Architecture of a Diels-Alderase Ribozyme with a Preformed Catalytic Pocket

Sonja Keiper,[1,4,5] Dirk Bebenroth,[1,4]
Burckhard Seelig,[2,6] Eric Westhof,[3]
and Andres Jäschke[1,*]
[1]Universität Heidelberg
Institut für Pharmazie
   und Molekulare Biotechnologie
D-69120 Heidelberg
Germany
[2]Freie Universität Berlin
Institut für Chemie
14195 Berlin
Germany
[3]UPR9002-CNRS
Université Louis Pasteur
67000 Strasbourg
France

Summary

Artificial ribozymes catalyze a variety of chemical reactions. Their structures and reaction mechanisms are largely unknown. We have analyzed a ribozyme catalyzing Diels-Alder cycloaddition reactions by comprehensive mutation analysis and a variety of probing techniques. New tertiary interactions involving base pairs between nucleotides of the 5′ terminus and a large internal loop forming a pseudoknot fold were identified. The probing data indicate a preformed tertiary structure that shows no major changes on substrate or product binding. Based on these observations, a molecular architecture featuring a Y-shaped arrangement is proposed. The tertiary structure is formed in a rather unusual way; that is, the opposite sides of the asymmetric internal loop are clamped by the four 5′-terminal nucleotides, forming two adjacent two base-pair helices. It is proposed that the catalytic pocket is formed by a wedge within one of these helices.

Introduction

One of the most fascinating features of RNA is its capacity to adopt defined three-dimensional foldings forming specific molecular pockets for binding other molecules or catalyzing chemical reactions. NMR spectroscopy and X-ray crystallography on RNA aptamer-target complexes have yielded valuable information about the molecular mechanisms of RNA binding [1–4]. A wide variety of biochemical and biophysical methods has been es-

*Correspondence: jaeschke@uni-hd.de
[4]These authors contributed equally to this work.
[5]Present address: School of Chemistry, Queen's University, Belfast BT9 5AG, United Kingdom.
[6]Present address: Department of Molecular Biology and Howard Hughes Medical Institute, Massachusetts General Hospital, Boston, Massachusetts 02114.

tablished to gain knowledge about the catalytic mechanisms of ribozymes [5–10].

While over the past two decades the catalytic mechanisms of the naturally occurring ribozymes have been studied intensively [11, 12], very little is known about how RNA catalyzes reactions other than phosphodiester hydrolysis and transesterification. To obtain a comprehensive picture of the catalytic abilities and limitations of ribozymes, it is thus important to expand mechanistic investigations to artificial ribozymes.

Our laboratory has reported the isolation of ribozymes from a combinatorial RNA library that catalyze the formation of carbon-carbon bonds by Diels-Alder reaction, a [4+2] cycloaddition reaction [13]. The isolated RNA molecules accelerated the reaction of an aromatic diene (anthracene) tethered to their 5′ ends with a biotinylated dienophile (maleimide) up to 20,000-fold. The majority of the selected sequences (13 independently evolved sequence families representing 32 individual sequences) contained a small secondary structure motif consisting of 3 helices, an asymmetric internal loop, and the formally single-stranded 5′ end (Figure 1A). A rationally designed 49-mer RNA (for the nucleotide sequence, see below) containing these elements was found to accelerate the reaction between a covalently tethered anthracene and biotin maleimide about as fast as the parental sequences (in *cis* reaction) [13]. These ribozymes were later found to accelerate the cycloaddition reaction between dienes and dienophiles free in solution (in *trans* reaction) in a truly bimolecular fashion. Catalysis proceeds with multiple turnover and high enantioselectivity, making this a unique ribozyme system [14].

In a previous study, the interactions of this 49-mer Diels-Alderase ribozyme with its substrates and products had been elucidated by chemical substitution analysis by using 44 different, systematically varied substrate and product analogs [15]. RNA substrate and RNA product interactions were found to be governed primarily by hydrophobic or *van der Waals* interactions, while hydrogen bonding and metal ion coordination appeared to be less important. These data provided a first insight into principles of substrate recognition and the molecular determinants of stereodifferentiation.
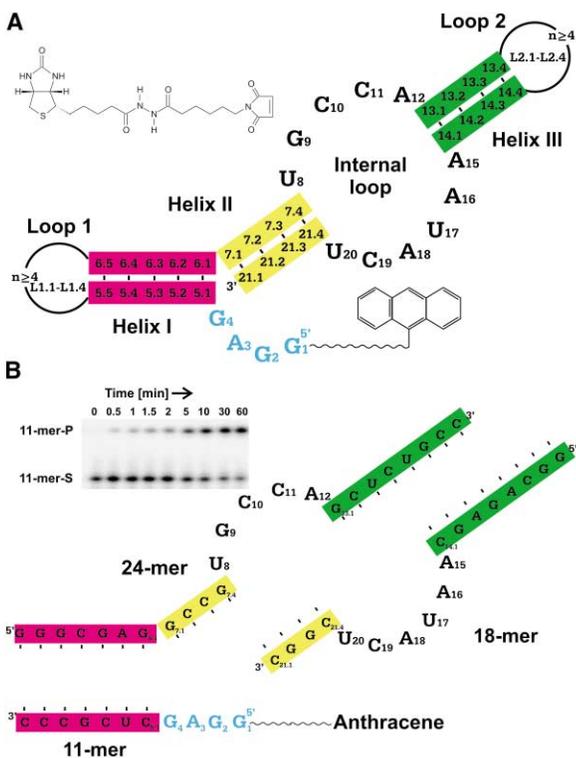
To gain an understanding of the structural features of the ribozyme itself, we have now systematically probed the existence of the individual structural elements and investigated the role of individual nucleotides in catalysis by mutation analysis. These experiments not only allow us to define conserved (or essential) residues, but they also allow us to search for tertiary interactions by compensatory double-mutation analysis. Enzymatic and chemical probing techniques were employed to explore the secondary and tertiary fold of the 49-mer RNA catalyst. Pb^{2+}-probing experiments were conducted to give information about general steric accessibility and about the localization of metal ions. To detect structural changes associated with substrate binding, probing experiments were either carried out with the 49-mer ribozyme alone, or with anthracene covalently attached to the 49-mer

Figure 1. Minimal Motif of Ribozymes with Diels-Alderase Activity

(A) Common secondary motif of Diels-Alderase ribozymes previously identified by in vitro selection, numbering scheme, and formula of biotin maleimide used.

(B) Tripartite ribozyme system consisting of a 24-mer, an 18-mer, and an 11-mer anthracene RNA strand (wild-type sequence). Inset: 20% PAGE gel assay of this tripartite system. Reactions were started by the addition of biotin maleimide, and aliquots withdrawn at the indicated time points were quenched with β-mercaptoethanol. For details, see the Experimental Procedures.

RNA. Alternatively, experiments were conducted in the presence of free substrates or reaction products. The results are compared with the consensus sequences derived from the original selection.

Results and Discussion

Probing the Secondary Structure: The Role of the Three Helices

Important features of the previously proposed secondary structure are (see Figure 1A): (1) three double-helical stems, (2) an asymmetric internal loop composed of a pentanucleotide, UGCCA, and a hexanucleotide, AAUACU, and (3) the formally single-stranded 5′-terminal GGAG tetranucleotide, with the first G being the attachment site for the anthracene.

To allow for rapid screening of single and double mutants, a tripartite ribozyme system was developed (Figure 1B). This ribozyme system of a 24-mer, an 18-mer, and an anthracene-tethered 11-mer spontaneously reassembles after mixing and has about 70% of the activity of the one-stranded in *cis* system. Reaction rates could be easily determined by a gel electrophoretic assay (see the inset in Figure 1B).

To probe the existence and importance of all three

helices, they were systematically destroyed and reassembled (Figure 2). When the respective strands in the helix I region were scrambled (i.e., each helix position replaced by its Watson-Crick counterpart), the combinations 11scr:24wt:18wt and 11wt:24scr:18wt gave no activity at all. The compensatory combination 11scr:24scr:18wt, however, restored activity to 70% wt. To test whether helix I is dispensable at all, it was replaced by a single-stranded tether of eight uridines that connects the 5′ GGAG to position 7.1 in helix II in a unimolecular assay format. No catalytic activity could be detected.

