Annu. Rev. Biophys. Biomol. Struct. 1999. 28:57–73

RNA FOLDS: Insights from Recent Crystal Structures

Adrian R. Ferré-D'Amaré and Jennifer A. Doudna
Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520-8114;
e-mail: ferre@csb.yale.edu, doudna@csb.yale.edu

KEY WORDS: coaxial stack, helix packing, strand crossover, tertiary interaction, cation binding site

ABSTRACT

An RNA fold is the result of packing together two or more coaxial helical stacks. To date, four RNA folds have been determined at near-atomic resolution by X-ray crystallography: transfer RNA, the hammerhead ribozyme, the P4–P6 domain of the *Tetrahymena* group I intron, and the hepatitis delta virus ribozyme. All four folds result in RNAs that are considerably more compact than isolated A-form duplexes. These structures illustrate, to varying degrees, three modes of fold stabilization: association of complementary molecular surfaces, stabilization of close RNA packing by binding of cations, and stabilization through pseudoknotting.

CONTENTS

1056-8700/99/0610-0057$08.00

SUMMARY AND PERSPECTIVES

Twenty-five years ago, the structure determinations of $tRNA^{Phe}$ demonstrated that ribonucleic acid can adopt a compact, globular fold (27, 48). In the last five years, the structures of three new RNA folds have been determined at near-atomic resolution: the hammerhead ribozyme (45, 51), the P4–P6 domain of the *Tetrahymena* group I intron (4), and the hepatitis delta virus (HDV) ribozyme (18a). Side-by-side comparison of the four known RNA folds shows that they all are constructed by packing together coaxially stacked helices. To different extents in the four folds, the packing appears to be stabilized by nonhelical or tertiary hydrogen bonding networks, coordination of metal ions, and by the connectivity of the RNA backbone.

In this review, we first summarize the overall features of the four known RNA folds, and then compare and contrast their size, complexity, and inferred or experimentally verified modes of fold stabilization. We are witnessing rapid growth of structural information on RNA and can expect many more structure determinations of new RNA folds in the near future. The structural information will be pivotal in elucidating the biochemical roles and capabilities of RNA in biological systems.

TRANSFER RNA AND CONCEPTUAL FRAMEWORK

The A-form antiparallel double helix composed of Watson-Crick base-pairs is the basic element of RNA structure. Helices can be as short as two base pairs, and in principle there is no maximum length. Most RNA helices are not comprised purely of Watson-Crick pairs, but contain noncanonical pairs and bulges that are accommodated within the helical framework. Local structural elements that are observed multiple times in near-atomic or atomic resolution structures are called motifs (33). Double helices can be formed by adjacent sequences, in which case the result is a hairpin, or by complementary nucleotides that are distant in the primary sequence. The set of helices that are formed by an RNA and their interconnections, sometimes incorrectly called topology, is the secondary structure. This usage is different from that in proteins, because nucleic acid helices are double stranded.

As a result of the strength of base-stacking interactions in solution (60), RNA helices form coaxial stacks, the next higher level of structure. We define an RNA fold as the result of the packing of at least two coaxial helical stacks against each other. Although some simple hairpins or coaxial stacks can perform limited

biochemical functions, such as providing binding sites for their cognate proteins (e.g. 9, 34, 39, 46) or small ligands (15, 17), most RNA molecules that perform sophisticated biochemistry are thought to have specific folds.

Because they are part of one molecule, the helices that form stacks and folds must be connected through their main-chain, or phosphate-ribose backbones. These connections are called strand crossovers. The crossovers can consist of just the phosphate backbone between two adjacent nucleotides that belong to different helical stacks, or of numerous nucleotides. In addition, segments of the RNA contact each other through base pairing (to form double helices), and through nonhelical hydrogen bonding, ionic, and van der Waals interactions. These interactions, when they occur between nonadjacent nucleotides, are called tertiary contacts and can involve both base and backbone atoms. Triplexes and quadruplexes are defined as tertiary structural elements in which atoms of the bases of three or four RNA strands make contact.

Transfer RNA (Figure 1; see color Figure 2 at end of volume) illustrates the concepts just defined. The sequence or primary structure of most tRNAs can be arranged into the classical cloverleaf consisting of four helical segments: the acceptor, anticodon, D, and T stems or arms (54). This RNA adopts an L-shaped fold in which two coaxial helical stacks meet at right angles (27, 48). The acceptor and T stems stack together, whereas the D and anticodon stems form a second continuous helix. The two extended helices interact through their terminal T and D loops as well as through nucleotides in the sequences joining the four helical segments. The key nucleotides involved in tertiary interactions are conserved, implying that most tRNAs have essentially the same fold. The majority of these tertiary interactions result in the formation of short triplexes. Magnesium cations located in nonhelical regions of the molecule may stabilize the fold (54). We now turn to the three RNA folds that have been determined at near-atomic resolution in recent years.

THE HAMMERHEAD RIBOZYME

Background

The hammerhead ribozyme is a catalytic RNA that was discovered in the RNA genomes of some plant pathogens and also in satellite RNAs from newt mitochondria (58). Like the unrelated hairpin, hepatitis delta virus, and Varkud satellite ribozymes, this ribozyme catalyzes a trans-esterification, cleaving its RNA backbone at a specific phosphodiester bond within one of the helices and yielding products with 5′-hydroxyl and 2′,3′-cyclic phosphate termini. The hammerhead RNA consists of three base-paired stems connected by two invariant single-stranded regions. The hammerhead appears to be a metalloenzyme, requiring one, or possibly several, metal ions for activity (40).

