the case of a purine residue). The 3′-end bases are not shown, for clarity of the drawing and to emphasize the apparent lack of sequence effects.

space of strand A towards the loop B residues important for catalysis (Berzal-Herranz *et al.*, 1993; Chowrira *et al.*, 1993b).

When attempting the determination of the relative angle between the helical axis of stems A and B, we took into account the data of Komatsu *et al.* (1996) in addition to the cross-linking data of the present work. These authors describe how they successively branched from helix H1 (H1-construct) or helix H4 (H4-construct) a four base-pair stem-loop, respectively preceded or followed by a stretch of six unpaired adenines. This resulted in the inactivation of the H1-construct while the H4-construct remained active. The final model building results in a tilt of $-80°$ of the helical axis of stem A towards the helical axis of stem B. With the H4-construct, the additional hairpin can stack naturally below domain B, while this would not be favourable for the H1-construct, which would imply a different angle between the two domains.
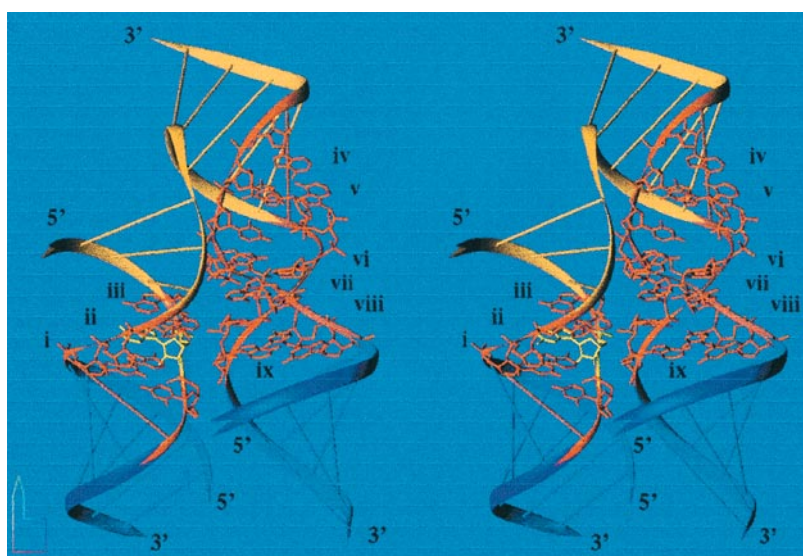
In addition, we took advantage of the lack of conformational constraints on the bulging residues U_{+2} and U_{39} to form a ribose-zipper (Cate *et al.*, 1996) between the residues for which the O2′ atoms were found essential for catalysis, i.e. 10, 11, 24, and 25 (Chowrira *et al.*, 1993b; Figure 8). In the model, the ribose hydroxyl groups for which O2′ removal is rescued by raising the magnesium concentration (11 and 24; Chowrira *et al.*, 1993b) interact with each other. This observation gives additional clues for the geometry of the $A_{24} \cdot A_{38}$ sheared pair. Indeed the latter involves the N6 atom of A_{24} with both the N3 and O2′ atoms of A_{38}. It has been shown that N6 or N7 modification of A_{24} or A_{38} results in a deleterious effect upon

catalysis, although to different levels (Grasby *et al.*, 1995). At face value, a change in the interacting atoms (N6 of A_{38} with N3 and O2′ of A_{24}) could therefore explicitly account for these results. But, at the same time, it would force A_{24} to slip towards the solvent side of domain B, making unlikely any direct relationship between A_{24} and G_{11}. In the geometry chosen for the model, A_{38}(N6) is involved in a direct hydrogen-bond with the *pro-S*$_P$ oxygen of the phosphate at A_{24}. Thus only definite loop structures allow for the formation of a basically sequence-independent ribose zipper between residues 10, 11, 24 and 25.

Finally to complete docking, U_{+2} was flipped into the shallow groove of loop A to allow interaction with the N2 atom of G_8, while the N3 atom of U_{+2} was exposed to the solvent, in agreement with probing data (Butcher & Burke, 1994b). This additional modification makes the ribose-phosphate backbone of three consecutive base-pairs $A_{-1} \cdot A_{10}$, $G_{+1} \cdot A_9$, and $U_{+2} \cdot G_8$ adopt a geometry very similar to that present in the hammerhead ribozyme in the region of the sheared tandem, $U_7 \cdot A_{14}$, $G_8 \cdot A_{13}$, and $A_9 \cdot G_{12}$.

The final docked model (Figure 9) shows these geometrical features described above and the inter-domain feature of the ribose zipper. Theoretical accessibilities of the Watson–Crick positions of residues in the vicinity of the ribose-zipper (A_9, A_{10} and A_{24}) were subject to slight modification upon docking (Figure 5). The cross-linking data proved to be able to be fitted reasonably well with the rank order of their distances in the modelling process. Thus at the conclusion, the cross-links with a marginal effect correspond to distances of 21 Å (U_{+2}-U_{39}), and 29 Å (U_{+2}-U_{42}). In the case of cross-links with a moderate effect, distances are 25 Å (C_{+3}-U_{39}) and 34 Å (C_{+3}-U_{42}), while for the cross-links with a strong effect the distances are 37 Å (U_{+2}-C_{44}) and 43 Å (C_{+3}-C_{44}) (Figure 10).