For helix II, the phylogeny of the selected sequences indicated a rather strong conservation, both in length and in the purine/pyrimidine pattern [13]. To investigate this relationship, helix II was destroyed by simultaneously replacing $G_{21.2}$ and $G_{21.3}$ by adenosines, thereby creating two A-C mismatches. The activity was reduced to zero. If at the same time $C_{7.2}$ and $C_{7.3}$ were replaced by uridines, activity was restored to about 80% of the wild-type level. The combination $G_{21.2},G_{21.3}:U_{7.2},U_{7.3}$ involving two G:U wobbles gave about 10% activity. These findings strongly support the assumption that helix II is an essential structural element (Figure 2).

Base pair 7.4:21.4, which constitutes the junction of helix II and the internal loop, was a G:C base pair in 11 of the 13 selected sequence families. Replacement by an A:U base pair gave about 70% of the wild-type activity, replacement by a G:U wobble gave less than 10% of the wild-type activity, and C:G, U:A, and U:G inversion mutants were almost completely inactive (<1% wild-type level), indicating that, at this position, standard base pairing with the proper purine-pyrimidine orientation is required for activity.

To test if the conserved length of 4 bp is an essential feature, insertion mutants 7.0A and 21.0U and deletion mutants 7.1del and 21.1del were prepared and were combined either with the respective wild-type strand or with each other. Insertion 21.0U, creating a 1 nt 3′ overhang, yielded 95% activity, which could be expected since all selected sequences contained extensive 3′ overhangs. Surprisingly, insertion 7.0A, which generates a 1 nt hinge between helices I and II showed 70% of the wild-type activity. Even the double mutant 7.0A:21.0U with a 5 bp helix II still had 55% activity. Deletion 7.1del was not tolerated (no activity), whereas deletion 21.1, i.e., shortening of the 3′ end by one nucleotide, was tolerated well (>90% wt).

The length and sequence of helix III were almost random in the selected sequences [13]. The only remarkable feature was its interface to the internal loop, where several occurrences of mismatches were found. To analyze the role of this interface, all 16 possible nucleotide combinations of the closing base pair 13.1:14.1 were now prepared. All combinations were within ±20% of the wild-type activity, except for a U:U mutant which showed a more dramatic reduction in activity (60% of the wild-type activity). It appears that these nucleotides are either not strongly paired, or that pairing does not influence the formation of a catalytically active structure.

Probing the Secondary Structure: Enzymatic Probing

Enzymatic probing was employed to further study the secondary structure of the Diels-Alderase ribozyme.
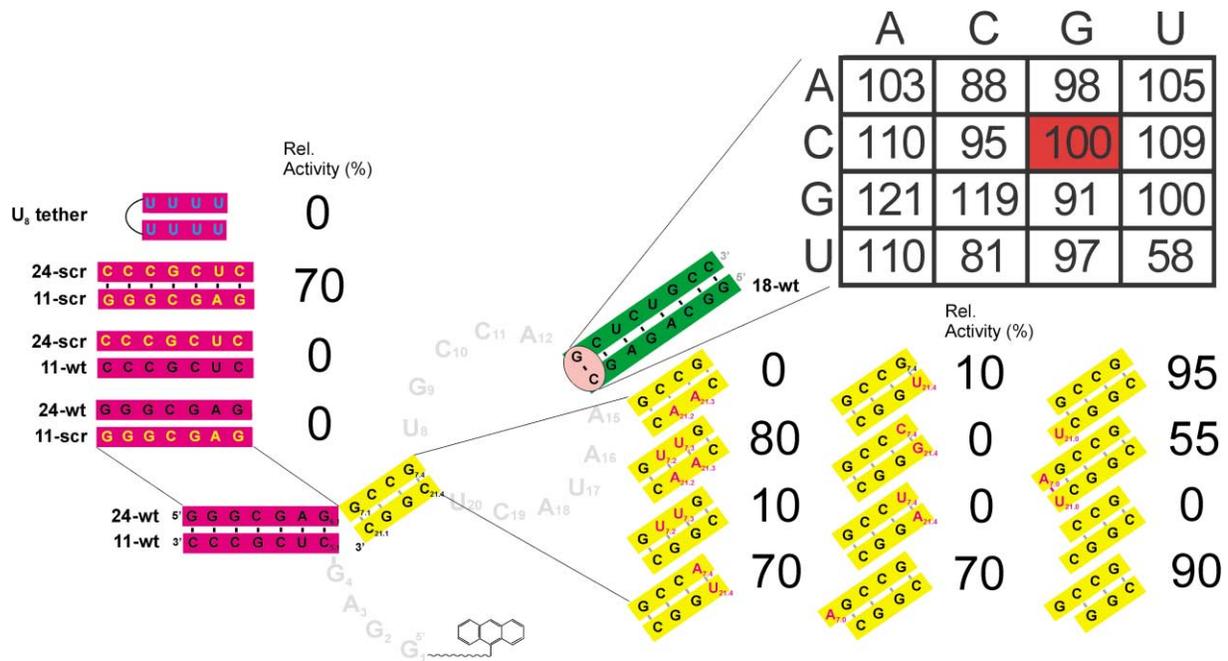
Figure 2. Probing the Secondary Structure by Mutation Analysis

Helices (red, yellow) or an individual base pair (pink) were mutated, and the effect on the ribozyme activity was investigated. The three strands were mixed, and reactions were started by the addition of biotin maleimide. Relative activity values represent single time point measurements (percent conversion after 30 min) relative to the tripartite wild-type ribozyme.

RNases S1, T2, T1, and U2 were used to induce cleavage of unpaired regions, while RNase V1 yielded information about helical double-stranded (and also stacked single-stranded) structures (Figure 3).

Digestion with nuclease S1 revealed cleavage at the two stable tetraloops at positions $U_{L1.1}$, $U_{L1.2}$, $C_{L1.3}$ and $U_{L2.1}$, $U_{L2.2}$, $C_{L2.3}$. Cleavage of the asymmetric internal loop was more distinct at its (lower) hexanucleotide stretch
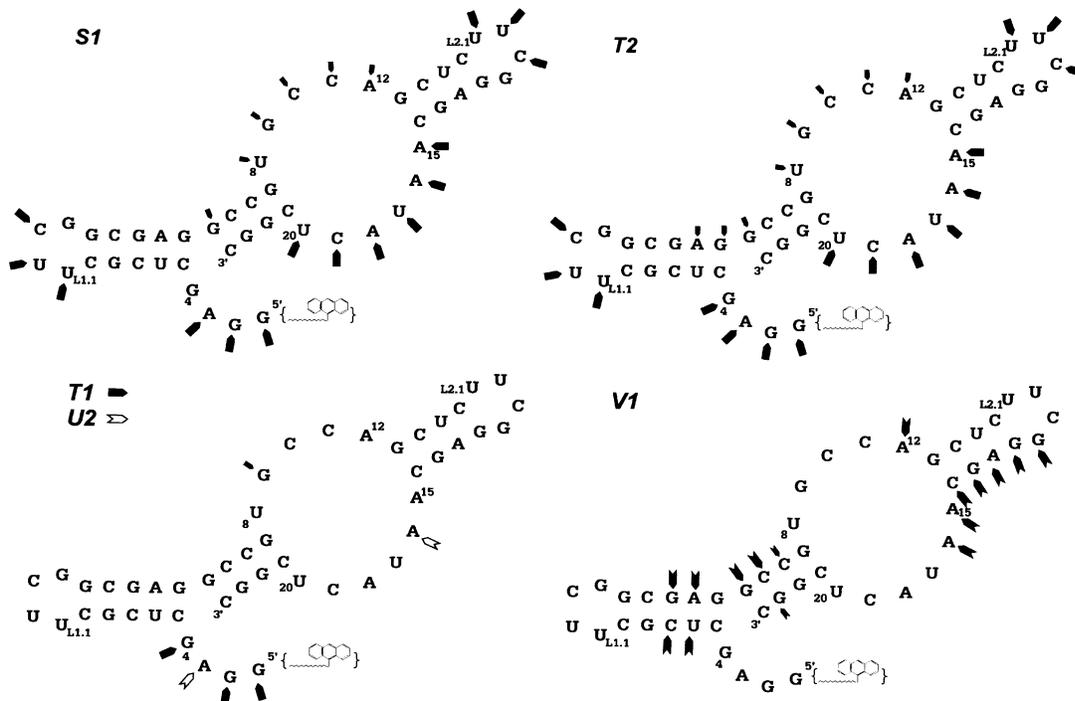

Figure 3. Enzymatic Secondary Structure Probing

The size of the arrows corresponds to the cleavage intensity at the respective position.