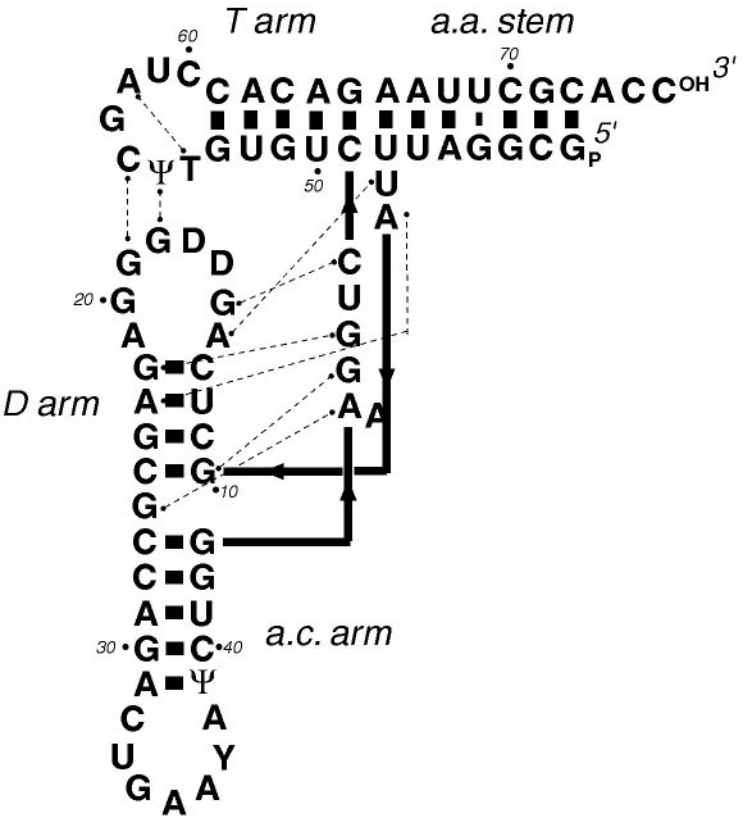

Figure 1 Secondary structure of $tRNA^{Phe}$ (after 22). Methylation of base and backbone atoms is not indicated for simplicity. Tertiary contacts are indicated by *thin dotted lines* with *dumbbell ends*. Main-chain crossovers are represented by *thick lines* with *embedded arrowheads*. Note the presence of four helices (aminoacyl acceptor—*a.a.*, anti-codon—*a.c.*, *D*, and *T*) in two coaxial stacks.

Structure

Two independent structure determinations of the hammerhead ribozyme have been reported. In the first, a two-stranded hammerhead RNA included an all DNA "substrate" strand. Because the nucleophile in the trans-esterification reaction is the 2′-hydroxyl group adjoining the scissile bond, this prevented self-cleavage during crystallization (45). In the second crystallographic study, the substrate strand was modified exclusively at the nucleophile by utilizing a synthetic RNA in which the reactive hydroxyl group had been replaced by a methoxyl function (51). The two crystal structures are similar

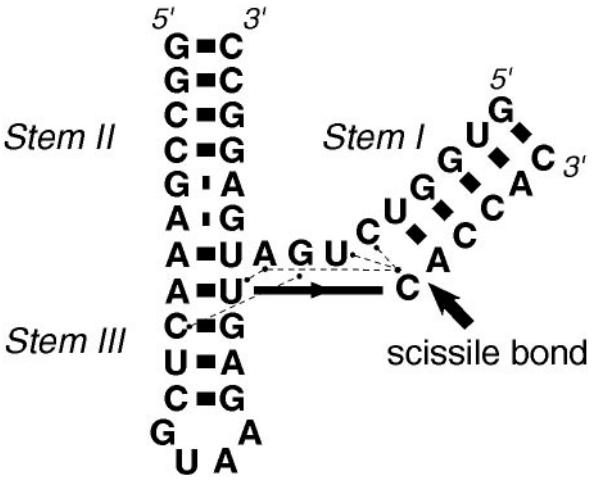

Figure 3 Secondary structure of a hammerhead ribozyme construct employed for structure determination (51).

(root-mean-square difference = 0.73 Å). RNA segments previously thought to be single-stranded form a helical segment, with noncanonical pairs, which stacks between stems II and III. Stem I branches off the stack to form a Y-shaped molecule with a U-turn at the junction (Figure 3; see color Figure 4). The nucleotides at the junction form a putative active site pocket.

These two structures of the hammerhead ribozyme appear to represent the ground state conformation of this catalyst, because the reactive groups are far away from the in-line conformation they must adopt in the transition state (16, 37). Several cation binding-sites have been observed crystallographically in the hammerhead ribozyme, and their roles in catalysis are the subject of numerous studies. These have been reviewed elsewhere (52, 61).

THE P4–P6 DOMAIN OF THE *TETRAHYMENA* SELF-SPLICING INTRON

Background

Group I introns, defined by a conserved catalytic core and reaction pathway, splice precursor RNAs to form mature ribosomal, transfer, or messenger RNAs (7). Half of the conserved core in the *Tetrahymena thermophila* intron resides in an independently folding domain consisting of the base-paired (P) regions P4 through P6 (P4–P6) (35). The P4–P6 domain alone folds into a structure whose chemical protection pattern is very similar to that seen for the P4–P6

region of the intact intron (8, 35, 36). Furthermore, the P4–P6 domain can assemble with the remainder of the intron in *trans* to reconstitute catalytic activity (12). Folding of both the P4–P6 domain and the intact intron occurs only in the presence of magnesium ions (8). Introns of the subclass to which the *Tetrahymena* intron belongs have a nonconserved extension comprised of three paired regions—P5a, P5b, and P5c—which fold first in the kinetic folding pathway of the intron (63).

Structure

The 2.8-Å-resolution crystal structure of the 160-nt P4–P6 domain (4) revealed that this RNA adopts a fold composed of two coaxial stacks (Figure 5). P4, P5, and P6 of the phylogenetically conserved core of the intron form an essentially continuous helical stack, as do the helices P5a and P5b of the variable extension. The third helical segment of the extension, P5c, forms a triple-helix junction with P5a and P5b. The two helical stacks are arranged side by side and are connected by two parallel strand crossovers at the top of the molecule, so that the fold resembles a candy cane (see color Figure 6). Two sets of tertiary interactions between loops in the P5abc extension and the minor groove of P4-P5-P6 clamp the two stacks together: An adenosine-rich corkscrew (the A-rich bulge) plugs into the minor groove of helix P4; and a GAAA tetraloop binds to a conserved 11-nt internal loop, termed the tetraloop receptor.

The backbone of the A-rich bulge makes a corkscrew turn held by two well-ordered magnesium ions. Two of its adenosine bases are flipped out and interact with adjacent residues of the P4 helix. A186 on the other side of the bulge fits into pockets in the three-helix junction (Figure 5). Site-directed mutagenesis and modification interference experiments have demonstrated that these interactions are crucial to the stability of the entire domain (36, 38).

The other half of the clamp involves the interaction of a GAAA tetraloop with a tetraloop receptor (color Figure 6). The tetraloop adopts the conformation seen previously in other molecular contexts (21, 44) and interacts with an 11-nucleotide (nt) motif with a widened minor groove that docks with the tetraloop in a highly specific manner (4). The tetraloop-receptor interaction incorporates a ribose zipper (4) and an adenosine platform (5), and its structure appears to be stabilized by binding of a potassium ion (1). The tetraloop receptor motif has been detected in many RNA sequences (10).