The interpretation of the distance data becomes clearer when some additional criteria are considered. Of the cross-links with a marginal effect, U_{+2}-U_{39} involves two bulged residues, one in each domain, the conformations of which are thus not really constrained. For example, rotation of these residues by about 180° along their respective P-O3′ axis shortens the distance that separates them by at least 8 Å. Therefore it would be easy to obtain an O2′-O2′ distance of less than 12 Å, i.e. within the expected span of the aryl and alkyl linkers. The weak effect of this cross-link can be interpreted as an ability of these bulged residues to adopt conformations that disturb neither the docking of loop A with loop B nor the catalytic transition state of the ribozyme. Of the two other cross-links involving bulged residues, U_{+2}-U_{42} and C_{+3}-U_{39}, marginal and moderate effects are observed, although the distances spanned are almost the same. This can be rationalized on the model by analysing the path followed by the cross-linking reagents for each configuration. The sugars of U_{+2} and U_{42} are both pointing out on the same side of the model result-
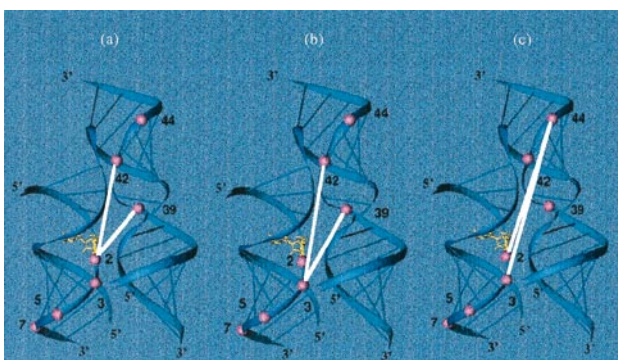
Figure 9. Side-by-side stereo-pair of the model of the hairpin ribozyme in the docked conformation. Helixes H2 and H3 are in orange, while helixes H1 and H4 are in dark blue. Base-pairs in these helixes are drawn schematically as straight bars. The nucleotides of the internal loops A (left) and B (right) are drawn in ball and stick mode in reddish colour. Domain A is docked onto residues within domain B in a geometry analogous to that between helix P1 on the loop J4/5 in group I introns (Michel & Westhof, 1990) and further held together by a ribose zipper (see Figure 8). Base-pairs indicated with roman numbering adopt unusual geometries (see Figures 3 and 4). Domain A: i, G_8 (N2)·U_{+2} (O4), ii, sheared A_9 (N6, N7)·G_{+1} (N3, N2), iii, sheared A_{10} (N6)·A_{-1} (N3), Domain B: iv, sheared G_{21} (N3, N2)·A_{43} (N6, N7), v, Hoogsteen *trans* A_{22} (N6, N7)·U_{42} (N3, O2), vi, sheared A_{23} (N3)·A_{40} (N6), vii, sheared A_{24} (N6)·A_{38} (N3), viii, *trans* C_{25} (N4)·U_{37} (O4), ix, Watson–Crick *cis* A_{26} (N1, N6)·G_{36} (N1, O6).

ing in the absence of distortion of the ribozyme either in loop A or in loop B. The O2′-O2′ distance here could also be shortened somewhat by further rotation of the unconstrained residue U_{+2}. By contrast, for C_{+3}-U_{39}, even if U_{39} were to flip, the cross-linking reagent must cross through the docked structure, a path which would produce the structural distortions experimentally observed, especially for the aryl cross-link which contains bulky aromatic rings.

Cross-linking from loop B to helix H1 as a test of the model

The cross-links used to help in molecular modelling were chosen deliberately between residues in internal loops that were not thought to be involved significantly in substrate-ribozyme binding. As a

Figure 10. The three categories of cross-links are represented on the 3D-model of the hairpin ribozyme: (a) cross-links with marginal effect; distances are 21 Å (U_{+2}-U_{39}) and 29 Å (U_{+2}-U_{42}), (b) cross-links with a moderate effect; distances are 25 Å (C_{+3}-U_{39}) and 34 Å (C_{+3}-U_{42}), (c) cross-links with a strong effect; distances are 37 Å (U_{+2}-C_{44}) and 43 Å (C_{+3}-C_{44}).

check of the model, we wished to see whether it is possible to extend the cross-links to residues expected to be further away in the model and obtain corresponding reductions in cleavage rate. The most obvious pyrimidine nucleoside sites for cross-linking which did not involve altering the strategy for determining the kinetic parameters were U_{+5} and U_{+7} located on the substrate strand of helix H1. Such cross-linking would involve residues involved in substrate-ribozyme binding. Nevertheless it was decided to observe what effects on catalytic cleavage such cross-linking might exert. Accordingly cross-linked hairpin ribozymes were constructed as before.