AAUACU, whereas the opposite pentanucleotide UGC CA was only weakly hydrolyzed by S1. At the 5′ terminus, the first three nucleotides, G_1, G_2, and A_3, were found to be susceptible to S1 digestion. The pattern of RNase T2-induced cleavage was found to be very similar.

The G-specific RNase T1 cleaved G_1, G_2, and G_4 in the 5′-terminal tetranucleotide and cleaved G_9, the only guanosine residue in the internal loop with minor intensity under native conditions. All other guanosine residues remain uncleaved, consistent with their involvement in helices or stable hairpin loops. Of the seven adenosines, only A_{16} and A_3 were cleaved by the A-specific RNase U2.

In the presence of RNase V1, the three proposed helices were all partly reactive. Cleavages were observed at positions $U_{5.2}$, $C_{5.3}$, $G_{6.3}$, and $A_{6.2}$ of helix I, $G_{7.1}$, $C_{7.2}$, $C_{7.3}$, and $G_{21.2}$ of helix II, and $G_{L2.4}$, $G_{14.4}$, $A_{14.3}$, $G_{14.2}$, and $C_{14.1}$ of helix III. The extension of the cleavages from helix III to the first nucleotides on both sides of the internal loop A_{12} and A_{15}, A_{16} may indicate a stacked conformation of adenosine residues in this region.

These results unambiguously confirm the overall secondary structure prediction as shown in Figure 1, but they begin to suggest the participation of nucleotides in the internal loop in tertiary interactions.

Probing the Tertiary Structure: Conserved Nucleotides in the Internal Loop Region

To investigate the role of individual nucleotides in the (formally) single-stranded regions, all possible single mutants were prepared and true initial reaction rates determined (Figure 4). Only four positions in the internal loop turned out to be absolutely immutable: U_8, C_{10}, and C_{11} in the upper part, and U_{20} in the lower part. Activity was reduced below the detection level if these positions were mutated. At position 9, there was a strong preference for G, with a 20- to 50-fold rate reduction in mutation. In position 12, purines were preferred 15-fold over pyrimidines. A and G differed only slightly from each other.

The lower part of the internal loop was found to be much more variable. Positions 15, 16, and 19 tolerated most point mutations while still retaining 10%–20% activity. At position 17, a U→C transition was accepted (35% of the wild-type activity), while transversions gave an over 50-fold rate reduction. At the nearly conserved position 18, the transition mutation A18G caused nearly complete loss of activity. Remarkably, the transversion mutants A18C and A18U showed higher activities than the transition mutant.

In general, the upper pentanucleotide loop segment U_8-A_{12} shows a much higher level of conservation than the lower hexanucleotide segment A_{15}-U_{20}.

Structural Elements of the 5′ Terminus

The 5′-terminal GGAG was part of the constant primer region during the selection of this ribozyme and therefore was not variable in the selected sequences. Mutation analysis now reveals (Figure 4) that this sequence is the most strongly conserved in the overall structure. No single mutations at all were tolerated in positions 1 and 2, and an over 50-fold drop in activity was seen at positions 3 and 4. This high level of conservation suggests its importance for the formation of the active structure.

Identification of Tertiary Interactions by Compensatory Double Mutations

To identify positions at which nucleotides interact with each other, two mutated strands were combined with one wild-type strand in the tripartite ribozyme activity assay, and the Diels-Alder reaction with biotin maleimide was studied (rescue experiments). All theoretically possible interstrand double mutants involving the 11 nucleotides of the internal loop and the four 5′-terminal nucleotides were prepared and assayed. Particular attention was paid to positions that indicated a high level of conservation in the single-mutation studies.

Two rescues were observed clearly, indicating a direct interaction between these positions. The first one was between the conserved U_{20} in the lower part of the internal loop and the highly conserved A_3 (Figure 5). While the reaction rate broke down almost completely in all single mutants, full activity could be restored when mutations were carried out in a pair-wise complementary fashion (at least for C_3:G_{20} and G_3:C_{20}). Interestingly, a U_3:A_{20} inversion was less well tolerated. The G_3:U_{20} "wobble" gave a weak but significant rescue, but not U_3:G_{20}. Some of the compensatory mutations are probably disfavored because of competition between alternative structures; for example, an adenine or a guanine at position 20 could pair directly with U8.

The second interaction was one between C_{19} and G_4 (Figure 5). Since C_{19} was not nearly as conserved as U_{20}, the magnitude of the rescues was much smaller. The general pattern, however, could be clearly recognized. Strong rescues were found in the diagonal, representing canonical base pairing, and weaker rescues were found for G:U, U:G, and U:U. These findings could be further validated by investigation of a panel of 49 different quadruple mutants (positions 3, 4, 19, and 20; data not shown).

Various other combinations were tested and were not found to yield compensatory increases in activity. In particular, this applied to the conceivable interactions G_1:C_{10}, G_1:C_{11}, G_2:C_{10}, G_2:C_{11}, G_9:C_{19}, and A_{12}:U_{17}.

Chemical Probing

The two chemical reagents diethylpyrocarbonate (DEPC) and dimethylsulfate (DMS) were used to probe the accessibility of individual nucleobases in end-labeled RNA (Figure 6) [16, 17]. DEPC modifies N7 of adenine, and DMS alkylates N3 of cytosine (direct probing) or N1 of adenines (indirect probing). To study the involvement of bases in secondary and tertiary interactions, the probing experiments were carried out under denaturing (1 mM EDTA, 80°C), semidenaturing (1 mM EDTA, 25°C), and native conditions (10 mM MgCl₂, 25°C), thereby gradually changing accessibility of certain positions to the chemical reagents (Figure 6).