THE HEPATITIS DELTA VIRUS RIBOZYME

Background

The hepatitis delta virus is a human pathogen with a 1700-nucleotide, circular, single-stranded RNA genome. The viral genome is replicated by the host RNA

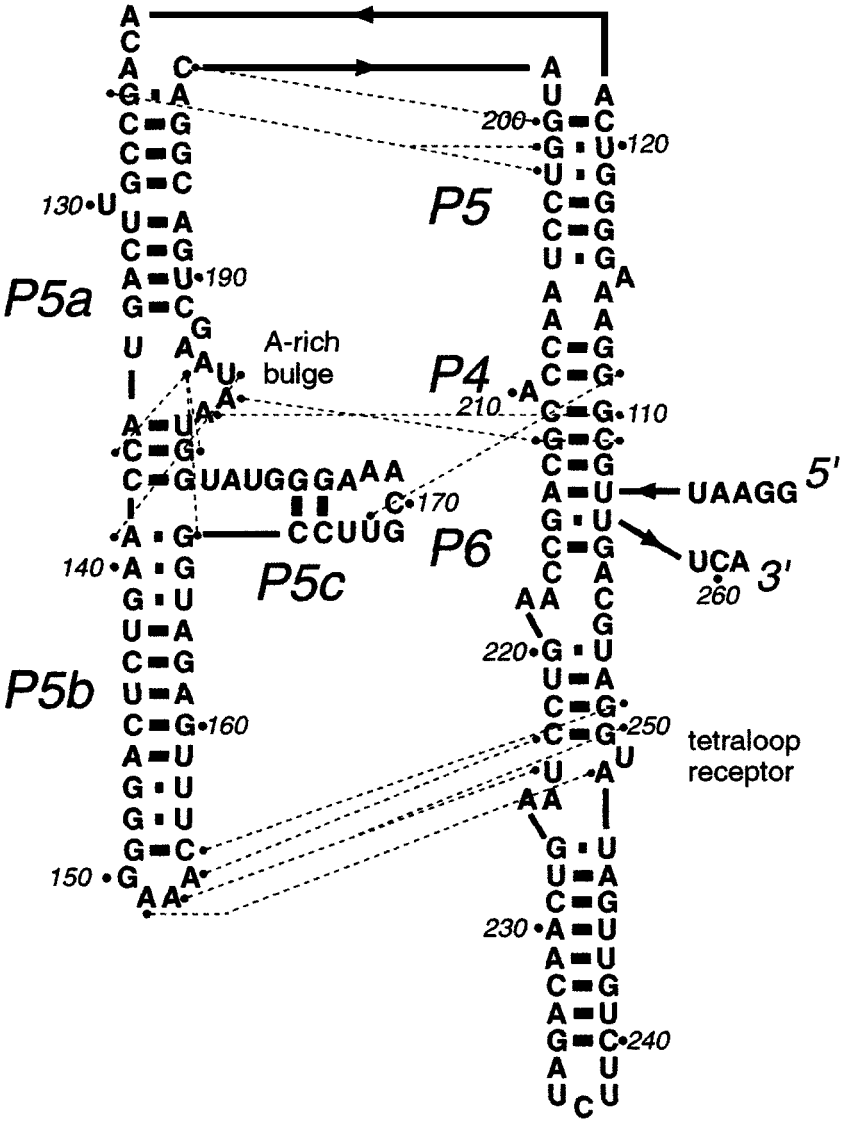

Figure 5 Secondary structure of the P4–P6 domain of the *Tetrahymena thermophila* group I self-splicing intron. Numbering corresponds to the intact intron.

polymerase II through a double rolling-circle mechanism, resulting in linear multimers of the genome and the complementary antigenome. The multimers are processed into unit-length RNAs by a self-cleaving activity, the HDV ribozyme, which is part of both RNAs. Eighty-five nucleotides of either strand can be taken out of their genomic or antigenomic context, and will catalyze a trans-esterification similar to that catalyzed by the hammerhead ribozyme. The HDV and hammerhead ribozymes, however, have distinctly different sequences and structures (2, 29).

No nucleotides 5′ of the cleavage site participate in the secondary structure of the HDV ribozyme (Figure 7). Consistent with this, the ribozyme will cleave a single nucleotide 5′ of the scissile bond. Furthermore, this (−1) nucleotide can have any purine or pyrimidine base. There is no requirement for particular sequences further upstream. Other remarkable features of the ribozyme

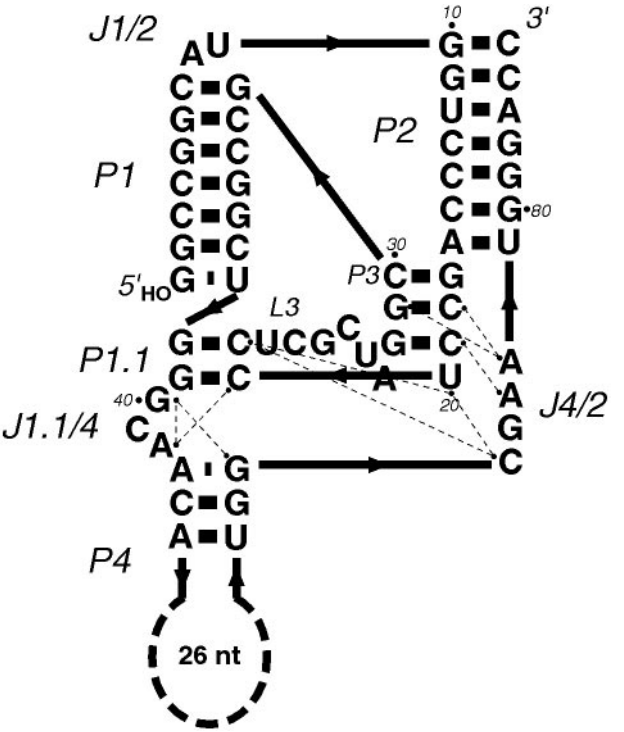

Figure 7 Secondary structure of a genomic HDV ribozyme. The P4 stem was altered for crystallization by introduction of a U1A protein binding site, without affecting HDV ribozyme activity (18a).

are its high reactivity (approximately 100 times faster than the hammerhead ribozyme), its activity in high concentrations of denaturants such as 8 M urea of 15 M formamide, and its nonspecific requirement for cations (2, 29).