The kinetic parameters under single turnover conditions of ribozymes cross-linked to U_{+5} unfortunately proved to be uninterpretable (data not shown). In practically all cases (irrespective of type of cross-link) the K'_M values for the cross-linked ribozymes differed greatly from the DTT-treated ribozymes (variances of five- to tenfold). Thus, despite the preannealing step to ribozyme A strand, it seems that the pendant group at the 2′-position of U_{+5} is deleterious in some way to obtaining correct substrate-ribozyme annealing. By contrast, apart from U_{+7}-C_{44} aryl, the K'_M values for cross-links to U_{+7} were generally more consistent (Table 4). Surprisingly however, the results did not show the expected reductions in ratio of catalytic rates. The U_{+7}-C_{44} alkyl cross-link showed a moderate loss of activity but there were only marginal effects for both U_{+7}-U_{39} and U_{+7}-U_{42}. These data can only be rationalized by assuming fraying of the end of helix H1. The base-paired state of U_{+7} is not a pre-requisite for catalysis, since it has been seen partially frayed in probing experiments (Butcher & Burke, 1994b). Further, the mismatch mutant U_{+7}-A has only a threefold effect on cleavage activity and a triple mismatch (U_{+7}-A;U_{+8}-

Table 4. Kinetic data for cleavage

Sub	RzB	DTT	K'_M [μM]	k'_{cat} [min^{-1}]	Ratio k'_{cat}
A. Aryl disulphide cross-linked and reduced hairpin ribozymes					
U_{+7}	U_{39}	−	0.028 ± 0.005	0.013 ± 0.001	
		+	0.033 ± 0.003	0.033 ± 0.001	0.39
	U_{42}	−	0.023 ± 0.003	0.0017 ± 0.0001	
		+	0.039 ± 0.004	0.0028 ± 0.0002	0.60
	C_{44}	−	0.014 ± 0.002	0.033 ± 0.002	
		+	0.13 ± 0.02	0.15 ± 0.01	N/D[a,b]
B. Alkyl disulphide cross-linked and reduced hairpin ribozymes					
U_{+7}	U_{39}	−	0.027 ± 0.004	0.0091 ± 0.0006	
		+	0.018 ± 0.004	0.0086 ± 0.0006	0.94
	U_{42}	−	N/D[a]		
		+	N/D[a]		
	C_{44}	−	0.026 ± 0.005	0.0011 ± 0.0001	
		+	0.044 ± 0.009	0.012 ± 0.001	0.092

[a] Not determined.
[b] K'_M differed by tenfold.

A;U$_{+9}$-A) has an identical k_{cat} to the unmodified substrate under multiple turnover conditions (Joseph *et al.*, 1993). The single turnover K'_M values for these mutants are, however, not reported. The ability of helix H1 to fray somewhat could provide sufficient leeway for the catalysis to occur. For cross-link U$_{+7}$-C$_{44}$, the stronger effect observed on cleavage implies in addition conformational strains in some important features of loop A. This distortion of loop A is most apparent for the cross-links having a strong effect on cleavage. Therefore, it is clearly impossible for loop A to adopt its native state in the cases where residues U$_{+2}$ or C$_{+3}$ are stretched to nucleotide C$_{44}$, although their associated distances are within 20% of the distance corresponding to cross-link U$_{+7}$-U$_{39}$.

Discussion

The main aim of this study was to obtain a preliminary model of the hairpin ribozyme. A second aim was to see to what extent inter-domain cross-linking techniques could be used to provide data helpful in distinguishing alternative ways of docking the two domains in the modelling process. We chose for this cross-linking study residues in the two loops A and B, and it was found possible to use the rank order of reductions in ribozyme activity to guide the docking process by comparisons to the O2′-O2′ distances obtained from U$_{+2}$ and C$_{+3}$ to each of U$_{39}$, U$_{42}$ and C$_{44}$ as the model was refined. There is, thus, good confidence that the docked model (Figure 9) has correctly oriented the two domains towards each other.

The molecular modelling procedure has involved data and assumptions concerning only the ground state structures of the two domains. This modelling was guided by both NMR data on loop A (Cai & Tinoco, 1996) and chemical accessibility data (Butcher & Burke, 1994b). The docking of domain A with domain B, which was guided by the cross-linking data, was carried out as far as possible without significant disturbance of these starting structures. Indeed only two residues, A$_{10}$ and U$_{39}$, show a decrease in accessibility following docking (Figure 5). The flexibility of the doubly bulged residues involved in U$_{+2}$-U$_{39}$ cross-linking allows for the approach of these two residues within the theoretical maximum distances spanned by the aryl disulphide (16 Å) and the alkyl disulphide (12 Å) linkages. Yet, it is clear that for some of the residues where there are only marginal effects of cross-linking, O2′-O2′ distances are longer (21 to 29 Å). This brings two remarks. First, the cleavage rate measurements reflect each time data for a single cross-link, while the calculated distances correspond to a model for which it was assumed that all cross-links are present and compatible. This simultaneous incorporation of the complete set of cross-links in a single model forces a consensus conformation which artefactually stretches the expected distances spanned by these cross-links. Indeed slight deformations could absorb the strains induced by each cross-link. Secondly, a more intimate approach of the two domains in the catalytically active configuration can be expected. Such an intimate contact would be reached by numerous small adjustments in the fit between the two domains, in addition to kinking of one or other helix and opening of one or both looped regions without bringing significant changes in the relative rank of the inter-residue distances as defined by the cross-links.