Under denaturing conditions, six of the seven adenosines were carboethoxylated at their N7 by DEPC. Adenosine $A_{14.3}$, which is located in helix III, was only modified under these conditions, while $A_{6.2}$ did not react at all.
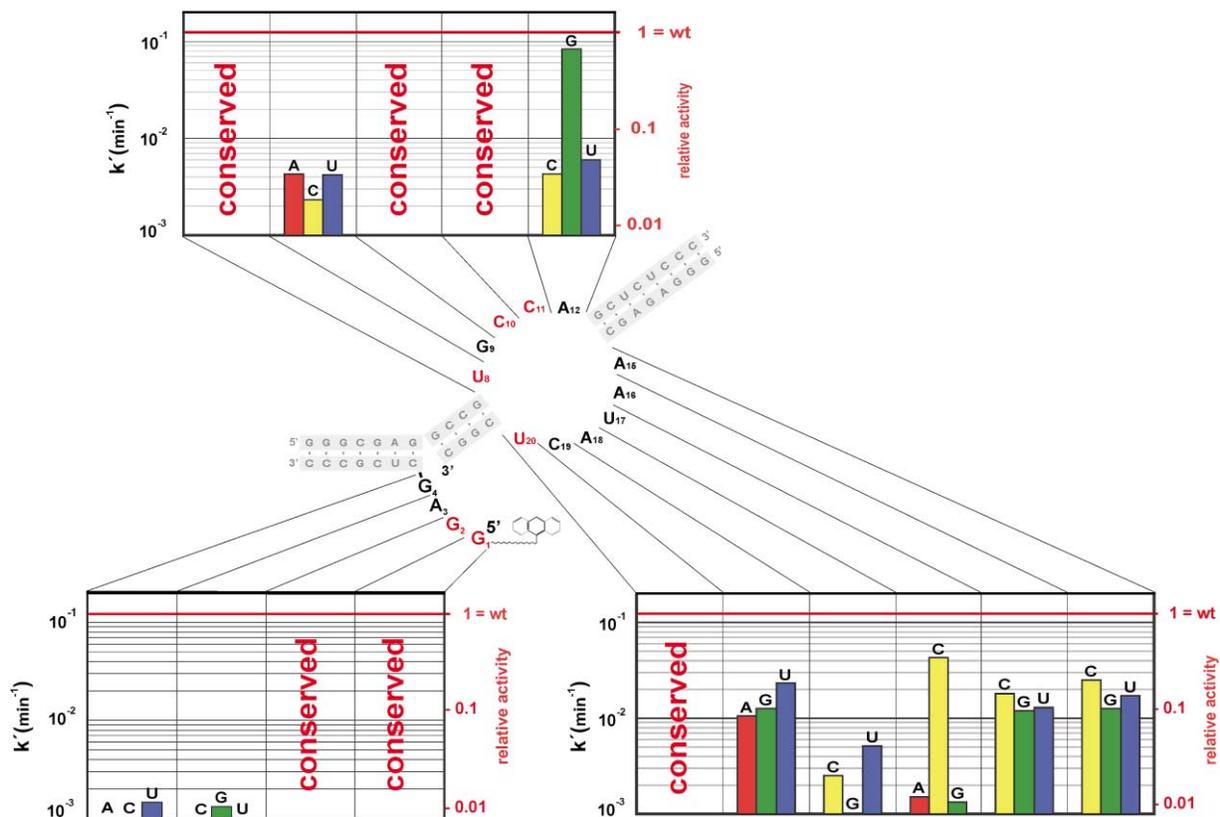
Figure 4. Single-Mutation Analysis of the Internal Loop and the 5′-Terminal Region

Values represent true initial rates measured over the first 15% of conversion. Numbers in red represent activities relative to the wild-type sequence.

Under semidenaturing conditions, the reactivity of the four adenosines in the internal loop toward DEPC was reduced (A_{12}, A_{15}, A_{16}, and A_{18}), while A_3 remained virtually unchanged. Under native conditions, the cleavage intensity at these five positions was further reduced, and for A_{12} and A_{18}, no modification at all could be detected.

Interestingly, probing of the adenine N1 position (DMS, indirect) gave a different picture. Most promi-

nently, A_{18} was more strongly modified under native conditions, compared to a semidenaturing environment (Figures 7A and 7C). Apparently, N1 of A_{18} became more exposed during tertiary structure formation. In contrast, A_{15} and A_{16} became less susceptible for modification upon tertiary structure formation, while A_{12} did not change. The same was found for A_3, which was only weakly modified at its N1 position.
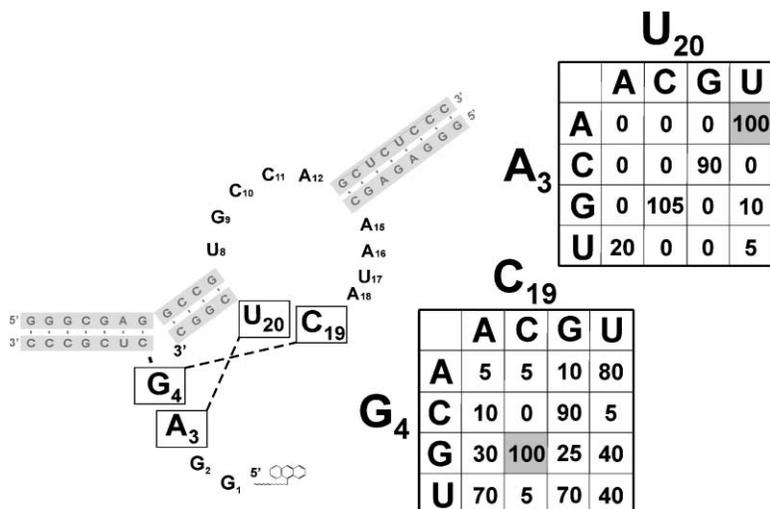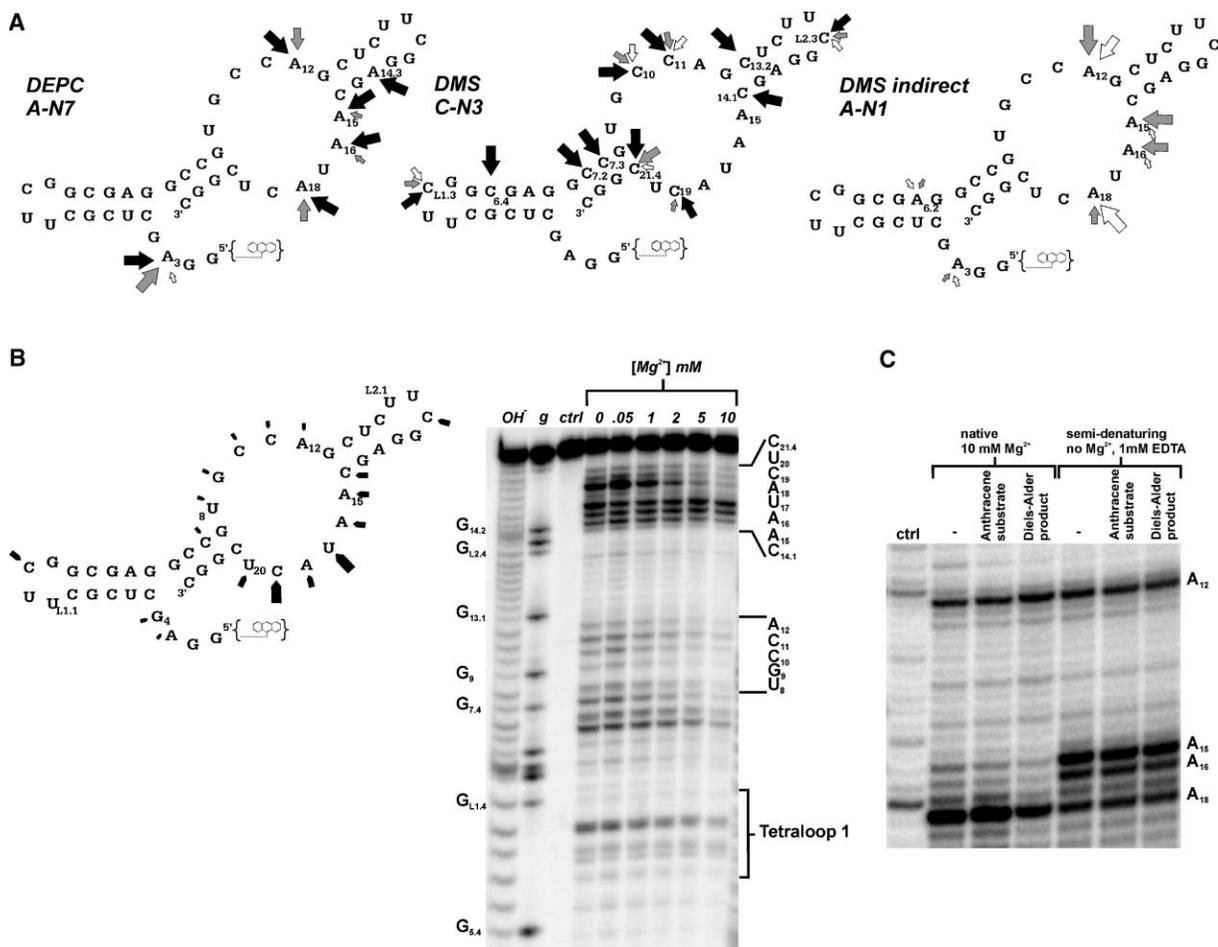

Figure 5. Compensatory Double-Mutation Analysis

Values represent single time point measurements (percent conversion after 30 min) relative to the tripartite wild-type ribozyme.