Structure

We recently determined the crystal structure of a genomic HDV ribozyme at 2.3 Å resolution (18a). The HDV ribozyme folds into a compact structure comprised of five helical segments, P1, P1.1, P2, P3, and P4, connected as a nested double pseudoknot. The five helical segments form two parallel stacks: P1, P1.1, and P4 stack nearly coaxially, whereas P2 and P3 form a second coaxial stack (see color Figure 8). The two helical stacks are in turn joined side by side by five strand crossovers located at the joining regions J1/2 and J4/2, the connection between P3 and P1, and at either end of P1.1. The helix P1.1 is the key interaction that constrains the ribozyme into its unique three-dimensional fold by bracing together the P1 substrate-bearing helix, the P2-P3 stack, the P4 helix, and the functionally important J4/2 region. No tightly bound metal ions were found in the HDV ribozyme; the fold is stabilized entirely by base-pairing, stacking, and tertiary base-backbone and backbone-backbone interactions.

A pseudoknot is a nucleic acid structure characterized by base pairing between nucleotides in the loop of a conventional hairpin duplex with complementary residues outside that hairpin (43). The first pseudoknot in the HDV ribozyme was predicted by Perrotta & Been on the basis of sequence comparison and compensatory mutagenesis (41), and results from the presence of P2, in which nucleotides in the loop of the hairpin P1 base-pair with sequences distal to it. The two-base-pair helix P1.1 introduces the second pseudoknot into the structure. This pseudoknot, which characterizes the HDV ribozyme fold, is nested between the paired segments that constitute the first pseudoknot because its two constituent helices P1.1 and P3 stack with the helices (P1 and P2, respectively) of the first pseudoknot. This order of stacking is different from that of classical pseudoknots where two consecutive stems stack coaxially, and the two crossovers make triplex interactions in the major and minor grooves (28, 42).

Hydroxy radical footprinting of genomic and antigenomic ribozymes demonstrated strong protection of the RNA backbone on the 5′ side of P3 and L3, the nucleotides at the junction of P3, P1, and J4/2, and weaker protection of guanosines 38–40 (49). In good agreement, the crystal structure shows that these segments line a deep active site cleft that buries the 5′-hydroxyl leaving group of the self-scission reaction. The position of this functional group identifies the active site of the HDV ribozyme. This active site cleft is formed on three sides by the substrate-bearing P1 helix, J4/2, and a niche formed by the minor groove of P3 and L3. It is enclosed on top by the arching RNA backbone in the

junction of P3 and P1, and supported from below by the P1.1 helix. The HDV ribozyme fold is, in principle, defined by the stacking of helical segments and the placement of strand crossovers. However, a number of tertiary interactions appear to stabilize the ribozyme and also position key functional groups in its active site. One of these, the trefoil turn, might activate the key C75 nucleotide for catalysis (18a).

RNA FOLDS AND FOLDING

Size, Complexity, and Compactness

Comparison of the four folds reveals some common themes of RNA organization. The majority of nucleotides in all four folds are involved in base stacking interactions. This is not surprising given the strength of hydrophobic and aromatic-aromatic interactions in aqueous solutions. Although individual RNA helices are relatively short, the prevalence of stacking results in long, pseudo-continuous helices. Thus, noncanonical base pairs and nucleotide bulges are often accommodated within the stacks.

The four RNA folds known at near-atomic resolution are comprised each of just two coaxial stacks, the minimum to qualify as folds. While this at least in part reflects technical limitations, such as the difficulty in obtaining well-ordered crystals of very large RNAs, it might also indicate that two-stack folds are common. Because RNA folding must overcome considerable coulombic repulsion, even when a large fraction of the negatively charged phosphates are neutralized by nonspecific counter-ion condensation (31) as well as specific metal ion binding (30, 40), folds that can accomplish their biological function while minimizing the number of interfacial buried charges might have a selective advantage.

The tRNA, hammerhead, P4–P6, and HDV ribozyme two-stack folds encompass a broad range of size and complexity. The smallest and simplest is the hammerhead fold, which can be assembled with less than 40 nucleotides and consists of a two-helix stack and a third helix packed against the junction of the first two, forming a small, buried interface. The tRNA fold ($\sim$80 nt) is intermediate both in size and complexity, whereas the 160-nt P4–P6 fold is the largest currently known. Functional HDV ribozymes can have as few as 60 nucleotides. This ribozyme fold bears a superficial resemblance to the P4–P6 domain in that two helical stacks are arranged side by side. However, the stacks of P4–P6 are brought together by two sets of loop-to-helix-groove contacts involving noncanonical interactions (the A-rich-bulge-P4 and tetraloop-receptor contacts), whereas the smaller HDV ribozyme achieves its compact structure by adopting the complex double pseudoknot fold.

How compact are folded RNAs compared to standard A-form duplexes? A measure of compactness is the solvent-accessible surface area of an RNA

normalized by its number of nucleotides. Table 1 shows that after correction for helix-end exposure, canonical A-form duplexes present about 180 Å^2 of surface per nucleotide to solvent. The smaller RNA domains (hammerhead and HDV ribozymes, tRNA, and the two subdomains of P4–P6, P5abc and P4-P5-P6) range between 160 and 167 Å^2/nt, with the HDV ribozyme the most compact.

The P4–P6 domain appears to be considerably more compact than the smaller folds; it has an accessible surface area of only 148 Å^2/nt. This might be somewhat surprising, since its constituent subdomains have accessible areas of 161 and 165 Å^2/nt. What this difference reflects is the large surface area that the P5abc subdomain buries on the core P4-P5-P6 stack, thanks in part to the nearly perpendicular arrangement of the distal portion of the P5c helix relative to the P5a-P5b stack. We have classified P4–P6 as a two-stack fold in spite of this perpendicular helix, because the proximal portion of P5c participates in the P5a-P5b stack. Nonetheless, the accessibility calculation shows that, at least by the criterion of compactness, P4–P6 is qualitatively more complex than the smaller two-stack folds.

