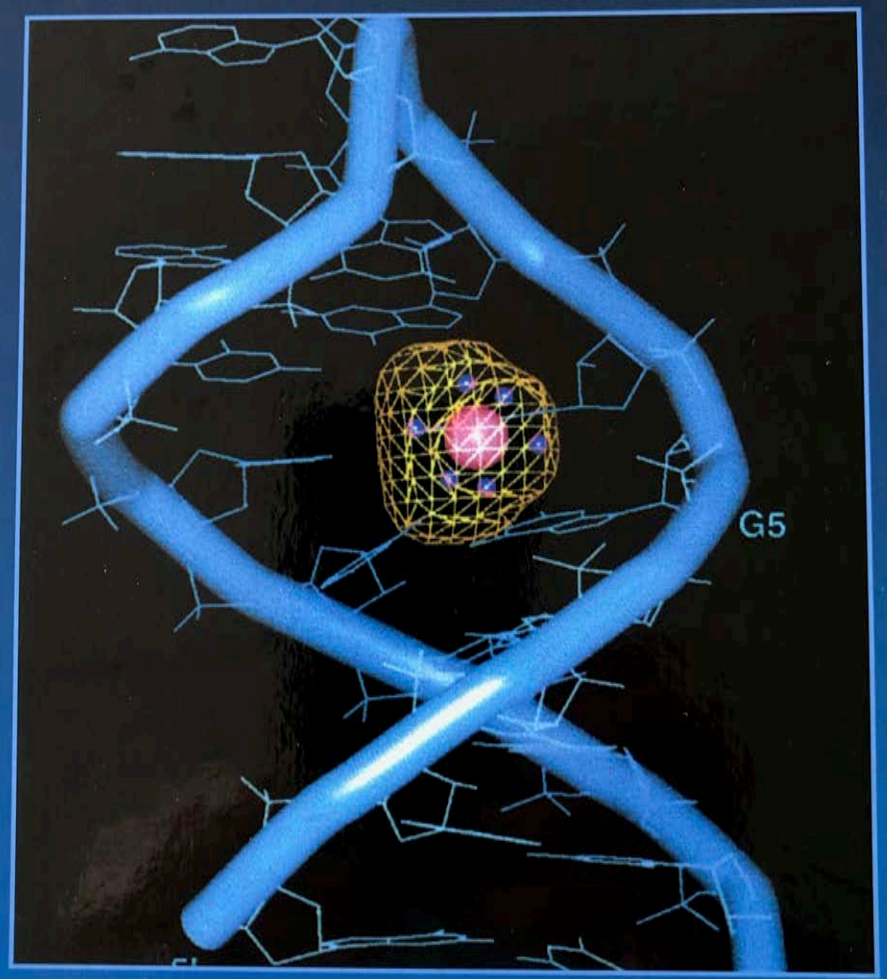
BioTechniques BOOKS

RIBOZYME

Biochemistry and Biotechnology



Edited by: Guido Krupp & Rajesh K. Gaur

CHAPTER 16

Recent Advances in RNA Crystallography

Martin Egli and George Minasov
Department of Molecular Pharmacology and Biological Chemistry and The Drug Discovery Program, Northwestern University Medical School, Chicago, IL, USA

Background: Until recently, the explosive growth of research on RNA that followed the discovery of its catalytic properties and the accumulation of a wealth of functional data were not accompanied by similar progress in the X-ray crystallographic determination of RNA structure. Chemical probing, various spectroscopic techniques, solution NMR, and modeling were the main sources of information regarding the three-dimensional structure of RNA beyond simple double helical fragments that had either canonical geometry or contained non-standard base pairs. Twenty years elapsed between the report of the first single crystal structure of an RNA, the phenylalanine transfer RNA, and the second one of an RNA molecule involving relatively complex tertiary structures, the hammerhead ribozyme. Fortunately, this situation has now changed, and the last five years have brought a steady flow of exciting new RNA crystal structures.

Results: RNA has become the focus of an increasing number of structural biology studies. Improvements in the chemical and enzymatic synthesis of RNA, sparse matrix crystallization screens tailored to RNA, and crystal engineering have paved the way to RNA crystals that are suitable for structure determination at high resolution. Recently reported crystal structures of RNA molecules that have complex tertiary interactions include the hammerhead, hepatitis delta virus (HDV), *Tetrahymena thermophila* group I intron ribozymes, the P4–P6 domain of that intron, a lead-dependent ribozyme, a 29-nucleotide fragment of rat 28S ribosomal RNA containing the sarcin/ricin loop (SRL), the HIV-1 *trans*-activation response (TAR) domain, a 62-nucleotide domain of *Escherichia coli* 5S rRNA, the RNA pseudoknot from beet western

Ribozyme Biochemistry and Biotechnology
Edited by Guido Krupp and Rajesh K. Gaur
© 2000, Eaton Publishing, Natick, MA

yellow virus (BWYV) that is involved in ribosomal frameshifting, and several aptamers.

Conclusions: The four crystal structures of ribozymes allow a refined interpretation of the available functional data and shed light on the structural bases of RNA catalysis. The hammerhead ribozyme displays an active site that is almost completely exposed to solvent. The *in-line* mechanism associated with the phosphoryl transfer reaction that is catalyzed by the hammerhead requires a local conformational change away from the ground state structure observed in the crystal. Conversely, the catalytic centers in the HDV ribozyme and the group I intron are sequestered from solvent and resemble the active sites of protein enzymes. Both catalysts appear to be conformationally preorganized to accommodate their substrates. Despite the still small number of RNA crystal structures, many novel and recurring tertiary structural motifs have emerged. These include the tetraloops, the tetraloop receptor, the U-turn, purine-rich bulges, A-platforms, cross-strand purine stacks, the minor groove triplex, and others. In addition, several structures demonstrate the important role that metal ions play in RNA folding and the stabilization of particular conformations. With the three-dimensional–structural database of RNA now rapidly expanding, we have definitely entered the post-tRNA (structure) era.

Techniques: RNA suitable for X-ray crystallographic studies can be prepared either by chemical synthesis or in vitro transcription. Enzymatic synthesis is the method of choice for larger RNAs, typically 50 mers and longer. Although still relatively expensive, improved chemical synthetic methods may allow production of RNAs as long as 100 nucleotides in reasonable yields. After purification through the use of preparative gel electrophoresis or column chromatography and desalting, aqueous stock solutions of RNA are stored at -20°C under metal-ion–free conditions. A battery of sparse matrix screens then allow rapid sampling of a large number of crystallization conditions. Major advances were made in the X-ray crystallographic analysis of macromolecular structures over the last years. The flash-freezing technique helps preserve the crystal for the duration of data collection. Synchrotron radiation has led to drastic improvements in data resolution and makes data collection using very small crystals possible. Tunability of the energy of the X-ray beam enables collection of anomalous diffraction data at multiple wavelengths. Provided that the crystal contains anomalous scatterers (e.g., bromine atoms from nucleotide analogs that are incorporated into an RNA via chemical synthesis), this so-called multiwavelength anomalous diffraction (MAD) technique can greatly facilitate the structure determination and model building processes.