Figure 6. Chemical Probing of Tertiary Structure Elements

(A) Probing with DEPC and DMS. Probing was performed under denaturing (black arrows), semidenaturing (gray arrows), and native (white arrows) conditions. The size of the arrows corresponds to the intensity of cleavage or modification. Indirect probing was only performed under native and semidenaturing conditions.

(B) Lead probing. Lanes "OH⁻," "g," and "ctrl" correspond to alkaline hydrolysis, G sequencing with RNase T1, and control incubation (probing buffer, no lead ions), respectively, on a 15% denaturing polyacrylamide gel.

(C) Effects of a high concentration of anthracene substrate (0.5 mM) and Diels-Alder product (1 mM) on the DMS modification of RNA, analyzed by primer elongation and 12% denaturing PAGE. Lane "ctrl" corresponds to a control incubation. For assignment, four sequencing reactions were run in parallel (lanes not shown).

When DMS probing of cytosines (Watson-Crick position N3) was performed under denaturing conditions, cleavage of 11 out of 16 cytosine residues could be observed. Under semidenaturing conditions, residues C_{10}, C_{11}, C_{19}, and $C_{21.4}$ were modified with lower intensity, and upon tertiary structure formation, C_{10}, C_{11}, and C_{19} were further protected from methylation.

Probing with Pb²⁺

Lead probing was used to assess the general accessibility, unusual backbone conformations, or specific metal ion binding [18, 19]. Under native conditions, cleavage of the hexanucleotide stretch $A_{15}A_{16}U_{17}A_{18}C_{19}U_{20}$ of the asymmetric internal loop was found to be considerably stronger than that on the opposite side of this loop (Figure 6B). Within this strongly cleaved region, Pb²⁺-induced hydrolysis is particularly intensive at positions U_{17} and C_{19} and is slightly less intensive at A_{15} and A_{16}.

Increasing concentrations of Mg²⁺ ions diminished cleavage at U_{17} and C_{19}, while A_{15} and A_{16} appeared to be unaffected. Remarkably, there was almost no cleavage at A_{18}. Moderate cleavage was also detected at positions $C_{L1.3}$ and $C_{L2.3}$ of the two tetraloops.

Comparative Investigations in the Presence of Substrates and Products

To investigate structural changes associated with substrate or product binding, the described chemical probing experiments were performed in parallel with unmodified RNA and with anthracenemethylene-hexaethylene glycol covalently attached to the 5′ terminus of 49-mer RNA (Figures 3 and 6). Alternatively, free substrates and products were added to the assays in concentrations exceeding the measured K_m and K_i values [14, 15].

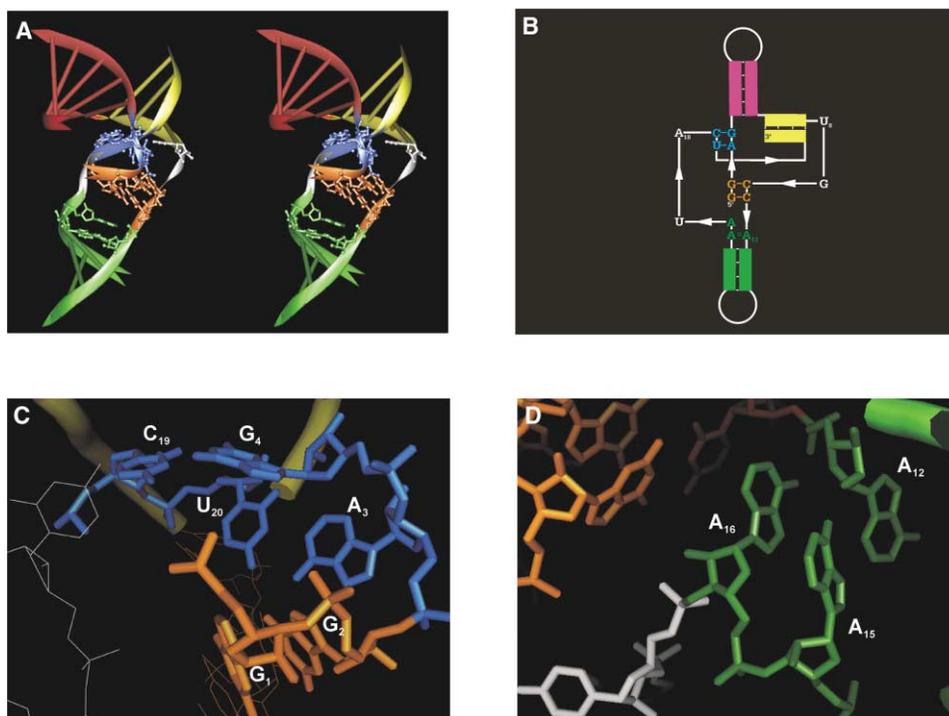Only very few and minor differences could be observed, all of which were located in the lower half of the

Figure 7. The Architecture of the Diels-Alderase Ribozyme

(A) Stereoview of the overall architecture of the ribozyme fold.
(B) Planar representation of the secondary structure with the proposed tertiary pairs.
(C) Tertiary interactions G_1:C_{11}, G_2:C_{10} (orange) and A_3:U_{20}, G_4:C_{19} (blue).
(D) Close-up of nucleotides A_{12}, A_{15}, and A_{16} (green).

asymmetric loop. The only more pronounced change was associated with position N1 in A_{18}, which showed the unusual exposure on tertiary structure formation (*vide supra*). At 1 mM Diels-Alder product concentration, this position became partly protected from modification (Figure 6C). This also applied to a lower extent to the other adenosines in the lower internal loop region. The presence of 0.5 mM of the anthracene substrate, however, did not affect the extent of modification. Adenosine A_{16} showed a slightly reduced cleavage by RNase S1 in the anthracene conjugate. C_{19} showed a weaker Mg^{2+} dependence of lead cleavage in the anthracene conjugate. At high concentrations of anthracene substrate, a slightly reduced lead cleavage in the lower internal loop region (A_{15}–U_{20}) could be observed. No influence of free dienophile or cycloaddition product (both up to 5 mM concentration) on the lead cleavage pattern could be detected.

No differences at all were observed for the upper half of the internal loop or for the 5′ terminus. There are no experimental data indicating dramatic changes in the overall structure upon substrate or product binding.

Data Integration

The data gathered from mutational analysis and chemical as well as enzymatic probing are in good agreement with each other and with the sequence information derived from the in vitro-selected families [13]. All proposed helices could be unambiguously supported. One central element of the secondary structure is the asymmetric internal loop, comprised of a pentanucleotide, $U_8G_9C_{10}C_{11}A_{12}$, and a hexanucleotide, $A_{15}A_{16}U_{17}A_{18}C_{19}U_{20}$. The hexanucleotide shows a significantly higher level of sequence variability than the pentanucleotide, and it is more susceptible to both enzymatic and lead-induced hydrolysis. DEPC and DMS probing data generally also support this notion (see Figures 3 and 6), but in addition indicate that both the adenosines and the cytidines become increasingly protected from chemical modification when folding into a tertiary structure. Four nucleotides of the internal loops were found to be absolutely invariant (at least at the single point mutation level). These are U_8, C_{10}, C_{11} (in the pentanucleotide), and U_{20} (in the hexanucleotide).

Remarkably, the four 5′-terminal nucleotides, G_1, G_2, A_3, and G_4, were found to be highly conserved, and the probing experiments indicate protection from modification under native conditions. This is another interesting example of a constant primer binding site being turned into an essential structural element in a SELEX experiment [20].