If the degree of compactness is roughly proportional to the number of helical stacks that are brought together, as suggested by the comparison of P4–P6 to the other three folds, we would expect even more tightly packed folds when larger RNA structures are determined. However, because RNA folds result from packing of double-stranded helices, a lower bound to the accessible area

Table 1 Comparison of solvent-accessible surface areas of selected ideal A-form duplex and folded RNAs

RNA	Accessible area (Å^2)[a]	Corrected area (Å^2)[b]	Nucleotides	Area per nucleotide (Å^2)
AC	1097	715	4	179
ACGC	1799	1417	8	177
ACGCAC	2508	2126	12	177
Hammerhead	7186	6804	41	166
HDV Rz[c]	9852	9470	59	160
tRNA^{Phe}	12686		76	167
p5abc[d]	12241	12050	75	161
P4-P5-P6[d]	13921	13730	83	165
P4-P6	23376		158	148

[a]Calculated with a 1.4-Å probe radius, excluding bound waters and cations.

[b]Corrected for exposed helical ends (191 Å^2 per end). Solvent-accessible areas of oligonucleotides and molecules taken out of their biological molecular context (the Hammerhead and HDV ribozymes, as well as the subdomains of P4–P6) were corrected.

[c]Nucleotides corresponding to the engineered U1A binding site were omitted from calculation.

[d]Atomic coordinates for the P4-P5-P6 subdomain were generated by deleting nucleotides 123 through 197 (the P5abc subdomain) from the P4–P6 structure. Structures of the isolated subdomains have not been determined experimentally.

per nucleotide must be set by the geometric limitations of helix—or, to a first approximation cylinder—close packing. Thus, even very tightly packed RNA folds are likely to be less dense than protein folds (47).

How Are RNA Folds Stabilized?

In a prescient review written before the structures of P4–P6 and the HDV ribozyme were determined, Draper proposed that three strategies could be used to stabilize RNA folds (13). First, hydrogen bonding networks might facilitate association of complementary RNA surfaces. Second, coordination of specific metal ions might mediate packing. Third, a convoluted fold such as a pseudoknot might be conformationally constrained by the fold itself. This last mode is not a tautology, because pseudoknots are defined by the pattern of Watson-Crick pairs, or secondary structure. The four RNA folds illustrate all modes of stabilization.

FOLD STABILIZATION BY COMPLEMENTARY SURFACES Tertiary interactions are responsible for association of complementary surfaces. Therefore, this mode of stabilization is present in all folds to some extent but is most graphically illustrated in the P4–P6 fold. The two helical stacks of this fold (the P5abc and P4-P5-P6 subdomains), which are covalently connected by a pair of crossovers only at one end and thus not conformationally restrained relative to each other, are brought into close parallel apposition by two discrete sets of loop-to-minor-groove contacts: the A-rich-bulge-P4 and tetraloop-receptor interactions. The importance of this type of fold stabilization is demonstrated dramatically by mutation of A186, which lies at the center of a dense network of hydrogen bonds in P5abc. Mutations at this position not only affect the structure locally, but result in unfolding of the entire domain (36).

Several sets of tertiary interactions, or motifs, have already been observed multiple times in RNA structures and are predicted to occur in other RNAs based on sequence analysis (33). The ribose zipper is present both in the P4–P6 domain and in the HDV ribozyme. Adenosine platforms occur in three different places in the P4–P6 domain and appear to mediate folding of the intact group I intron (5). The adenosine platform also participates in formation of the tetraloop receptor, which binds to a GAAA tetraloop, itself a motif. This last tertiary interaction motif is modular in the sense that it can be placed in different molecular contexts and still participate in the same type of association (18).

FOLD STABILIZATION BY CATION BINDING Fold stabilization by coordination of metal ions would involve, in the strict sense, association of two helical stacks mediated purely by cations. Such an RNA structure has not yet been observed. Nevertheless, at least in one case, crystallographically observed metal ions have been demonstrated to mediate RNA folding indirectly by stabilizing RNA

conformations, which in turn participate in tertiary interactions. Five magnesium ions were observed crystallographically to bind to the three-way P5abc junction of the P4–P6 domain, and single-atom changes in any one of four experimentally accessible cation binding sites abolished folding of the junction, and as a consequence, the entire 160-nt P4–P6 domain (6).

A monovalent ion binding site was demonstrated underneath adenosine platforms by both X-ray crystallography and nucleotide analog interference mapping. A single-atom substitution of an oxygen ligand with sulfur changed the specificity of a site within the P4–P6 domain from the hard cation K^+ to the soft metal ion Tl^+. A similar effect was observed in a group I intron from *Azoarcus*. Furthermore, activity assays on this intron demonstrated that K^+ binding is important for splicing activity, and that this cation cannot efficiently be replaced by Na^+, Li^+, or Cs^+ (1).

Starting with classical experiments on tRNA (56, 57), cation involvement in RNA folding has been extensively documented. Metal ions have also been observed to participate intimately in the formation of local RNA structures, such as guanine quadruplexes (26), cation mediated loops (23, 25), and distorted helical grooves (9). However, their detailed manner of participation in the stabilization of RNA folds mostly remains to be established and constitutes an exciting area of inquiry.

FOLD STABILIZATION BY CHAIN CONNECTIVITY The HDV ribozyme illustrates stabilization of an RNA fold by the order of connection of helical segments. This doubly pseudoknotted ribozyme is of similar size to the simply connected hammerhead, yet achieves a compact structure in which two helical stacks are tightly packed side by side to produce a deep catalytic cleft. This structure is so stable that the ribozyme is active in high concentrations of denaturants. Other naturally occurring ribozymes, such as the hammerhead, are strongly inactivated by denaturants and also display catalytic rates which are about 100 times slower than that of the HDV ribozyme.

Two lines of evidence support the notion that the connectivity of the main chain stabilizes the HDV ribozyme fold. First, deletion or mutagenesis of the nucleotides that form P1.1 to nucleotides that would not form Watson-Crick pairs reduces ribozyme activity by three to five orders of magnitude. P1.1 is the helix that establishes the second of the two nested pseudoknots of the HDV fold. Disruption of this helix would delete two out of the five crossovers present in the HDV fold, converting it into a single pseudoknot. Second, disruption of the other pseudoknot by cleaving the ribozyme at the J1/2 crossover at the top of the molecule (Figure 7) produces molecules that are active in *trans*. However, these *trans*-ribozymes have cleavage rates that are about two orders of magnitude slower than those of the parent *cis*-ribozymes. These slower rates

are comparable to those of the hammerhead ribozyme. Furthermore, the *trans*-ribozymes are inactive in denaturants, showing that the pseudoknot involving J1/2 provides an important part of the stabilization of the fold (2).