Clues to the catalytically active structure of the hairpin ribozyme have come from studies of the functional group requirements for cleavage. This method involves substitution of individual nucleotide residues in the hairpin with nucleotide analogues carrying specific base or sugar modifications. Regarding base substitution data, an early finding was that when G$_{+1}$ in the substrate strand was substituted by inosine all detectable cleavage activity was lost (Chowrira & Burke, 1991), suggesting that the 2-amino group of G$_{+1}$ may be involved in the cleavage mechanism. More recently, the functional group requirements of the

essential purine residues in loops A and B of a 3-stranded ribozyme were studied (Grasby *et al.*, 1995). Loop A substitution results were generally consistent with cross-strand interactions proposed in the NMR structure (Cai & Tinoco, 1996). By contrast in loop B, it was found that the exocyclic amino groups of A_{40} and A_{43} are not required for catalytic cleavage (Grasby *et al.*, 1995), whereas they are clearly necessary for formation of cross-strand pairing in the initially folded structure of domain B (Butcher & Burke, 1994b). These functional group studies could favour a model with structural rearrangements in the loops as the two domains interact in the transition state. However, adenines A_{40} and A_{43} could be crucial for correct folding of loop B without being involved in the ensuing catalytic process itself, as previously observed for the mutation A_{14}-G in the hammerhead ribozyme (Bassi *et al.*, 1996).

It can be seen from our model that the base of G_{+1} (depicted in yellow in Figure 9) is in the *anti* configuration to allow pairing with A_9, as seen in the NMR structure of isolated domain A (Cai & Tinoco, 1996). The exocyclic amino of G_{+1} must play an extremely important participatory role in the cleavage of the adjacent phosphate (Chowrira & Burke, 1991) but its positioning in our model is too distant to give any clues as to how this amino group may be involved in cleavage. Further, we can find no evidence for the base to flip into a *syn* configuration as suggested previously (Chowrira & Burke, 1991). At this point it is worth noting that in the hammerhead ribozyme removal of the N2 amino groups of guanine G_{12}, involved in a sheared G·A pair (and also to some extent G_8) leads to severe reduction in catalytic activity despite the fact that they are far away from the cleaved phosphodiester (Fu & McLaughlin, 1992; Slim & Gait, 1992; Tuschl *et al.*, 1993).

Data concerning the cleavage requirements for sugar residues and 2′-hydroxyl groups proved very valuable in the modelling. Since U_{39} can be replaced by a propyl linker, this residue acts as a flexible bulge and plays no part in catalytic cleavage (Schmidt *et al.*, 1996a). This information was crucial to the modelling of loop B. Moreover, there are four hydroxyl groups on the ribozyme strand which have been found to be essential for cleavage, A_{10}, G_{11}, A_{24}, and C_{25} (Chowrira *et al.*, 1993b). The final docked configuration was reached after inclusion of a ribose zipper (Cate *et al.*, 1996) between the two domains of the hairpin ribozyme (Figure 8) in addition to the cross-linking data which alone would have left rotational ambiguity. The ribose zipper takes into account the essential hydroxyl groups (A_{10} to C_{25} and G_{11} to A_{24}). The differentiation between the interacting riboses is supported by experiment, since the 10-25 and 11-24 coupled riboses behave differently upon an increase in magnesium ion concentration. Whilst inhibition by 2′-modification at sugars 11 or 24 can be rescued by an increase in magnesium ion concentration, 2′-modification at 10 or 25 cannot.

Therefore, the 2′-hydroxyl groups of A_{24} and G_{11} have been proposed to be functional in magnesium binding in the transition state (Chowrira *et al.*, 1993b). However, such a configuration of hydroxyl groups could also help in the precise positioning of the two-domain ribozyme and especially of A_{10} which pairs with the cleavable residue A_{-1}.

The molecular model of the hairpin ribozyme we have obtained provides a realistic and invaluable starting point for further experimental design. Further checks of inter-domain distances might be possible by choosing residues which would be expected to be reasonably close in the model and thus tolerant of cross-linking, such as C_{-4} in helix H2 to U_{39} in loop B. This is beyond the scope of the current study, since hairpin ribozyme cleavage would release a different fragment, the kinetics of which would require separate validation. Further, we have also learnt that caution is needed when obtaining cross-linking data for residues involved in substrate-ribozyme binding (such as C_{-4}). It would also be interesting to carry out tests of some of the other features of the docked configuration, such as chemical substitutions of the hydroxyl groups involved in the ribose zipper motif, and how the precise conformations of loops A and B allow for the correct positioning of the ribose zipper. A range of complementary techniques can also be considered, such as fluorescence resonance energy transfer, *in vitro* selection and footprinting, which together with further cross-linking studies may provide additional data helpful in model refinement.