The search for additional interactions demonstrated that A_3 and G_4 pair with U_{20} and C_{19}, respectively. Single mutations dramatically reduce the catalytic activity, while compensatory double mutations yield significant rescues [21]. In addition, we propose here that G_1 and G_2 may interact, respectively, with C_{11} and C_{10}. This proposal is based mainly on the observations that all four nucleotides have been found to be absolutely invariant

in the single-mutation analysis, and that C_{10} and C_{11} became increasingly protected from chemical modification when going from denaturing to native conditions. While these findings are suggestive with regard to the involvement of these nucleotides in tertiary contacts, direct proof of specific interactions between these positions could not be obtained. The fact that we could not confirm these interactions by compensatory double mutations does not mean that they do not exist, as the G:C base pairs may interact with the substrates, transition state, or product in a specific manner and, consequently, could therefore not be replaced by other base pairs.

The overall structure of this molecule appears to be rather robust under native conditions. No fundamental structural changes (like formation or breakdown of helices or tertiary interactions) are observed upon substrate or product binding. The only region where there are some subtle changes in accessibility to chemical probes is the A_{15}–A_{18} region of the internal loop, which may – although not highly conserved – contribute directly or indirectly to the catalytic pocket.

Molecular Modeling

Based on these new experimental data, we constructed various three-dimensional models of the ribozyme in order to examine whether all constraints could be satisfied within one three-dimensional structure. Although the relative orientations of the three helices cannot be deduced from the data provided, the experimental data still contain enough geometrical and stereochemical constraints to derive an overall architecture of the ribozyme that rationalizes the experimental data and is compatible with the formation of a catalytic pocket (Figure 7).

According to this model, the three helices point into three different directions in space in a Y-shaped fashion. The base is formed by helix III and the two arms by helices I and II (the closing tetraloop sequences have been omitted in the drawings for clarity). The minor/shallow groove of helix I faces the major/deep groove of helix II. The binding and catalytic pockets form at the junction between the base helix and the two arm helices (Figure 7A). This arrangement leads to a compact overall structure. The catalytic pocket is formed primarily from the asymmetric internal loop and the 5′-terminal GGAG end, and this formation involves various tertiary interactions (vide infra). The central element of the structural model is made up by the internal loop. The positions of the highly conserved nucleotides within this structural model are all located in close proximity to each other.

The adenines in the internal loop present interesting reactivities at N1 and N7. For A_{12} and A_{15}, we interpreted those data as indicating the formation of a sheared $A_{12}oA_{15}$ (*trans* sugar/Hoogsteen) [22] base pair (Figure 7D). Such a choice fits with the observation that A_{12} can be replaced favorably by guanosine (which would lead to the more stable sheared GoA pair), but not by pyrimidine bases. For A_{12}, protection at N7 is due to stacking, and for A_{15}, protection is due to H bonding to C2 of A_{12}. At N1, A_{12} should be fully accessible, while at A_{15}, it should be partially protected by H bonding to the hydroxyl group of A_{12}. A_{16} is protected at N1 and at N7. It was

built with stacking contacts on the $A_{12}oA_{15}$ sheared base pair and with a potential H bond between the amino N6 and the phosphate group of A_{12}. These positions are compatible with the observations that A_{16} is prone to substitution by the three other nucleobases. Interestingly, U_{17} can only be substituted by cytidine. Residue 17 faces the entrance of the catalytic pocket, and purine substitutions may not allow access of the large anthracene ring to the catalytic cavity for steric reasons. Residue 17 is also important for the looping out of the single strand. A_{18} is reactive at position N1 and is conserved to a very high degree; in particular, no substitution by guanosine is possible. This conservation is probably due to the closeness to the entrance of the catalytic pocket. All three adenine residues on this side of the internal loop show altered reactivity toward modification by DEPC under the different reaction conditions, and this altered reactivity confirms their involvement in tertiary interactions.

The opposite parts of the internal loop are clamped by the four nucleotides on the 5′ end, forming the interacting base pairs G_1C_{11}, G_2C_{10}, A_3U_{20}, and G_4C_{19} (Figure 7C). The two sets of base pairs are not co-axial: there is a large kink between G_2C_{10} and A_3U_{20}. Furthermore, while G_1C_{11} and G_2C_{10} are stacked regularly, there is a wedge between A_3U_{20} and G_4C_{19}. We propose that this wedge forms the main binding pocket of the catalytic site. While position 19 tolerates substitutions, U_{20} is one of the conserved positions. Interestingly, C_{19} is only weakly modified by DMS even under denaturing conditions, thus its involvement in base pairing seems likely. At positions 19:4, non-Watson-Crick pairs like GoU, GoA or GoG are compatible with catalysis. C_{10} and C_{11} are both highly conserved. The purine nucleotides face the catalytic cavity, which could explain the constraints at these positions. C_{10} and C_{11} are less modified at their N3 atoms under semidenaturing conditions compared to denaturing conditions, and this indicates the involvement of this position in base pairing.

For each of the base pairs G_1C_{11} and G_2C_{10}, the Watson-Crick scheme is postulated. These base pairs point into the pocket, providing numerous ring nitrogens and amino groups for interactions with the substrates (Figure 7C), which might explain why these base pairs cannot be substituted by others. The pair G_4C_{19} is also of the canonical Watson-Crick type, which is in good agreement with the weak modification of N3 of C_{19} by DMS. A_3U_{20} is proposed to form another Watson-Crick pair. This way, the tetranucleotide clamp can efficiently hold together the opposite sides of the internal loop. Interestingly, all conserved nucleotides are located in close proximity to each other and form a pocket that is wide enough to accommodate substrates or products. Residues U_{17} and A_{18} are at the entrance to the pocket, while residues A_{12}, A_{15}, and A_{16} form a motif stacking below the first set of base pairs G_1C_{11} and G_2C_{10}. All those loop residues show some sensitivity to product binding.

Based on these interactions, we now propose a more complex structure of the Diels-Alderase ribozyme (Figure 7B). The ribozyme forms an unusual pseudoknot structure resembling a clamp or a push-button. The 5′-terminal GGAG serves to clamp together the opposite sites of the asymmetric internal loop. The tertiary inter-

actions basically cut the internal loop into two halves, with the (rather conserved) U_8 and G_9 on one side of the clamp and the less conserved A_{12}, A_{15}, A_{16}, U_{17}, and A_{18} on the other.

The pseudoknot structure, actually a double pseudoknot, is highly unusual, since it involves two antiparallel base-pairing interactions. There are four direct connections between different helical elements without unpaired connectors (G_4, helix I; U_{20}, helix II; G_2-A_3, helix I–helix II), which could impose severe strain on the system. On the other hand, this appears to be an efficient way to achieve a rather dense and stable packing of helical elements.

The constraints of this model are mainly dictated by the stereochemistry and the bulkiness of the helices and base pairs. Other attempted models, involving alternative tertiary interactions, showed disagreement with parts of the experimental data. Depending on the significance of the individual data, different structural features of this model may have a different reliability. While the existence of the three secondary helices has been unambiguously proven, and very strong support exists for the A_3U_{20} and G_4C_{19} interactions, support is somewhat weaker for the stacked A_{12} A_{15} A_{16} formation, followed by the G_1C_{10} and G_2C_{11} tertiary interaction. Only little data are available to evaluate the positions of the formally unpaired nucleotides U_8, G_9, U_{17}, and A_{18} (see Figure 7B), and no experimental data at all are available for deducing the relative orientation of helices I, II, and III with respect to each other.

One important question is whether all identified elements of higher structure (helices I–III, A_3:U_{20}, G_4:C_{19}, G_1:C_{11}, G_2:C_{10}) are present at the same time in one ribozyme conformation, or whether there are major structural changes during the catalytic cycle, with one set of interactions (i.e., one conformation) being responsible for substrate binding and another one being responsible for actually accelerating the reaction. The comparison of probing data in the presence and absence of substrates and products indicates that there are no fundamental changes in the secondary and tertiary structure upon binding. Apparently, the formation of the higher-order structure is independent on the presence of substrates. Unlike many previously described aptamers, the Diels-Alderase ribozyme seems to form a stable, pre-structured catalytic pocket that can accommodate the substrates and accelerate the reaction.