The HDV ribozyme fold may occur in other RNAs. For instance, a nested double pseudoknot structure was proposed for the translational repressor of the α operon of *Escherichia coli* (59). Intriguingly, it was found that the thermal unfolding of the two inner helices of the α repressor (corresponding to P2 and P3 in the HDV ribozyme) is coupled, which suggests that they might be stacked as seen in the HDV ribozyme (19). Another double pseudoknot fold has been proposed for a very active RNA ligase identified by in vitro selection (14). Because of the abundance of base-pairing interactions between segments that are distant in sequence, observed in large RNAs such as ribosomal RNA and RNase P, we expect fold stabilization by intricate main-chain connectivity to be a strategy often adopted by RNA.

Stability and Function

When rationalizing the importance of RNA folds in terms of structure-function relationships, it should be kept in mind that RNA duplexes are very stiff, at least in solutions of physiological ionic strength (20). When the structure of $tRNA^{Phe}$ was determined, it was suggested that placement of the anticodon and acceptor moieties on the far ends of the molecule was important. Images of tRNA molecules in the active site of ribosomes suggest that the fold, as well as the rigidity of the molecule, is important for bridging between the decoding and peptidyl-transferase sites (55).

Although the structure of the hammerhead in the active state remains to be determined, it is possible that the molecule twists somewhere along its Y junction to reach the transition state. If this is the case, the helical stacks might move as rigid bodies, with most of the flexing taking place in the interhelical interface made by tertiary contacts. This is supported by the observation that the angle between the two helical stacks varies between molecules that are in different crystal environments (61). Alternatively, movement might be confined to the immediate vicinity of the scissile bond (37, 53).

In the case of P4–P6, folding kinetics studies have shown that the P5abc subdomain is the first segment of the entire intron to fold, followed by assembly of the P4–P6 domain (50, 63). The structure of the fold can then be taken to indicate that the P5abc subdomain associates with a less stable P4-P5-P6 helix, and facilitates its helical stacking by lining one face of the helix. This would set the stage for assembly of the complete intron active site, of which the P4-P5-P6 constitutes an important part.

The HDV ribozyme fold is perhaps the most suggestive of the function of the RNA because of the resemblance of this RNA fold to that of many protein

enzymes. The presence of a deep catalytic site cleft between the two helical stacks of the fold where the 5′-hydroxyl leaving group is snugly placed between potentially reactive backbone and base functional groups immediately suggests an analogy with protein enzyme active sites, where substrates are bound, positioned, and activated for catalysis.

CONCLUSION

As discussed above, some modes of stabilization of RNA folds might result in weakly formed structures. This suggests that whereas many RNAs could in principle have intramolecular or interstack interfaces that are conducive to folding, the free energy of folding might be small enough to require the binding of a protein to drive the equilibrium in the folded direction (1a). If this is true, it would be expected that many RNA folds within ribonucleoproteins (RNPs) would be similar to those of RNAs that fold spontaneously. Such a role for proteins is also consistent with most biologically important RNAs in present-day cells occurring as RNPs (e.g. ribosomes, spliceosomes, signal recognition particle, mobile group II introns, etc). In the case of group I introns, two studies imply that proteins can play precisely such a role in stabilizing RNA folds that, in other contexts, are stable by themselves (32, 62).

How many RNA folds exist? If there is a finite, relatively small number of folds, is this a consequence of the intrinsic chemical stability of the folds, or a reflection of the number of RNA folds present at the origin of life? The discovery that the GNRA tetraloop motif, common in biological RNAs, forms part of the substrate binding pocket in an artificially selected molecule (11, 24), argues that some RNA motifs are prevalent due to their intrinsic chemical stability. Only the determination of more structures will tell whether this holds for RNA folds as well.

Literature Cited

1. Basu S, Rambo RP, Strauss-Soukup J, Cate JH, Ferré-D'Amaré AR, et al. 1998. A specific monovalent metal ion integral to the A-A platform of the RNA tetraloop receptor. *Nat. Struct. Biol.* 5(11):968–92

1a. Batey RT, Williamson JR. 1998. Effects of polyvalent cations on the folding of an rRNA three-way junction and binding of ribosomal protein S15. *RNA* 4:984–97

2. Been MD, Wickham GS. 1997. Self-cleaving ribozymes of hepatitis delta virus RNA. *Eur. J. Biochem.* 247:741–53

3. Carson M. 1991. Ribbons 2.0. *J. Appl. Crystallogr.* 24:958–81

4. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, et al. 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* 273:1678–85

5. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, et al. 1996. RNA tertiary

structure mediation by adenosine platforms. *Science* 273:1696–99
6. Cate JH, Hanna RL, Doudna JA. 1997. A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.* 4:553–58
7. Cech TR. 1990. Self-splicing of group I introns. *Annu. Rev. Biochem.* 59:543–68
8. Celander DW, Cech TR. 1991. Visualizing the higher order folding of a catalytic RNA molecule. *Science* 251:401–7
9. Correll CC, Freeborn B, Moore PB, Steitz TA. 1997. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell* 91:705–12
10. Costa M, Michel F. 1995. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* 14:1276–85
11. Dieckmann T, Suzuki E, Nakamura GK, Feigon J. 1996. Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* 2:628–40
12. Doudna JA, Cech TR. 1995. Self-assembly of a Group I intron active site from its component tertiary structural domains. *RNA* 1:36–45
13. Draper DE. 1996. Strategies for RNA folding. *Trends Biochem. Sci.* 21:145–49
14. Ekland EH, Szostak JW, Bartel DP. 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* 269:364–70
15. Ellington AD. 1994. RNA selection. Aptamers achieve the desired recognition. *Curr. Biol.* 4:427–29
16. Feig AL, Scott WG, Uhlenbeck OC. 1998. Inhibition of the hammerhead ribozyme cleavage reaction by site-specific binding of Tb. *Science* 279:81–84
17. Feigon J, Dieckman T, Smith FW. 1996. Aptamer structures from A to ζ. *Chem. Biol.* 3:611–17
18. Ferré-D'Amaré AR, Zhou K, Doudna JA. 1998. A general module for RNA crystallization. *J. Mol. Biol.* 279:621–31

18a. Ferré-D'Amaré AR, Zhou K, Doudna JA. 1998. Crystal of a hepatitis delta virus ribozyme. *Nature* 395:567–74