Materials and Methods

Preparation of oligoribonucleotides

Oligoribonucleotides were synthesized on solid-phase controlled pore glass (CPG) on a 1 μmol scale using 2′-O-*t*-butyldimethylsilylnucleoside-3′-*O*-(2-cyanoethyl-*N*,*N*-diisopropyl) phosphoramidite monomers having phenoxyacetyl amino protection for A, isopropylphenoxyacetyl for G and benzoyl protection for C (Glen Research *via* Cambio) as previously described (Gait *et al.*, 1991; Grasby *et al.*, 1995; Schmidt *et al.*, 1996a,b). 2′-Deoxy-2′-trifluoroacetylamino-3′-*O*-(2-cyanoethyl-*N*,*N*-di-isopropyl)phosphoramidite derivatives of uridine and cytidine were generous gifts from Dr Wolfgang Pieken (NeXstar Pharamaceuticals Inc.) and required no special conditions for coupling or deprotection (Pieken *et al.*, 1991). Oligoribonucleotides were deprotected by treatment of the CPG with methanolic ammonia followed by triethylamine trihydrofluoride/DMF (3:1) at 55°C for 1.5 h (Wincott *et al.*, 1995; Schmidt *et al.*, 1996a,b) and desalting *via* Sephadex NAP-10 (Pharmacia) filtration or by butanol precipitation (Gait *et al.*, 1991).

Purification of the oligoribonucleotides was carried out by anion exchange chromatography on a NucleoPac PA-100 column (Dionex) (Sproat *et al.*, 1995; Schmidt *et al.*, 1996b) using buffer A: 10 mM and buffer B: 400 mM sodium perchlorate, 20 mM Tris-HCl (pH 6.8), 25% formamide, flow rate 3 ml min^{-1} with gradients of 15 to 45% B over 20 minutes. Desalting was achieved *via* extensive dialysis against water. The purity of the chemi-

cally synthesized RNA was checked by 5′-^{32}P-labelling (Slim & Gait, 1991) followed by electrophoresis on a 20% denaturing polyacrylamide gel (PAGE).

Preparation of cross-linked substrate-ribozyme B complexes

The 2′-amino modified substrate strand (2.5 nmol) was 5′-^{32}P radiolabelled (Slim & Gait, 1991), mixed with 2′-amino modified ribozyme strand B (12.5 nmol) and evaporated to dryness. The resulting residue was dissolved in 2 μl of boric acid/borax buffer (0.07 M, pH 8.6) and 2 μl of the cross-linking reagent (3-isothiocyanatobenzyl 2-pyridyl disulphide, Linking agent A (Sigurdsson *et al.*, 1995a), or 2-isocyanatoethyl 2-pyridyl disulphide, Linking agent B (Sigurdsson *et al.*, 1996b)) (100 mM in DMF) was added. The reaction mixture was left overnight at 37°C and vortexed at intervals. The solution was diluted with water (15 μl) and 1 M Tris-HCl (pH 7.5, 1 μl). The pyridyl protecting groups were removed by addition of DTT (100 mM, 50 μl), sodium chloride solution (0.5 M, 10 μl), Tris-HCl (1 M, pH 7.5, 2 μl) and water (65 μl) for three hours at room temperature. The RNAs were precipitated at −20°C for two hours by addition of ammonium acetate solution (3 M, 150 μl), glycogen (2 μl, 20 mg ml^{-1}) and ethanol (1.2 ml). The RNA was collected by centrifugation and the supernatant removed by decantation. The pellet was washed with a further 200 μl of ethanol to ensure any excess DTT was removed. The RNAs were then air oxidized by treatment in a 2 ml microfuge tube with 300 μl of buffer C (50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM MgCl$_2$) overnight at 37°C. To purify the hetero-linked RNAs from the homo-linked RNAs, the reaction products were subjected to 20% polyacrylamide gel electrophoresis in the presence of 7 M urea (Gait *et al.*, 1991). The radioactive hetero-dimer was located by auto-radiography as the slowest migrating band and the RNA eluted by soaking with 3 M ammonium acetate solution (100 μl) plus water (200 μl) for six hours at room temperature followed by a second elution with 3 M ammonium acetate solution (50 μl) and water (100 μl). The combined eluates were washed with phenol:chloroform (1:1, 200 μl) and the cross-linked RNA was precipitated by addition of ethanol (1.2 ml) and the pellet briefly dried.