1. INTRODUCTION

For two decades the crystal structures of a handful of transfer RNAs, either alone (52,65,84) or in some cases in complex with the corresponding aminoacyl synthetases (11,85,87), were the only sources of information concerning the detailed tertiary interactions underlying the overall fold of a complex RNA molecule (95). In the mid-seventies, RNA's principal roles were considered to be those of the messenger in biological information transfer (mRNA), adaptor molecule in protein synthesis (tRNA), and major constituent of the ribosome (rRNA). In addition, there was an appreciation for the multitude of interactions between proteins and RNA required in a properly functioning cell (83).

We now know that RNA is considerably more versatile and plays a much more active role in many cellular and recognition processes than was previously anticipated. For one, several RNAs were found to catalyze phosphoryl transfer reactions in a highly sequence-specific fashion (3,18,100). Many new, specific interactions between proteins and RNA have been reported, with RNA being a more active partner in the recognition process as compared with DNA (38,70). In addition, the in vitro selection technique, SELEX (5,40,49,102), has led to the identification of a flurry of RNA molecules that serve as aptamers (48,89) or catalyze chemical reactions that are normally performed by protein enzymes [e.g., site-specific RNA cleavage (75), phosphorylation (58), RNA ligation (27), amino acylation (57), and ribonucleoside polymerization (28)]. Indeed, there is now good evidence that, in the ribosome, it is the RNA portion that is responsible for catalyzing polypeptide synthesis (71,73).

The expanded list of functions rendered the need for detailed three-dimensional structures of RNA more urgent. A battery of chemical probing techniques, molecular modeling (55,61), and solution NMR (2,59,80,81) were the main sources of structural information. However, no crystallographic information was available beyond the transfer RNAs, and while it was possible to generate sufficient amounts of pure RNA for crystallization experiments using in vitro transcription (63,112), growing well-diffracting RNA crystals appeared to be more difficult. The introduction of sparse matrix crystallization screens that were specially suited for growing crystals of RNA (23,91) and improved strategies for structure determination (12,13) have helped to dramatically change this situation. A dozen crystal structures of RNAs, among them four catalytic species, have been reported over the last five years (Table 1). Moreover, a large number of RNA structures were determined by solution NMR and, similarly, the database of RNA-protein complex structures, analyzed by both NMR and X-ray crystallography, is growing steadily. However, in this review we will focus on recent crystal structures of RNA alone and summarize some of the structural and functional insights that were gained from them.

2. RESULTS AND DISCUSSION

In the following sections we will briefly describe each of the recently determined RNA crystal structures and discuss biological implications and remaining questions. However, we have kept at a minimum the description of the accumulated data for the RNAs prior to their crystal structure determinations and the reader is referred to the cited primary sources and reviews for more background information. An overview of the observed tertiary structural motifs is given in Table 2.

Table 1. Recent RNA Crystal Structures

RNA (number of residues)	Organism	Source	Resolution [Å]	Reference / NDB entry code
Hammerhead ribozyme 47 (RNA-DNA)	plant virus	chem. synthesis	2.6	(78,79) UHX026
Hammerhead ribozyme 41 (all-RNA)	plant virus	chem. synthesis	3.1	(69,92) URX057, URX071
Group I intron 160 (P4–P6 domain)	*Tetrahymena thermophila*	T7 transcription	2.8	(14) URX053
Group I intron 247 (P4–P6/P3–P9 domains)	*Tetrahymena thermophila*	T7 transcription	5.0	(41) UR0003
HDV Ribozyme 72 (P4–P6 domain)	Hepatitis Delta Virus	T7 transcription	2.3	(33) URX053
Leadzyme 24 (all-RNA)	in vitro selection (transfer RNA)	chem. synthesis	2.7	(108) 1DRZ (PDB)
5S rRNA 62 (helices I, IV and loop E)	*E. coli*	T7 transcription	3.0	(20) URL065, URL064
TAR RNA 27 (trans-activ. resp. region)	HIV-1	chem. synthesis	1.3	(46) URX075
28S rRNA (SRL) 29 (sarcin/ricin loop)	rat	chem. synthesis	2.1	(21) 430D (PDB)
RNA frameshift pseudoknot 32	beet western yellow virus	T7 transcription	1.6	(97) 437D (PDB)
RNA aptamers (F5 and F7) 18 / 14 (all-RNA)	in vitro selection	chem. synthesis	2.8	(86) 5MSF, 7MSF (PDB)
Deoxyribozyme 82 (10–23 enzyme)	in vitro selection	chem. synthesis	2.6	(74) 1BR3 (PDB)

2.1 Hammerhead ribozymes

The hammerhead ribozyme is derived from a conserved sequence motif that is responsible for self-cleavage in satellite RNAs of certain viruses (103). The hammerhead motif consists of three base-paired stems that flank a non-helical core of mostly conserved residues. Cleavage occurs 3′-adjacent to a dimer UH (H = C, U, and A, but not G, Figure 1), proceeds via inversion of the phosphate group, and requires the presence of divalent metal cations (60). Several crystal structures of hammerhead ribozymes were reported over the last years, starting with the one composed of an RNA enzyme and a DNA substrate strand (78).

This structure established the overall wishbone-like shape of the hammerhead ribozyme. Accordingly, two of the three A-form helices (stems II and III, Figure 1) stack upon one another through a series of mostly-conserved non-canonical base pairs. Stem I is linked to them via the cleavage-site nucleotide C17 and the catalytic pocket, which assumes a U-turn conformation. The cleavage-inactive RNA-DNA construct furnished only limited information with respect to the conformational basis of the cleavage specificity. The observed conformation of the substrate strand at the cleavage site (a hydroxyl group was modeled onto the 2′-deoxyribose of residue C17) was not consistent with an 'in-line' attack that is a stretched angle between the 2′-oxygen and the scissile P-O5′ bond. This arrangement is required for the reaction to proceed via inversion of the attacked phosphate group. Moreover, no metal ions were found near the active site. However,