19. Gluick TC, Draper DE. 1994. Thermodynamics of folding a pseudoknotted mRNA fragment. *J. Mol. Biol.* 241:246–62
20. Hagerman PJ. 1997. Flexibility of RNA. *Annu. Rev. Biophys. Biomol. Struct.* 26: 139–56
21. Heus HA, Pardi A. 1991. Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science* 253:191–94
22. Holbrook SR, Sussman JL, Warrant RW, Kim SH. 1978. Crystal structure of yeast phenylalanine transfer RNA II. Structural features and functional implications. *J. Mol. Biol.* 123:631–60
23. Ippolito JA, Steitz TA. 1998. A 1.3 Å resolution crystal structure of the HIV-1 transactivation response region RNA stem reveals a metal ion-dependent bulge conformation. *Proc. Natl. Acad. Sci. USA.* 95: 9819–24
24. Jiang F, Kumar RA, Jones RA, Patel DJ. 1996. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature* 382:183–86
25. Kieft JS, Tinoco I. 1997. Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt (III) hexammine. *Structure* 5:713–21
26. Kim J, Cheong C, Moore PB. 1991. Tetramerization of an RNA oligonucleotide containing a GGGG sequence. *Nature* 351:331–32
27. Kim SH, Suddath FL, Quigley GJ, McPherson A, Sussman JL, et al. 1974. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* 185:435–40
28. Kolk MH, van der Graaf M, Wijmenga SS, Pleij CWA, Heus HA, Hilbers CW. 1998. NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. *Science* 280:434–38
29. Lai MM. 1995. The molecular biology of hepatitis delta virus. *Annu. Rev. Biochem.* 64:259–86
30. Laing LG, Gluick TC, Draper DE. 1994. Stabilization of RNA structure by Mg ions. *J. Mol. Biol.* 237:577–87
31. Manning GS. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11: 179–246
32. Mohr G, Caprara MG, Guo Q, Lambowitz AM. 1994. A tyrosyl-tRNA synthetase can function similarly to an RNA structure in the *Tetrahymena* ribozyme. *Nature* 370:147–50
33. Moore PB. 1999. Structural motifs in RNA. *Annu. Rev. Biochem.* In press
34. Munishkin A, Wool IG. 1997. The ribosome-in-pieces: binding of elongation factor EF-G to oligoribonucleotides that mimic the sarcin/ricin and thiostrepton domains of 23S ribosomal RNA. *Proc. Natl. Acad. Sci. USA* 94:12280–84
35. Murphy FL, Cech TR. 1993. An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry* 32:5291–300
36. Murphy FL, Cech TR. 1994. GAAA tetraloop and conserved bulge stabilize

tertiary structure of a group I intron domain. *J. Mol. Biol.* 1994:49–63
37. Murray JB, Terwey DP, Maloney L, Karpeisky A, Usman N, et al. 1998. The structural basis of hammerhead ribozyme self-cleavage. *Cell* 92:665–73
38. Ortoleva-Donnelly L, Szewczak AA, Gutell RR, Strobel SA. 1998. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA* 4:498–519
39. Oubridge C, Ito N, Evans PR, Teo C-H, Nagai K. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* 372:432–38
40. Pan T, Long DM, Uhlenbeck OC. 1993. Divalent metal ions in RNA folding and catalysis. In *The RNA World*, ed. RF Gesteland, JF Atkins, pp. 271–302. Cold Spring Harbor, New York: Cold Spring Harbor Lab. Press
41. Perrotta AT, Been MD. 1991. A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature* 350:434–36
42. Pleij CWA. 1994. RNA pseudoknots. *Curr. Opin. Struct. Biol.* 4:337–44
43. Pleij CWA, Rietveld K, Bosch L. 1985. A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res.* 13:1717–31
44. Pley HW, Flaherty KM, McKay DB. 1994. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature* 372:111–13
45. Pley HW, Flaherty KM, McKay DB. 1994. Three-dimensional structure of a hammerhead ribozyme. *Nature* 372:68–74
46. Puglisi JD, Chen L, Blanchard S, Frankel AD. 1995. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* 270: 1200–3
47. Richards FM. 1974. The interpretation of protein structures: total volume, group volume distributions and packing density. *J. Mol. Biol.* 82:1–14
48. Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, et al. 1974. Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* 250:546–51
49. Rosenstein SP, Been MC. 1996. Hepatitis delta virus ribozymes fold to generate a solvent-inaccessible core with essential nucleotides near the cleavage site phosphate. *Biochemistry* 35:11403–13
50. Sclavi B, Woodson S, Sullivan M, Chance MR, Brenowitz M. 1997. Time-resolved synchrotron X-ray "footprinting," a new approach to the study of nucleic acid structure and function: application to protein-DNA interactions and RNA folding. *J. Mol. Biol.* 266:144–59
51. Scott WG, Finch JT, Klug A. 1995. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell* 81:991–1002
52. Scott WG, Klug A. 1996. Ribozymes: structure and mechanism in RNA catalysis. *Trends Biochem. Sci.* 21:220–24
53. Scott WG, Murray JB, Arnold JRP, Stoddard BL, Klug A. 1996. Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science* 274:2065–69
54. Söll D. 1993. Transfer RNA: an RNA for all seasons. In *The RNA World*, ed. RF Gesteland, JF Atkins, pp. 157–84. Cold Spring Harbor, New York: Cold Spring Harbor Lab. Press
55. Stark H, Orlova EV, Rinke-Appel J, Junke N, Mueller F, et al. 1997. Arrangement of tRNAs in pre- and posttranslational ribosomes revealed by electron cryomicroscopy. *Cell* 88:19–28
56. Stein A, Crothers DM. 1976. Conformational changes of transfer RNA. The role of magnesium (II). *Biochemistry* 15:160–68
57. Stein A, Crothers DM. 1976. Equilibrium binding of magnesium(II) by *Escherichia coli* tRNAfMet. *Biochemistry* 15:157–60
58. Symons RH. 1997. Plant pathogenic RNAs and RNA catalysis. *Nucleic Acids Res.* 25:2683–89
59. Tang CK, Draper DE. 1989. Unusual mRNA pseudoknot structure is recognized by a protein translational repressor. *Cell* 57:531–36
60. Turner DH, Bevilacqua PC. 1993. Thermodynamic considerations for evolution of RNA. In *The RNA World*, ed. RF Gesteland, JF Atkins, pp. 447–64. Cold Spring Harbor, New York: Cold Spring Harbor Lab. Press
61. Wedekind JE, McKay DB. 1998. Crystallographic structures of the hammerhead ribozyme: relationship to folding and catalysis. *Annu. Rev. Biophys. Biomol. Struct.* 27:475–502
62. Weeks KM, Cech TR. 1996. Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science* 271:345–48
63. Zarrinkar PP, Williamson JR. 1994. Kinetic intermediates in RNA folding. *Science* 265:918–24