Determination of ribozyme kinetics

Each cross-linked RNA was redissolved in water (200 μl) and split into 2 aliquots. To one aliquot was added DTT (5 μl, 100 mM) to obtain a final concentration of 5 mM and the reaction was vortexed several times through a six-hour period. To the second aliquot was added an identical volume of water (5 μl). The cleavage rates of each the cross-linked and reduced forms of the ribozymes aliquots were measured under single turnover conditions (Fedor & Uhlenbeck, 1992). Each reaction mixture (90 μl) consisted of an aqueous solution of the cross-linked or DTT-treated substrate-RzB strand RNA (about 5 to 20 nM), 40 mM Tris-HCl (pH 7.5) and one of six concentrations of ribozyme strand A usually in the range 10 to 150 nM. In the cases of the DTT-treated RNAs, the reaction mixture also contained DTT (5 mM). Each solution was incubated at 70°C for one minute and then 37°C for 15 minutes. Each cleavage reaction was initiated by addition of 100 mM MgCl$_2$ solution (10 μl) to give a final magnesium concentration of 10 mM. Aliquots (10 μl) were removed at six suitable time intervals and the reactions quenched by addition to 10 μl of urea

stop mix (7 M urea, 50 mM EDTA, 0.04% (w/v) xylene cyanol, 0.04% (w/v) bromophenol blue). Samples were loaded on to a 20% denaturing polyacrylamide gel and subjected to electrophoresis at 12 W for 80 minutes. The resultant gels were dried and scanned using a phosphorImager (Molecular Dynamics). The data were processed using the programme Image Quant (Molecular Dynamics) and quantitated by use of the Geltrak programme (Smith & Thomas, 1990; Smith & Singh, 1996). The single-turnover rates of cleavage (k_{obs}) were plotted as a function of ribozyme concentration over k_{obs} using Eadie−Hofstee plots as previously described (Fedor & Uhlenbeck, 1992) such that k'_{cat} values were obtained from the Y-intercept and K'_M values from the slopes.

Molecular Modelling

The molecular modelling method is as previously described (Westhof, 1993). The independent atomic 2D-structure elements were built with the programs FRAGMENT and NAHELIX. The 3D-generated elements were then linked together, incorporating the biochemical data, interactively using the software programs FRODO (Jones, 1978; Pflugrath *et al.*, 1983), designed for nucleic acid modelling (STG version) (Amerein *et al.*, 1987), and MANIP (Massire & Westhof, in preparation). The resulting working models were then subjected to stereochemical and geometrical refinement in order to ensure proper geometry and to prevent bad contacts using the Konnert−Hendrickson (Konnert & Hendrickson, 1980) restrained least-squares refinement program NUCLIN/NUCLSQ (Westhof *et al.*, 1985). At the end of each modelling step, the theoretical accessible surfaces (in Å^2) of the Watson-Crick (N1[R], N3[Y]) and Hoogsteen (N7[R]) positions (Richmond, 1984) are compared to the reactivities of these atoms in solution (Butcher & Burke, 1994b). Surfaces below 10 Å^2 are usually considered spanning the range from unreactive to weakly reactive in native conditions (Brunel, 1992), since molecular dynamics are not explicitly taken into account in a rigid 3D model. Colour pictures were designed with the software DRAWNA (Massire *et al.*, 1994). The docking of the two domains was achieved by monitoring in real time the distances between all the cross-linked residues while torsional angle variations (between residues 14 and 15) and overall helix rotations were performed using the program MANIP (Massire & Westhof, in preparation).

References

Allerson, C. R. & Verdine, G. L. (1995). Synthesis and biochemical evaluation of RNA containing an intrahelical disulphide crosslink. *Chem. Biol.* **2**, 667–675.

Amerein, B., Ripp, R. & Dumas, P. (1987). PUCK : a real time modification of sugar pucker on a PS300. *J. Mol. Graph.* **5**, 184–189.

Bassi, G. S., Murchie, A. I. H. & Lilley, D. M. J. (1996). The ion-induced folding of the hammerhead ribozyme: core sequence changes that perturb folding into the active conformation. *RNA,* **2**, 756–768.

Berzal-Herranz, A., Joseph, S., Chowrira, B. M., Butcher, S. E. & Burke, J. M. (1993). Essential nucleotide sequences and secondary structure elements of the hairpin ribozyme. *EMBO J.* **12**, 2567–2574.

Brunel, C. (1992). La conformation en solution de l'ARN ribosomique 5 S. Mecanisme de régulation traductionnelle de la threonyl-ARNt synthetase d'*Escherichia coli*: un exemple de mimétisme moléculaire. U.F.R. des Sciences de la vie et de la terre. PhD thesis, p. 335, Univérsité Louis Pasteur, Strasbourg.

Burke, J. M. (1994). The hairpin ribozyme. In *Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D. M. J., eds), pp. 105–118, Springer-Verlag, Berlin.

Burke, J. M. (1996). Hairpin ribozyme: current status and future prospects. *Biochem. Soc. Trans.* **24**, 608–615.

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., Heckman, J. E. & Burke, J. M. (1995). Reconstitution of hairpin ribozyme activity following separation of functional domains. *J. Biol. Chem.* **270**, 29648–29652.

Cai, Z. & Tinoco, I. (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. (1996). Crystal structure of a group 1 ribozyme domain: principles of RNA packing. *Science,* **273**, 1678–1684.

Chowrira, B. M. & Burke, J. M. (1991). Novel guanosine requirement for catalysis by the hairpin ribozyme. *Nature,* **354**, 320–322.