Significance

RNA molecules can catalyze various chemical reactions, including carbon-carbon bond formation by Diels-Alder reaction. To obtain insight into the chemical basis of catalysis, single and compensatory double-mutation analysis as well as chemical and enzymatic probing experiments on a previously identified Diels-Alderase ribozyme were conducted. The experimental results were used to construct an architecture of this ribozyme. The secondary structure was found to consist of three helices, a large internal loop, and four highly conserved nucleotides at the 5′ terminus. The overall shape of the proposed three-dimensional

structure is based on two additional tertiary base interactions between bases at the 5′ end and bases at the internal loop. First, the tertiary interactions between A_3 and G_4 at the 5′ end and U_{20} and C_{19} at the longer side of the internal loop are indicated by double-mutation rescue experiments. Second, base pairings between G_1 and G_2 at the 5′ end and C_{11} and C_{10} at the shorter side of the internal loop that are indicated by DMS probing of the N3 position of the two cytosines have been proposed. Thus, the four 5′-terminal bases connect both sides of the internal loop in a very unusual fashion by a double pseudoknot with a large kink between the two 2 base pair helices. One of these helices presents a wedge that is proposed to constitute the catalytic pocket. Entrance to the pocket is monitored by the other residues of the 3′ side of the conserved loop. Comparative probing experiments in the absence and presence of substrates or products reveal no major structural changes, indicating a preformed catalytic pocket. This preformed environment provides a valuable asset in the context of RNA catalysis.

Experimental Procedures

Oligonucleotide Synthesis and Labeling

DNA and RNA oligonucleotides were synthesized by phosphoramidite chemistry (Expedite 8900 synthesizer, standard protecting groups, standard deprotection). T7 transcription reactions were performed as described [23]. Incorporation of oligo-(ethylene glycol)-tethered anthracene was performed either chemically by using the respective phosphoramidite, or by transcription initiation [24]. Oligonucleotides were purified by HPLC and/or PAGE. Oligonucleotides and conjugates were ^{32}P-labeled at their 5′ or 3′ end with polynucleotide kinase, T4 RNA ligase, or Klenow fragment of DNA Polymerase I, according to standard procedures [25]. For the mutation studies, 11-mer anthracene conjugates were synthesized chemically, while 18-mers and 24-mers were prepared by run-off transcription.

Ribozyme Activity Assays

Two different gel-electrophoretic assays were used. For verification of secondary structure elements (Figure 2) and for the rapid screening for compensatory rescues (Figure 5), half-quantitative data were considered sufficient, and single time point measurements with a standard reaction time of 30 min were carried out. For single-mutation analysis (Figure 4), true initial rates of product formation were determined by monitoring the first 15% of conversion. A wild-type combination of the unmutated strands was always run in parallel and used as reference. All measurements were done in triplicate.

0.5 μM (5 pmol) 24 nt RNA, 1 μM (10 pmol) 18 nt RNA, 1 nM (10 fmol) 3′ ^{32}P-labeled 11 nt-RNA-hexa(ethylene glycol)-anthracene were hybridized in 30 mM Tris-HCl (pH 7.4), 300 mM NaCl by heating to 90°C and slow cooling. After addition of $MgCl_2$ (final concentration of 80 mM), the reaction was started by adding fresh solution of biotin maleimide dissolved in DMSO (final concentration of 5 μM biotin maleimide, 2% DMSO). In appropriate intervals, aliquots were withdrawn and quenched by adding two volumes of stop mix (20 mM Na_2EDTA, 0.7% [v/v] β-mercaptoethanol, 80% [v/v] formamide). The quenched reaction aliquots were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE), typically on 20% gels. The bands of 11-mer-linked anthracene and Diels-Alder products were quantified by phosphorimaging or scintillation counting (see the inset in Figure 1B).

To analyze the effect of the deletion of helix I (Figure 2, construct "U_8 tether"), a one-stranded assay format had to be used. The reaction was studied by fluorescence spectroscopy, monitoring the decay of fluorescence emission at 419 nm of a covalently tethered anthracene as described [13]. A corresponding one-stranded wild-type ribozyme (Figure 1A) with tethered anthracene was used as a positive control.

Nuclease Probing

The ^{32}P-labeled RNA (50.000 cpm/sample) was supplemented with tRNA carrier to a final concentration of 0.15 μg/μl in H$_2$O and renatured by heating to 65°C for 3 min and cooling to room temperature for 20 min. Limited digestions with nucleases S1, T2, T1, U2, and V1 were carried out in 40 mM Tris-HCl (pH 7.5), 40 mM NaCl, and 10 mM MgCl$_2$. For reactions with nuclease S1, ZnCl$_2$ was added to a final concentration of 1 mM. Reactions were performed at room temperature for 10 min with 30 units of nuclease S1, 0.1 units of nuclease T2, 0.05 units of nuclease T1, 2.5 units of nuclease U2, and 0.2 units of nuclease V1. The reactions were terminated by addition of an equal volume of 0.6 M sodium acetate, 3 mM EDTA, and 0.1 μg/μl tRNA and subsequent phenol extraction. The RNA was precipitated with the 10-fold volume of acetone containing 2% lithium perchlorate, and the pellet was rinsed with acetone. Samples were mixed with a 7 M urea/dyes/20 mM EDTA solution and loaded on 15% polyacrylamide denaturing gels. Electrophoresis was performed at 1500 V for 3 hr, followed by phosphorimaging.

Chemical Modification

Direct Probing

For chemical modification of adenine residues with DEPC and cytosine residues with DMS, ^{32}P-labeled RNA was supplemented with tRNA carrier to a final concentration of 0.03 μg/μl in H$_2$O and renatured by heating to 65°C for 3 min and cooling to room temperature for 20 min. Buffer was added to a final concentration of 50 mM sodium cacodylate (pH 7.5), 300 mM KCl, and 10 mM MgCl$_2$ for native conditions or 50 mM sodium cacodylate and 1 mM EDTA for semidenaturing and denaturing conditions. 5 vol% of DEPC were added, and the samples were agitated thoroughly for 50 min at room temperature. 5 vol% of a solution of 10% DMS in ethanol was added, and the samples were incubated at room temperature for 9 min. Reactions under denaturing conditions took place at 80°C for 2 min in both cases. Reactions were terminated by the addition of sodium acetate to a final concentration of 0.3 M and standard ethanol precipitation (twice). The pellet of the DMS reaction was dissolved in 20 μl of a 10% solution of hydrazine in water and incubated at 0°C for 10 min, followed by ethanol precipitation (twice). Pellets of DEPC and DMS reactions were dissolved in 15 μl of 9% aniline-acetate (pH 4.5) buffer and incubated at 60°C for 10 min in the dark. After ethanol precipitation, the RNA was dissolved in 7 M urea and loaded on a 15% denaturing polyacrylamide gel. Electrophoresis was performed at 1500 V for 3 hr, followed by phosphorimaging.

Indirect Probing

For chemical modification of adenine residues with DMS, an elongated RNA construct was used that consisted of the 49-mer ribozyme sequence extended at its 3′ end by a 20 nt spacer sequence and a 20 nt primer binding site. The RNA (15 pmole) was supplemented with tRNA carrier to a final concentration of 0.12 μg/μl in H$_2$O and renatured by heating to 65°C for 3 min and cooling to room temperature for 20 min. Buffer was added to a final concentration of 50 mM sodium cacodylate (pH 7.5), 300 mM KCl, and 10 mM MgCl$_2$ for native conditions or 50 mM sodium cacodylate and 1 mM EDTA for semidenaturing conditions. 5 vol% of a solution of 10% DMS in ethanol was added, and the samples were incubated at 25°C for 5 and 15 min. Reactions were terminated by the addition of sodium acetate to a final concentration of 0.3 M and standard ethanol precipitation (twice). For primer extension, the pellets were dissolved in 4 μl H$_2$O and 100.000 cpm/sample 5′-labeled primer was added. Samples were heated for 2 min at 90°C and immediately placed on ice. 1 μl of 0.1 M DTT and 2 μl of 5× First Strand Buffer (Invitrogen) were added, and the samples were kept for 20 min at room temperature. 2.5 μl of a solution containing 2 mM of each dNTP was added, followed by 120 U Superscript II RTase (Invitrogen), and the samples were incubated for 30 min at 48°C. 10 μl of 7 M urea was added, and the samples were loaded on a 12% denaturing polyacrylamide gel. Electrophoresis was performed at 1500 V for 3 hr, followed by phosphorimaging.