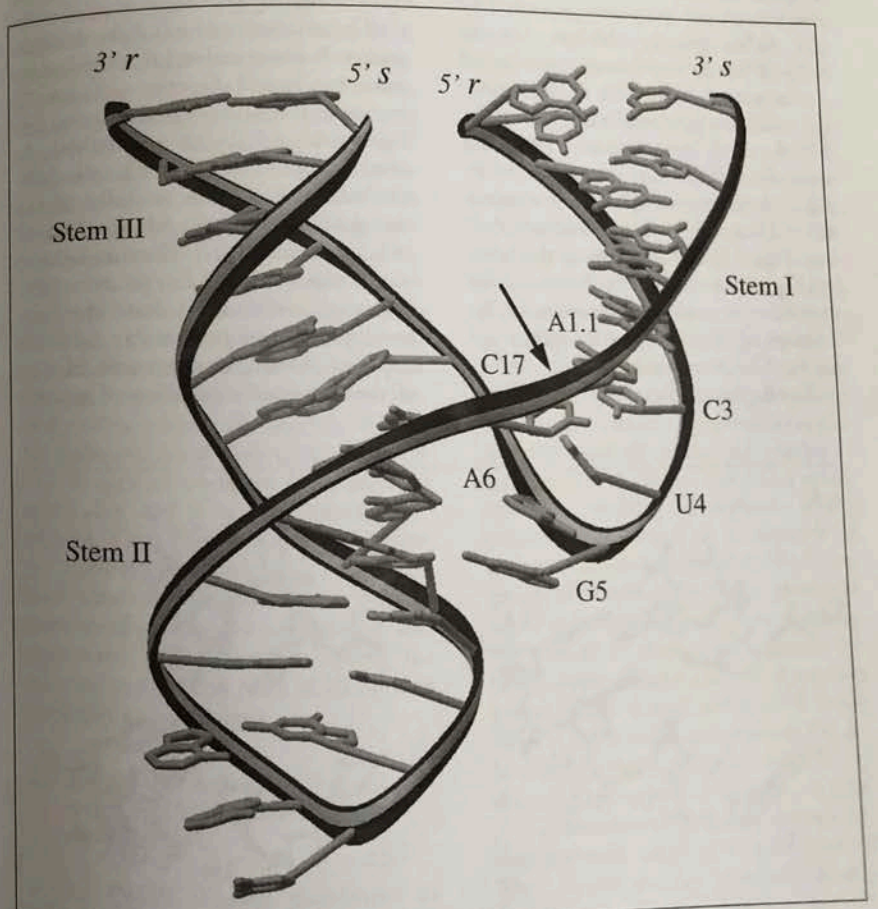


Figure 1. Overall view of the hammerhead ribozyme (nucleic acid database entry NDB URX057) (92). Ribbons trace the paths of the ribose-phosphate backbones and terminal residues of both strands are labeled (r = ribozyme, s = substrate). The arrow marks the cleavage site between residues C17 and A1.1, and the nucleotides C3, U4, G5, and A6 forming the uridine turn are labeled.

the structure offered some suggestions as to the role of conserved core residues for ion coordination and possible changes in the conformation of the substrate strand that would bring about an in-line conformation. In addition, the structure revealed that the active site in the hammerhead ribozyme is exposed to solvent.

The crystal structure of an all-RNA hammerhead–substrate complex that has a 2′-*O*-methyl nucleotide at position 17 that prevents cleavage in the presence of magnesium ions confirmed the overall conformational properties observed in the earlier enzyme-inhibitor complex (92). While this structure, too, showed an "adjacent" orientation of the 2′-oxygen relative to the scissile bond, it established three magnesium-ion–binding sites. One of the ions is coordinated to the exocyclic amino groups of residues C3, C17, and A1.1 via its hydration shell (see Figure 1 for orientation, the latter residue is not conserved). Based on this coordination mode, a mechanism for the line-up of the 2′-hydroxyl group of residue C17 was postulated (92).

Two further studies that used cleavage-active hammerhead ribozymes in the crystal instead of enzyme-inhibitor complexes demonstrated conformational changes at the active site toward a more in-line orientation but fell short of settling open questions regarding the precise mechanism of hammerhead-mediated cleavage. The first study used time-resolved X-ray crystallographic techniques and freeze-trapping to probe metal ion coordination and active-site conformational flexibility (93). One of the trapped structures displayed a major shift of the sugar moiety of nucleotide C17 relative to residues of the catalytic pocket. It also revealed a further divalent metal ion bound directly to the pro-R_P oxygen of the A1.1 phosphate group (see Figures 1 and 2 for orientation). In another experiment, crystals of a hammerhead ribozyme in complex with a chemically modified substrate strand (A1.1 features a (*S*)-5′-C-methyl-ribose) were soaked in metal ion solutions (69). The modification slows down the cleavage reaction and presumably facilitates trapping of reaction intermediates. One of these showed a rotation of approxi-

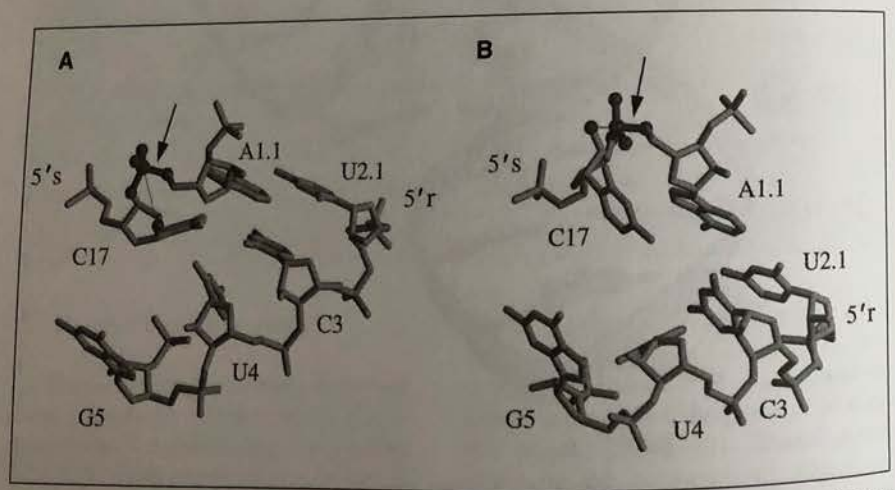


Figure 2. Close-up views. (A) the "ground state" (or *adjacent*) conformation (NDB entry URX075) (92) and (B) an "intermediate" (closer to *in-line*) conformation (NDB entry URX071) (69) of the hammerhead ribozyme cleavage site. The attacking O2′ of the active site cytosine 17 and the phosphate group of residue A1.1 are highlighted with spheres, and a thin line connects O2′ and P. The scissile bond is marked by an arrow and residues are labeled.

mately 60° by the ribose and base portions of residue C17 relative to the control structure, resulting in a movement of around 9 Å by the base moiety. Figure 2 depicts a comparison of the active site conformations in the structures of hammerhead ribozymes with 2′-*O*-methyl- (92) and (*S*)-5′-C-methyl-modifications (69) of residues C17 and A1.1, respectively, in their substrate strands.