Chowrira, B. M., Berzal-Herranz, A. & Burke, J. M. (1993a). Ionic requirements for RNA binding, cleavage, and ligation by the hairpin ribozyme. *Biochemistry,* **32**, 1088–1095.

Chowrira, B. M., Berzal-Herranz, A., Keller, C. F. & Burke, J. M. (1993b). Four ribose 2′-hydroxyl groups essential for catalytic function of the hairpin ribozyme. *J. Biol. Chem.* **268**, 19458–19462.

Dahm, S. A. & Unlenbeck, O. C. (1991). Role of divalent metal ions in the hammerhead RNA cleavage reaction. *Biochemistry,* **30**, 9464–9469.

Dahm, S. A., Derrick, W. B. & Uhlenbeck, O. C. (1993). Evidence for the role of solvated metal hydroxide in the hammerhead cleavage reaction. *Biochemistry,* **32**, 13040–13045.

Favre, A. & Fourrey, A. (1995). Structural probing of small endonucleolytic ribozymes in solution using thio-substituted nucleobases as intrinsic photolabels. *Acc. Chem. Res.* **28**, 375–382.

Fedor, M. J. & Uhlenbeck, O. C. (1992). Kinetics of intermolecular cleavage by hammerhead ribozymes. *Biochemistry,* **31**, 12042–12054.

Feldstein, P. A. & Bruening, G. (1993). Catalytically active geometry in the reversible circularization of mini-monomer RNAs derived from the complementary strand of tobacco ringspot virus satellite RNA. *Nucl. Acids Res.* **21**, 1991–1998.

Fu, D.-J. & McLaughlin, L. W. (1992). Importance of specific purine amino and hydroxyl groups for efficient cleavage by a hammerhead ribozyme. *Proc. Natl Acad. Sci. USA,* **89**, 3985–3989.

Gait, M. J., Pritchard, C. E. & Slim, G. (1991). Oligoribonucleotide synthesis. In *Oligonucleotides and Analogues: A Practical Approach* (Eckstein, F., ed.), pp. 25–48, Oxford University Press, Oxford, UK.

Gautheret, D., Konings, D. & Gutell, R. R. (1994). A major family of motifs involving G·A mismatches in ribosomal RNA. *J. Mol. Biol.* **242**, 1–8.

Goodwin, J. T. & Glick, G. D. (1994). Synthesis of a disulphide stabilized RNA hairpin. *Tetrahedron Letters,* **35**, 1647–1650.

Goodwin, J. T., Osborne, S. E., Scholle, E. J. & Glick, G. D. (1996). Design, synthesis, and analysis of yeast tRNAPhe analogs possessing intra- and inter-helical disulphide cross-links. *J. Am. Chem. Soc.* **118**, 5207–5215.

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.

Hampel, A., Tritz, R., Hicks, M. & Cruz, P. (1990). 'Hairpin' catalytic RNA model: evidence for helices and sequence requirement for substrate RNA. *Nucl. Acids Res.* **18**, 299–304.

Jones, T. A. (1978). A graphics model building and refinment system for macromolecules. *J. Appl. Crystallog.* **11**, 268–272.

Joseph, S., Berzal-Herranz, A., Chowrira, B. M., Butcher, S. E. & Burke, J. M. (1993). Substrate selection rules for the hairpin ribozyme determined by *in vitro* selection, mutation, and analysis of mismatched substrates. *Genes Dev.* **7**, 130–138.

Koizumi, M. & Ohtsuka, E. (1991). Effects of phosphorothioate and 2–amino groups in hammerhead ribozymes on cleavage rates and Mg²⁺ binding. *Biochemistry,* **30**, 5145–5150.

Komatsu, Y., Koizumi, M., Nakamura, H. & Ohtsuka, E. (1994). Loop-size variation to probe a bent structure of a hairpin ribozyme. *J. Am. Chem. Soc.* **116**, 3692–3696.

Komatsu, Y., Kanzaki, I., Koizumi, M. & Ohtsuka, E. (1995). Modification of primary structures 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.

Konnert, J. H. & Hendrickson, W. A. (1980). Restrained parameters thermal factors refinement procedures. *Acta Crystallog. sect. A,* **36**, 344–349.

Long, D. M. & Uhlenbeck, O. C. (1993). Self-cleaving catalytic RNA. *FASEB J.* **7**, 25–30.

Massire, C., Gaspin, C. & Westhof, E. (1994). DRAWNA: a program for drawing schematic views of nucleic acids. *J. Mol. Graph.* **12**, 201–206.

Michel, F. & Westhof, E. (1990). Modelling of the three-dimensional architecture of group 1 catalytic introns

based on comparative sequence analysis. *J. Mol. Biol.* **3**, 585–610.

Pflugrath, J. W., Saper, M. A. & Quiocho, F. A. (1983). Molecular Modelling with the PS300: a new generation graphics display system. *J. Mol. Graph.* **1**, 53–54.

Pieken, W. A., Olsen, D. B., Benseler, F., Aurup, H. & Eckstein, F. (1991). Kinetic characterization of ribonuclease-resistant 2'-modified hammerhead ribozymes. *Science,* **253**, 314–317.