Pb^{2+} Ion-Induced Cleavage

For the cleavage reaction with Pb^{2+}, 10 μl ^{32}P-labeled RNA (50.000 cpm/sample) was supplemented with tRNA carrier to a final RNA concentration of 0.125 μg/μl in H$_2$O and renatured by heating to

65°C for 3 min and cooling to room temperature for 20 min. 20 μl buffer was added to a final concentration of 40 mM Tris-HCl (pH 7.5), 40 mM NaCl, and 10 mM MgCl$_2$. The reaction was started by the addition of freshly prepared lead acetate solution (final concentration 0.75 mM). Aliquots were withdrawn after 15, 30, and 45 min of incubation at room temperature, and the reaction was terminated by mixing with 7 M urea/dyes/20 mM EDTA solution. Samples were loaded on 15% polyacrylamide denaturing gels. Electrophoresis was performed at 1500 V for 3 hr, followed by phosphorimaging.

Identification of Cleavage Sites

The RNA cleavage products were assigned by running in parallel with products of alkaline RNA hydrolysis and limited ribonuclease T1 digestion of the same RNA (supplemented with tRNA carrier to a final RNA concentration of 1.1 μg/μl). Alkaline hydrolysis ladders were generated by incubating the RNA solution with 10 μl of 50 mM sodium bicarbonate (pH 9.0) at 90°C for 8 min. Partial T1 nuclease digestion was performed under denaturing conditions (12.5 mM sodium citrate [pH 4.5], 0.5 mM EDTA, 3.5 M urea) with 0.05 units of the enzyme at 55°C for 10 min.

For the indirect probing of adenine with DMS, sequencing reactions were run in parallel. For this purpose, 6.25 μl of a solution containing the RNA construct (20 pmole) and 200.000 cpm 5′-labeled primer were heated at 90°C for 2 min, followed by cooling to room temperature over 15 min on the bench. A total of 37.5 μl of a solution was prepared by adding 5 μl 0.1 M DTT, 10 μl 5× First Strand Buffer, 600 U Superscript II RTase (Invitrogen), and water. Aliquots (7.5 μl) were mixed with 2.5 μl of a solution containing 2 mM of each dNTP and 4 mM of the respective ddNTP for the four sequencing reactions and incubated at 48°C for 30 min.

Molecular Modeling

The molecular assembly of the architecture of the ribozyme was performed by using the programming and modeling tools contained in the MANIP package [26]. The geometrical and stereochemical refinement was done by using NUCLIN/NUCLSQ [27], and the drawing of Figure 7A was done with DRAWNA2.1 [28]. Figures 7C and 7D were prepared by using PYMOL [29].

References

1. Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.-P., and Pardi, A. (1997). Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. Nat. Struct. Biol. *4*, 644–649.
2. Patel, D.J., and Suri, A.K. (2000). Structure, recognition and discrimination in RNA aptamer complexes with cofactors, amino acids, drugs and aminoglycoside antibiotics. J. Biotechnol. *74*, 39–60.
3. Baugh, C., Grate, D., and Wilson, C. (2000). 2.8 Å crystal structure of the malachite green aptamer. J. Mol. Biol. *301*, 117–128.
4. Tereshko, V., Skripkin, E., and Patel, D.J. (2003). Encapsulating streptomycin within a small 40-mer RNA. Chem. Biol. *10*, 175–187.
5. Hampel, A., and Cowan, J.A. (1997). A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage. Chem. Biol. *4*, 513–517.

6. Lebruska, L.L., Kuzmine, I.I., and Fedor, M.J. (2002). Rescue of an abasic hairpin ribozyme by cationic nucleobases: evidence for a novel mechanism of RNA catalysis. Chem. Biol. *9*, 465–473.

7. Scott, E.C., and Uhlenbeck, O.C. (1999). A re-investigation of the thio effect at the hammerhead cleavage site. Nucleic Acids Res. *27*, 479–484.

8. Suzumura, K., Yoshinari, K., Tanaka, Y., Takagi, Y., Kasai, Y., Warashina, M., Kuwabara, T., Orita, M., and Taira, K. (2002). A reappraisal, based on (31)P NMR, of the direct coordination of a metal ion with the phosphoryl oxygen at the cleavage site of a hammerhead ribozyme. J. Am. Chem. Soc. *124*, 8230–8236.

9. Ryder, S.P., Ortoleva-Donnelly, L., Kosek, A.B., and Strobel, S.A. (2000). Chemical probing of RNA by nucleotide analog interference mapping. Methods Enzymol. *317*, 92–109.

10. Lilley, D.M. (2004). Analysis of global conformational transitions in ribozymes. Methods Mol. Biol. *252*, 77–108.

11. Doherty, E.A., and Doudna, J.A. (2001). Ribozyme structures and mechanisms. Annu. Rev. Biophys. Biomol. Struct. *30*, 457–475.

12. Doudna, J.A., and Cech, T.R. (2002). The chemical repertoire of natural ribozymes. Nature *418*, 222–228.

13. Seelig, B., and Jäschke, A. (1999). A small catalytic RNA motif with Diels-Alderase activity. Chem. Biol. *6*, 167–176.

14. Seelig, B., Keiper, S., Stuhlmann, F., and Jäschke, A. (2000). Enantioselective ribozyme catalysis of a bimolecular cycloaddition reaction. Angew. Chem. Int. Ed. Engl. *39*, 4576–4579.

15. Stuhlmann, F., and Jäschke, A. (2002). Characterization of an RNA active site: interactions between a Diels-Alderase ribozyme and its substrates and products. J. Am. Chem. Soc. *124*, 3238–3244.

16. Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.P., and Ehresmann, B. (1987). Probing the structure of RNAs in solution. Nucleic Acids Res. *15*, 9109–9128.

17. Brunel, C., and Romby, P. (2000). Probing RNA structure and RNA-ligand complexes with chemical probes. Methods Enzymol. *318*, 3–21.

18. Brown, R.S., Dewan, J.C., and Klug, A. (1985). Crystallographic and biochemical investigation of the lead(II)-catalyzed hydrolysis of yeast phenylalanine tRNA. Biochemistry *24*, 4785–4801.

19. Giege, R., Helm, M., and Florentz, C. (1999). Chemical and enzymatic probing of RNA structure. In Comprehensive Natural Products Chemistry, D. Barton and K. Nakanishi, eds. (London: Pergamon Press), pp. 63–80.

20. Burke, D.H., and Hoffman, D. (1998). A novel acidophilic RNA motif that recognizes coenzyme A. Biochemistry *37*, 4653–4663.

21. Leontis, N.B., Stombaugh, J., and Westhof, E. (2002). The non-Watson-Crick base pairs and their associated isostericity matrices. Nucleic Acids Res. *30*, 3497–3531.

22. Leontis, N.B., and Westhof, E. (2001). Geometric nomenclature and classification of RNA base pairs. RNA *7*, 499–512.

23. Milligan, J.F., Groebe, D.R., Witherell, G.W., and Uhlenbeck, O.C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. *15*, 8783–8798.

24. Seelig, B., and Jäschke, A. (1999). Ternary conjugates of guanosine monophosphate as initiator nucleotides for the enzymatic synthesis of 5′-modified RNAs. Bioconjug. Chem. *10*, 371–378.

25. Sambrook, J., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, Third Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

26. Massire, C., and Westhof, E. (1998). MANIP: an interactive tool for modelling RNA. J. Mol. Graph. Model. *16*, 197–205, 255–197.

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

28. Massire, C., Gaspin, C., and Westhof, E. (1994). DRAWNA: a program for drawing schematic views of nucleic acids. J. Mol. Graph. *12*, 201–206, 196.

29. DeLano, W.L. (2002). The PyMOL Molecular Graphics System (http://www.pymol.org).