While all crystallographic studies of hammerhead ribozymes revealed a more or less A-form helical conformation at the active site region, the latter structure showed a significantly altered conformation that was closer to an in-line state. However, it remains unclear as to what extent the conformational change was caused by the chemical modification. There is some evidence, from a crystallographic study of an RNA fragment with an incorporated (*S*)-5′-*C*-methyl-guanosine per strand, that this modification leads to a significant distortion of the canonical A-form backbone geometry (M. Egli, G. Minasov, L. Beigelman, unpublished results). The hammerhead ribozyme constitutes the best characterized system among all catalytic RNAs. Nevertheless, the combined biochemical and structural evidence has failed to unveil all secrets of this ribozyme (24,60, 94,106). Thus, the geometric and hydrogen bonding properties of a hammerhead–substrate complex that has an active conformation remain elusive, and controversy still exists concerning the number and precise roles of metal ions participating in catalysis.

2.2 The P4–P6 domain of the group I intron RNA from *Tetrahymena thermophila*

The P4–P6 domain constitutes an autonomously folding unit of the self-splicing group I intron RNA from *T. thermophila* (67,68). Its crystal structure reveals a hairpin-like overall fold with close side-by-side packing of helical stacks (14) (Figure 3). Three structural motifs mediate the tight packing of helices through minor groove contacts. A more detailed description of these is given in Table 2. An 11-nucleotide conserved domain, termed tetraloop-receptor, provides an anchor point for an apical GAAA tetraloop (Figure 3). Embedded in the tetraloop receptor is a so-called A-platform, two contiguous adenosines lying next to one another and extending the helix by one stacking unit, thus forming a pseudo base pair and a sort of platform (15) (Figure 4). The accommodation of two purine residues next to each other widens the minor groove, a feature that likely facilitates lateral interactions between duplexes. The fact that this structural motif occurs three times in the P4–P6 domain may be taken as an indication of the importance of this building block for RNA folding. A second motif that stabilizes the sharp bend of the molecule and clamps together helical regions is an A-rich bulge. Its corkscrew conformation exposes the bases with their hydrogen bonding functions on the surface while magnesium ions glue together multiple phosphate groups in the core (17). A further motif that mediates close packing of duplexes is the ribose zipper, which is a network of hydrogen bonds between the 2′-hydroxyl groups of riboses in the minor grooves of adjacent helices (Table 2). It is noteworthy that all of these close contacts between duplexes take place in the minor groove, which is wide and shallow in the RNA A-form stems and features the 2′-hydroxyl groups that are capable of donating and accepting hydrogen bonds.

The structure of the P4–P6 domain has demonstrated important principles of RNA packing and has provided evidence for the need of tight packing for the function of large catalytic RNAs such as the spliceosome and the ribosome. One question that had intrigued many in the field

Table 2. Tertiary structural motifs observed in RNA X-ray crystal structures

Motif	Structure(s)	Explanation
A-platform	P4–P6 domain, Group I intron	Two adjacent adenosines in the sequence lie side-by-side and form a pseudo base-pair within a helix (Figure 4)
Cross-strand purine stack	hammerhead ribozyme, P4–P6 domain Group I intron, 5S rRNA - helix I, IV and loop E, sarcin/ricin loop (SRL) 28S rRNA	Stacking between a purine from one strand and a purine from the other under formation of stabilizing intra- and inter-strand contacts involving 2′-hydroxyl groups (Figure 9)
C-turn	beet western yellow virus (BWYV) frameshift pseudoknot	Transition between a stem and a loop region sharing certain features with the U-turn, but being less sharp; the turn involves three residues C1, A, and C2 and formation of a hydrogen bond between N4 of C2 and a phosphate oxygen from the adenosine
Four-way junction	10–23 DNA enzyme/RNA complex	The deoxyribozyme and its RNA target strand unexpectedly form a 2:2-complex with a four-way junction; its three-dimensional structure is reminiscent of the Holliday junction that occurs during genetic recombination.
G•U tandem	P4–P6 domain, Group I intron	Adjacent wobble base pairs, either at 5′-GU-3′/3′-UG-5′ or 5′-GG-3′/3′-UU-5′ steps, that form metal ion binding sites (Mg(II), Os(III) hexammine) in the major groove
Metal ion binding	hammerhead ribozyme (DNA/RNA)	Cd(II), Mn(II)
	hammerhead ribozyme (all-RNA)	Cd(II), Co(II), Zn(II), Mg(II), Mn(II), Tb(III) (31)

Table 2. Continued

Motif	Structure(s)	Explanation
	P4–P6 domain, Group I intron	Co(III), Cs(I), Mg(II), Os(III), Sm(III), Tl(I)
	5S rRNA, loop E	Mg(II) (Figure 10)
	TAR RNA	Ca(II) (Figure 11)
	BWYV frameshift pseudoknot	Mg(II), Na(I)
	leadzyme	Ba(II), Mg(II), Pb(II) (Figure 8)
Minor groove triplex	BWYV frameshift pseudoknot	Stabilized by 16 hydrogen bonding contacts in the minor groove, 10 of which are mediated by 2′-OH groups (Figure 14)
	Group I intron	Predicted between the P1 substrate and the J8/7 segment in the catalytic core, based on nucleotide analogue interference suppression experiments (101)
Pseudoknot	BWYV pseudoknot	Regulates translation of polycistronic mRNA by inducing a -1 ribosomal frameshift (Figure 14)
	HDV ribozyme	Nested double pseudoknot that mediates side-by-side packing of helical regions (Figure 6)
Purine-rich bulge	P4–P6 domain, Group I intron leadzyme	An A-rich bulge AAUAA bridges the P4 helix of the group I intron core and the three-way junction formed by helices P5a, P5b, and P5c; the corkscrew turn exposes the bases on the outside with the phosphate groups forming the core, stabilized by divalent metal ions (Figure 3). In the leadzyme, the G24-A25-G26 bulge 3′-adjacent to the cleavage site is part of an internal loop; the three bases are extrahelical and stacked.