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

Pyle, A. M. (1993). Ribozymes: a distinct class of metalloenzymes. *Science,* **261**, 709–714.

Richmond, T. J. (1984). Solvent accessible surface area and excluded volume in proteins. *J. Mol. Biol.* **178**, 63–89.

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

Schmidt, S., Grenfell, R. L., Fogg, J., Smith, T. V., Grasby, J. A., Mersmann, K. & Gait, M. J. (1996b). Solid phase synthesis of oligoribonucleotides containing site-specific modifications. In *Innovation and Perspectives in Solid Phase Synthesis and Combinatorial Libraries, 4th International Symposium Proceedings 1996* (Epton, R., ed.), pp. 11–18, Mayflower, Birmingham, UK.

Schmidt, S., Earnshaw, D. J., Sigurdsson, S. T., Eckstein, F. & Gait, M. J. (1997). Probing the tertiary structure of the hairpin ribozyme by chemical cross-linking. *Coll. Czech. Chem. Commun.* **61** (special issue), 276–279.

Scott, W. G., Finch, J. T. & Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism of action 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.

Shin, C., Choi, J. N., Song, J. T., Ahn, J. H., Lee, J. S. & Choi, Y. D. (1996). The loop B domain is physically separable from the loop A domain in the hairpin ribozyme. *Nucl. Acids Res.* **24**, 2685–2689.

Sigurdsson, S. T. & Eckstein, F. (1996a). Isolation of oligoribonucleotides containing intramolecular cross-links. *Anal. Biochem.* **235**, 241–242.

Sigurdsson, S. T. & Eckstein, F. (1996b). Site specific labelling of sugar residues in oligoribonucleotides: reactions of aliphatic isocyanates with 2'-amino groups **24**, 3129–3133.

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

Sigurdsson, S., Tuschl, T. & Eckstein, F. (1995b). Probing the tertiary structure of the hammerhead ribozyme using spectroscopy and interhelical cross-linking. In *Proceedings from the Ninth Converstaion in the Discipline Biomolecular Stereodynamics* (Sarma, R. H. & Sarma, M. H., eds), vol. 2, pp. 57–60, Adenine Press, New York.

Siwkowski, A., Shippy, R. & Hampel, A. (1997). Analysis of hairpin ribozyme base mutations in loops 2 and 4 and their effects on *cis*-cleavage *in vitro*. *Biochemistry,* **36**, 3930–3940.

Slim, G. & Gait, M. J. (1991). Configurationally defined phosphorothioate-containing oligoribonucleotides in the study of the mechanism of cleavage of hammerhead ribozymes. *Nucl. Acids Res.* **19**, 1183–1188.

Slim, G. & Gait, M. J. (1992). The role of the exocyclic amino groups of conserved purines in hammerhead ribozyme cleavage. *Biochem. Biophys. Res. Commun.* **183**, 605–609.

Smith, J. & Singh, M. (1996). System for accurate one-dimensional gel analysis including high resolution quantitative footprinting. *BioTechniques,* **20**, 1082–1087.

Smith, J. M. & Thomas, D. J. (1990). Quantitative analysis of one-dimensional gel electrophoresis profiles. *Compl. Appl. Biosci.* **6**, 93–99.

Sproat, B., Colonna, F., Mullah, B., Tsou, D., Andrus, A., Hampel, A. & Vinayak, R. (1995). An efficient method for the isolation and purification of oligoribonucleotides. *Nucleosides and Nucleotides,* **14**, 255–273.

Symons, R. H. (1992). Small catalytic RNAs. *Annu. Rev. Biochem.* **61**, 641–671.

Szewczak, A. A., Moore, P. B., Chan, Y. L. & Wool, I. G. (1993). The conformation of the sarcin/ricin loop from 28 S ribosomal RNA. *Proc. Natl Acad. Sci. USA,* **90**, 9581–9585.

Tuschl, T., Ng, M. M. P., Pieken, W., Benseler, F. & Eckstein, F. (1993). Importance of exocyclic base functional groups of central core guanosines for hammerhead ribozyme activity. *Biochemistry,* **32**, 11658–11668.

Santos Vitorino, Dos D., Fourrey, J.-L. & Favre, A. (1993). Flexibility of the bulge formed between a hairpin ribozyme and deoxy-substrate analogues. *Biochem. Biophys. Res. Commun.* **190**, 377–385.

Westhof, E. (1993). Modelling the three-dimensional structure of ribonucleic acids. *J. Mol. Struct.* **286**, 203–210.

Westhof, E., Dumas, P. & Moras, D. (1985). Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol.* **184**, 119–145.

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

Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S. & Usman, N. (1995). Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucl. Acids Res.* **23**, 2677–2684.

Young, K. J., Gill, F. & Grasby, J. A. (1997). Metal ions play a passive role in the hairpin ribozyme catalysed reaction. *Nucl. Acids Res.* **25**, 3760–3766.

Edited by A. R. Fersht