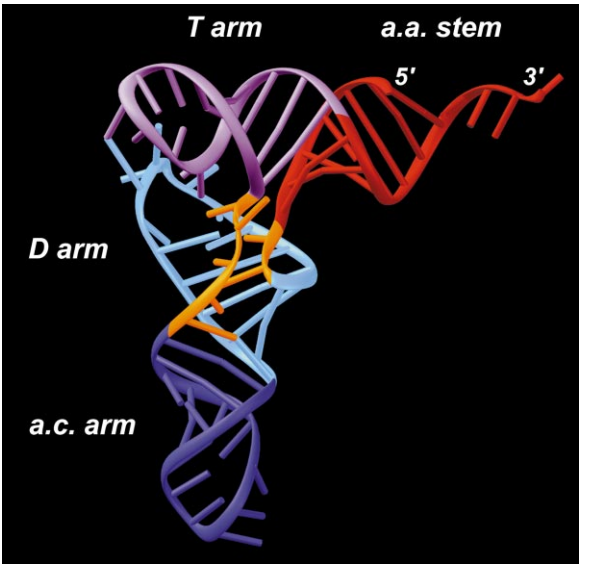

Figure 2 Ribbon and stick representation of the three-dimensional structure of $tRNA^{Phe}$. Crossovers are colored *gold*. Individual helical segments (named as in Figure 1) are in different colors. Color figures prepared with RIBBONS (3).

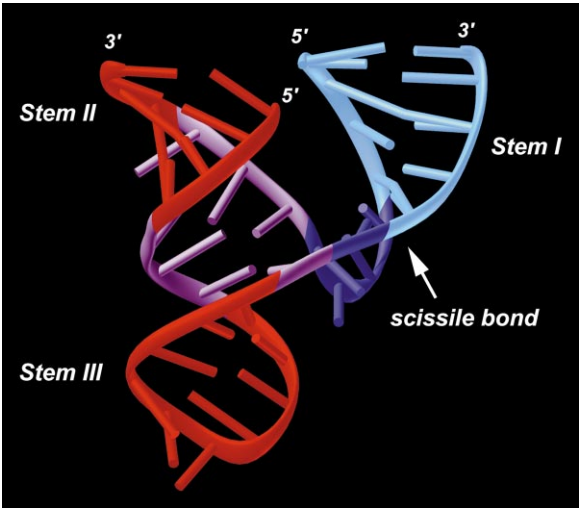

Figure 4 Ribbon and stick representation of the three-dimensional structure of the hammerhead ribozyme RNA of Figure 3. Nucleotides that form the U-turn are in *dark blue*; those that form the noncanonical helix that stacks between stems II and III are in *magenta*.

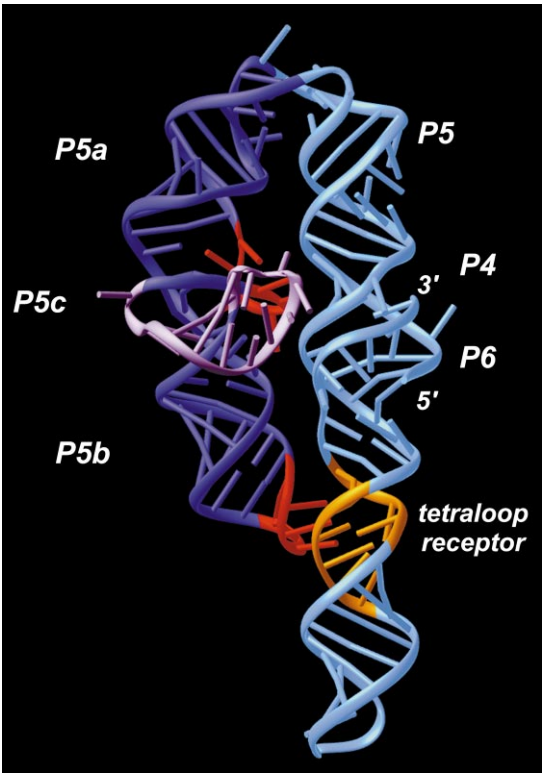

Figure 6 Ribbon and stick representation of the three-dimensional structure of the P4–P6 domain of the Tetrahymena group I intron. The P5abc and P4–P5–P6 subdomains are in *dark* and *light blue*, respectively. The P5c stem is in *magenta*, the tetraloop receptor in *gold*, and the GAAA tetraloop and A-rich bulge in *red*.

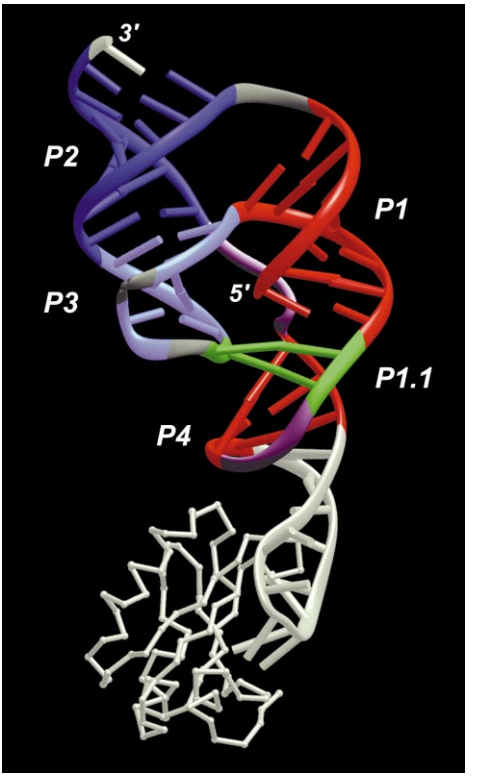

Figure 8 Ribbon and stick representation of the HDV ribozyme. The U1A protein bound for crystallization is in *white* at the bottom of the figure. The central P1.1 helix is in *green*. Sticks representing bases are omitted in single-stranded segments (*gray* and *magenta*) for clarity.