Table 2. Continued

Motif	Structure(s)	Explanation
		Unstacking of G24 sharply kinks the backbone at the scissile bond and may be associated with cleavage (Figure 8). (108)
Ribose zipper	P4–P6 domain, Group I intron HDV ribozyme	Interdigitated riboses mediating close contacts between sugar-phosphate backbones; the ribose zipper is characterized by direct hydrogen bonds between 2′-hydroxyl groups and bases
S-turn	SRL region, 28S rRNA	Backbone conformation near the bulged G4319 in the sarcin/ricin loop with two sharp bends giving it an S-shaped appearance. The bends are accompanied by C2′-endo puckers of sugars and reversal of chain direction; the S-turn is stabilized by base stacking and hydrogen bonds between 2′-hydroxyl and phosphate groups (Figure 13)
Tetraloop	hammerhead ribozyme P4–P6 domain, Group I intron TAR RNA aptamers	Hairpin loop of sequence GNRA (N, any nucleotide; R, purine nucleotide) frequently encountered in large structured RNAs; the three-dimensional structure bears resemblence to that of the U-turn (50) (Figures 1, 12)
Tetraloop receptor	P4–P6 domain, Group I intron (hammerhead ribozyme)	11-nucleotide motif that interacts specifically with the GAAA subclass of tetraloops. In the P4–P6 domain, it consists of two adjacent C-G base pairs, a 5-nucleotide internal loop, and a G•U base pair. The tetraloop docks in the minor groove of the tetraloop receptor and the

Table 2. Continued

Motif	Structure(s)	Explanation
		interactions involve stacking and hydrogen bonds specific to A (Figure 3); in the crystal structure of a hammerhead ribozyme this motif was observed as a lattice contact between the GAAA loop and a pseudo-receptor from an adjacent helix (79)
U-turn	hammerhead ribozyme	First observed in the anticodon and pseudo-uridine loops of $tRNA^{Phe}$; the consensus sequence for a U-turn is UNR (N, any nucleotide; R, purine nucleotide); Bases of the U-turn in hammerhead ribozymes are absolutely conserved and either mediate binding of the catalytic metal ion or stabilize the active site through a number of interactions (92) (Figure 1)

prior to the visualization of a large RNA by X-ray crystallographic techniques was how RNA could tightly pack despite the negatively charged backbones on its surface? One of the surprising observations in the structure of the P4–P6 domain is that close spacing of stem regions can occur in the absence of metal cations that act as coulombic shields against the repulsions between negatively charged phosphate groups. While the ribose zipper will be a recurring motif in future structures, there

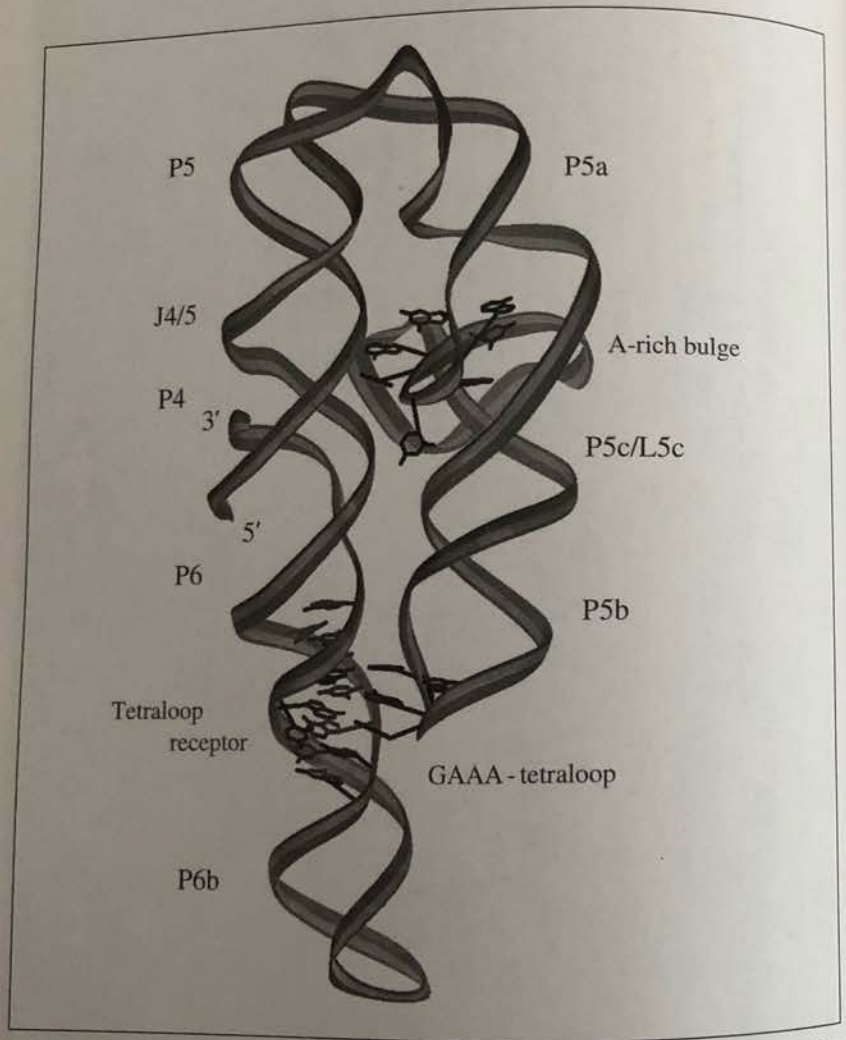


Figure 3. Ribbon diagram of the overall fold of the P4–P6 domain of the *T. thermophila* group I intron RNA. The drawing illustrates specific contacts between bases of the A-rich bulge (center, the bases of residues U182 to G188 are shown) and the minor groove of the P4 helix and the P5abc three-helix junction as well as the docking of the L5b GAAA tetraloop in the tetraloop receptor between helices P6a and P6b. Individual regions of the P4–P6 domain are labeled and darker regions are closer to the viewer.

is some indication from oligodeoxynucleotide crystal structures that DNA that lacks the 2′-hydroxyl group may also be capable of adopting tightly packed arrangements (26).

Despite these novel packing-mediating motifs, metal ions do play important roles in enabling close intramolecular contacts in the P4–P6 domain. Three different types of metal-ion–binding sites were identified in its crystal structure (8). Magnesium ions form outer sphere complexes in the major groove of helices, with non-canonical regions that comprise G•U tandems constituting preferred coordination sites (16). As mentioned above, magnesium ions also form the core of the domain by directly coordinating to phosphate oxygens and carbonyl groups of guanosines (17). In addition, nucleotide analogue interference mapping and soaking of P4–P6 domain crystals in solutions of electron-rich monovalent metal cations have provided strong evidence that potassium ions are associated with A-platforms and make inner-sphere contacts to a G under the platform (8). Potassium coordination is crucial for folding of the P4–P6 domain and was shown to enhance the activity of the Azoarcus group I intron.

2.3 The *T. thermophila* ribozyme

The crystal structure of the catalytic core of this group I intron was determined at 5 Å resolution (41). The fragment comprises the P4–P6 domain described in the previous section and the P3–P9 domain that consists of five helices including P8, P3, P7, and P9 (a figure with the details of the secondary structure of the *T. thermophila* ribozyme can be found in Reference 41 by Golden and coworkers). However, the fragment lacks the second substrate of the group I ribozyme, the 5′-splice site within the P1 duplex. The conformation of the P4–P6 domain in the structure of this 247-nucleotide core appears more or less unchanged compared with the crystal structure of the isolated domain (Figure

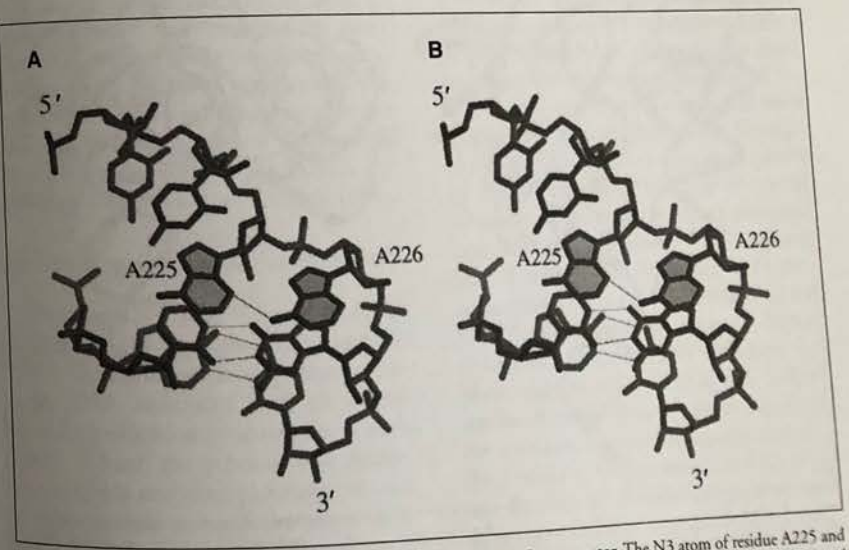


Figure 4. Stereo drawing of the adenosine platform in the J6a/6b tetraloop receptor. The N3 atom of residue A225 and the N6 atom of residue A226 are hydrogen bonded (thin solid line). The base planes of A225 and A226 are filled and residues are labeled.

5A). The P8, P3, P7, and P9 regions display coaxial stacking and thus form a pseudocontinuous helix with a bend of about 50° that is located between P3 and P7 and within P7 (Figure 5). The junction of the P4–P6 and P3–P9 domains in the group I intron structure involves a close approach of the P4, P6, P3, and P7 helices. This close packing is mediated by base triples that comprise the J3/4 and J6/7 regions. In the case of the latter junction, the base triples also help to orient the P7 helix by spanning the entire length of this helix and thus limiting the orientation of the guanosine-binding site relative to the P4 and P6 helices.

The guanosine-binding site, one of the substrate binding sites of the group I intron, is located in the P7 helix. Keeping in mind the low resolution of the structure, it appears that the geometry of this helix deviates significantly from the A-form geometry. In the current model, residue A263 is unpaired and stacked within the P7 helix. The unpaired nucleotide and additional backbone–backbone interactions with the P4 helix result in a compression of the major groove. This provides a snug binding site for the guanosine substrate. The close packing of the P4–P6 and P3–P9 domains generates a shallow cleft that appears capable of binding the short helix (P1) that contains the 5′ splice site. The sides of this concave binding site are constituted by the junction J4/5 and P3, and the junction J8/7 forms its floor. Because a contact between P1 and J4/5 at the active site has been implicated on the basis of functional group substitution data (96), it was concluded that the ribozyme core is preorganized for substrate binding (41). Further evidence for the requirement of only a minor reorganization of the core to accommodate P1 appears to come from the fact that four phosphate groups that span the guanosine binding site and are known to be sites of phosphorothioate interference would surround the phosphate at the 5′ splice site when P1 is docked at the cleft present in the core.

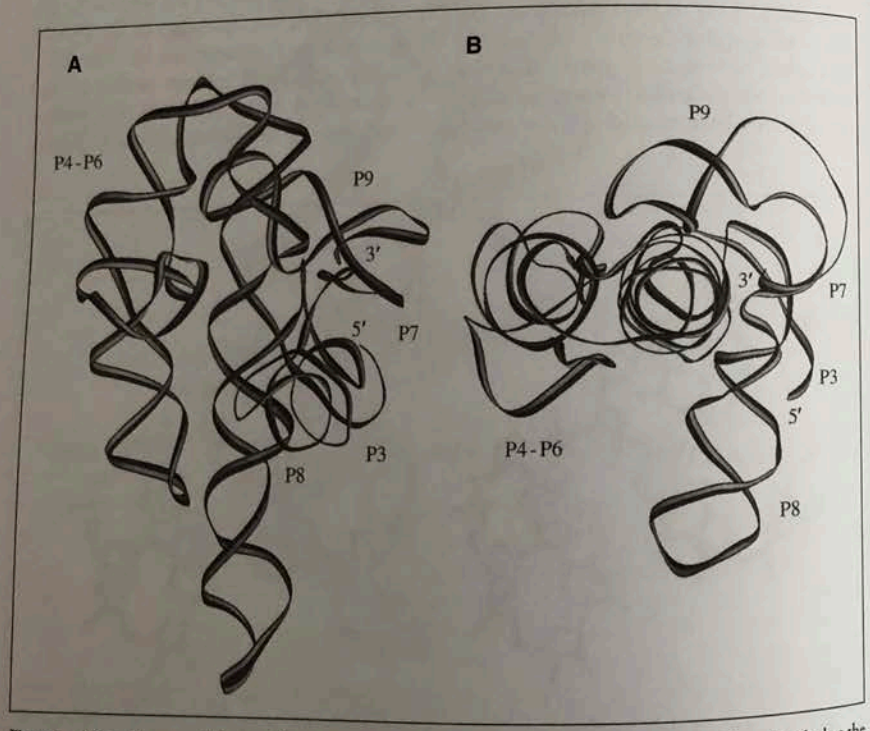


Figure 5. Ribbon diagrams of the overall structure of the *Tetrahymena thermophila* ribozyme. (A) Viewed such that the P4–P6 domain is on the left, such that its orientation is rotated by approximately 180° around the vertical relative to the one in Figure 3. (B) The ribozyme rotated roughly 90° around the horizontal and viewed along the P4–P6 helical stacks, which illustrates the bend between the P3 and P7 regions.

However, a recent publication based on nucleotide analog interference suppression (NAIS), an extension of nucleotide analog interference mapping (NAIM) (37), has reported that the P1 substrate helix and the single-stranded junction region J8/7 form an extended minor groove triple helix in the core of the group I ribozyme (101). The conclusions drawn from the crystallographic work with respect to the conformational preorganization of the ribozyme for substrate binding and the results based on the NAIS study are not consistent. In order for the J8/7 strand to form base triples in the minor groove of the P1 helix as well as with helix P3, it needs to rotate by 180° and translate by more than 20 Å relative to the position it occupied in the 5 Å crystal structure of the intron core. This would suggest that the J8/7 junction in the crystal structure of the ribozyme lacking the P1 helix is not preorganized for substrate binding (101).

2.4 The Hepatitis Delta Virus ribozyme

The hepatitis delta virus is an RNA satellite virus of hepatitis B virus, and the HDV ribozyme is the only known catalytic RNA associated with an animal virus. Its replication produces linear multimers of both the genomic and antigenomic RNAs and their processing to unit length proceeds via a self-cleaving activity of the RNA in the absence of any protein (9,54). The catalyzed reaction is the same as in the case of the hammerhead ribozyme: attack by a 2′-hydroxyl group at the adjacent phosphate under inversion of the phosphate group and formation of a free 5′-hydroxyl terminus and a 2′,3′-cyclic phosphate. However, there are several differences between the hammerhead and the HDV ribozymes. In the case of the latter, the presence of a single nucleotide 5′ to the cleavage site is sufficient to trigger cleavage, and its identity has practically no effect on the cleavage rate. The HDV ribozyme is faster than any other naturally ocurring ribozyme (>1 cleavage per second, hammerhead approximately 1 per minute), is resistant to denaturants, and does not require the specific involvement of metal ions. Cleavage takes place in solutions of several divalent metal cations with lower than 0.1 mM concentrations. By comparison, hammerhead ribozyme self-cleavage typically requires a magnesium ion concentration of around 10 mM.

The crystal structure of the self-cleaved form of a genomic HDV ribozyme was determined at medium resolution (Table 1) (Reference 33). The earlier observation that variation of the length of the P4 stem does not significantly affect activity of the HDV ribozyme was used to graft the RNA-binding domain of the U1A spliceosomal protein onto that stem (Figure 6, see Methods section). Crystals of this HDV construct were then grown in the presence of U1A. The co-crystal structure shows a very compact ribozyme core with the 5′-hydroxyl leaving group buried inside the active site cleft and sequestered from solvent. Similar to the P4–P6 domain in the group I intron crystal structure, the HDV ribozyme displays side-by-side packing of helical stacks. A total of five helical segments form two stacks, the first comprising P1, P1.1, and P4, the second P2 and P3 (Figure 6). The P1.1 helix consists of two base pairs and had not been anticipated based on previous work. However, mutagenesis showed that the four residues of

P1.1 are important for ribozyme function, presumably because their mutation or deletion would disrupt one of the pseudoknots in the structure. Compared with the P4–P6 domain, the side-by-side packing of helices in the HDV ribozyme is much more compact and convoluted. Instead of being mediated by two sets of loop-to-helix groove contacts, as in the case of the intron domain, the HDV ribozyme features five crossovers, generating a nested double pseudoknot with the P1.1–P3 pseudoknot nesting in the P1–P2 pseudoknot (Figure 6). Interestingly, no metal ions were located in the electron density maps, and the stability of the structure is thus based on base-pairing, stacking, base-backbone, and backbone-backbone (e.g., ribose zipper) interactions. The convoluted fold of the HDV ribozyme is consistent with its stability to denaturants such as urea or formamide even at elevated temperatures.

The globular structure of the HDV ribozyme is more similar to those of protein enzymes and its buried active site differs from the solvent accessible region around the scissile bond in the hammerhead ribozyme (Figure 7). The active site cleft is enclosed by the P1 helix that carries the substrate, the single stranded J4/2 region, and the niche constituted by the P3 minor groove and the L3 loop (Figure 6). In addition, the cross-strand between helices P1 and P3 seals the active site in the front. The active site residue G1 is engaged in a wobble base pair with U37 at the lower end of the P1 helix. The resulting shift of G1 into the minor groove allows it to engage in inter-strand stacking with the G38 residue of the P1.1 minihelix, itself composed of base pairs C21–G39 and C22–G38. Formation of the G•U base pair goes along with a distortion of the backbone that appears to be one factor that facilitates the in-line approach of the 2′-hydroxyl group relative to the scissile bond. The lack of any specific involvement of metal ions in the HDV-catalyzed phosphoryl transfer reaction suggests that a particular nucleotide acts as the general base in the activation of the 2′-hydroxyl group for attack. This role appears to be played by residue C75 that is located in the single-stranded J4/2 region and projects into the active site. From the crystal structure one can conclude that the active site of the HDV ribozyme is preorganized for catalysis and sequestered from solvent and thus bears some resemblance to the active sites of protein enzymes. In addition, the correlation of the structural data with the results from earlier mutation experiments has led to the identification of residues that either directly or indirectly participate in catalysis (33).

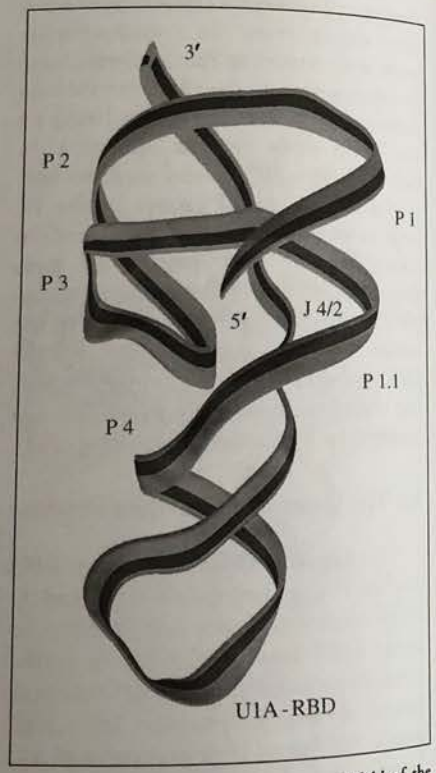


Figure 6. Ribbon diagram of the overall fold of the HDV ribozyme with the engineered U1A RBD at the bottom. The individual helical regions and the terminal residues of the RNA are labeled. The view illustrates the tight fit of the J4/2 region between the helical stacks.

2.5 The leadzyme

The leadzyme is a small catalytic RNA motif that emerged from in vitro selection experiments that were stimulated by the finding that lead hydroxide produces a site-specific cleavage in tRNA (75). As in the case of the naturally occurring ribozymes, the catalyzed reaction is a site-specific 2′-hydroxyl-mediated cleavage, which results in a free 5′-hydroxyl group and a 2′,3′-cyclic phosphate. In addition, the leadzyme processes the cyclic phosphate to the 3′-phosphate product. Cleavage requires lead ions (76), but the cleavage rate (around 1 per minute) is enhanced in the presence of certain earth alkali and rare earth ions (98). The leadzyme is a rather simple RNA motif; two base-paired stems enclose an asymmetric internal loop that contains the cleavage site. This core is rich in purines, with the tetranucleotide C*GAG (the asterisk marks the cleavage site) bridging the stem regions in one strand while the loop portion of the opposite strand features A and G. The leadzyme is an attractive target for structural studies as its analysis

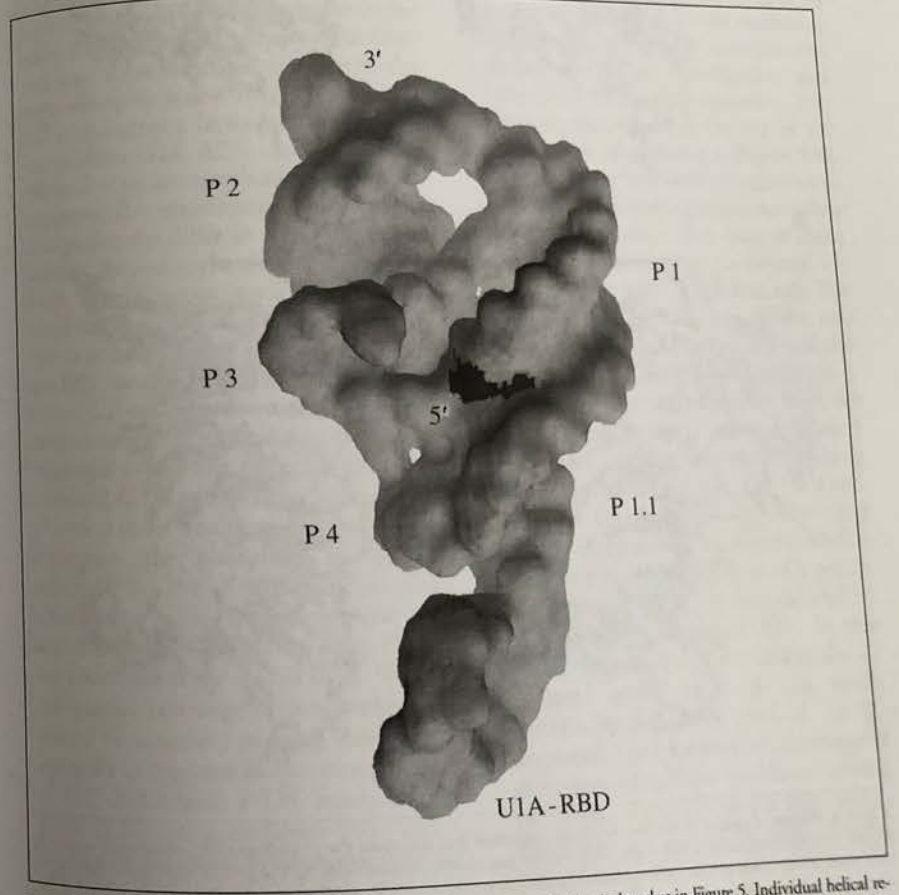


Figure 7. Surface model of the HDV ribozyme. The orientation is nearly identical to that in Figure 5. Individual helical regions are labeled and the substrate-bearing active site nucleotide G1 (the 5′-terminal residue in the present construct) is highlighted in black. The drawing illustrates the formation of a deep solvent-inaccessible active-site cleft by the HDV ribozyme.

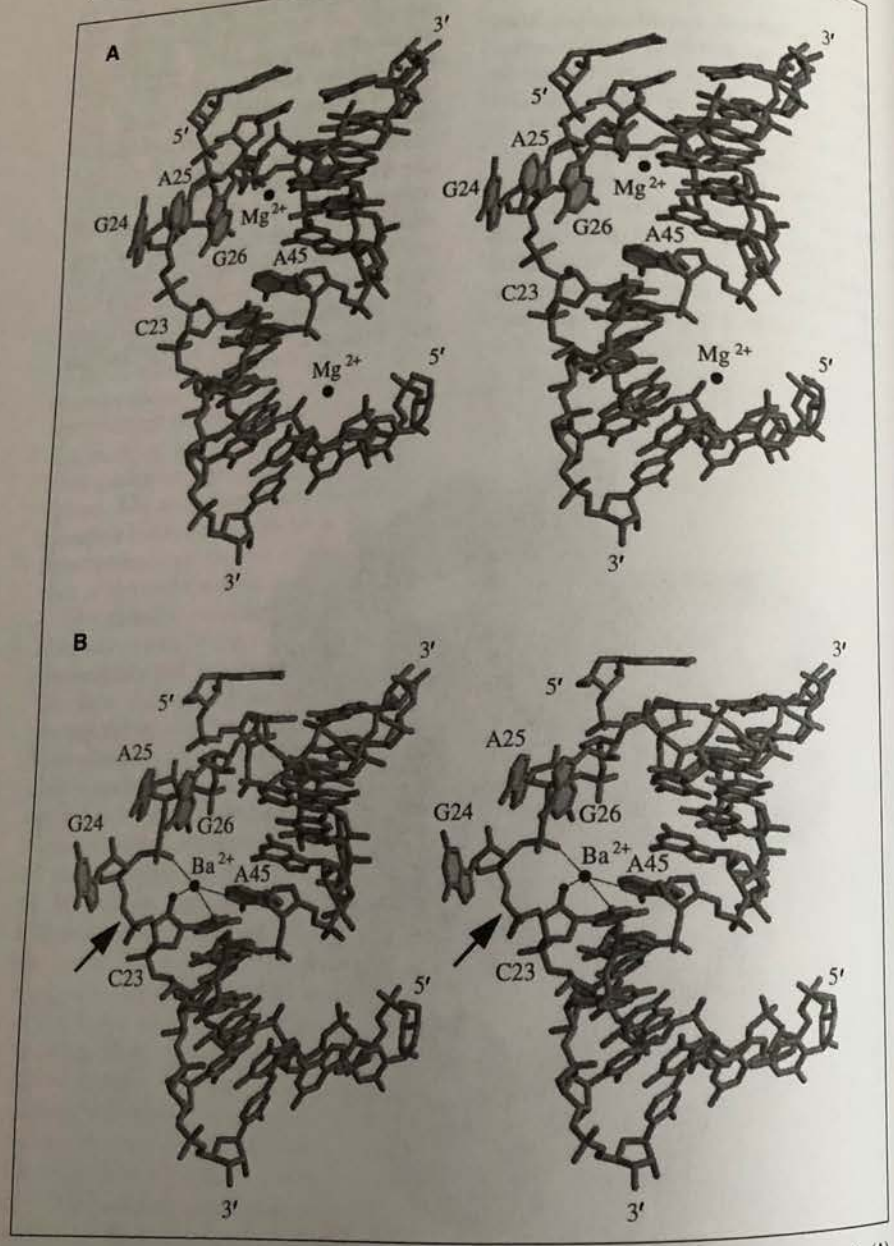


Figure 8. Ribbon diagrams of the two leadzyme molecules that constitute the crystallographic asymmetric unit. (A) Stereo drawing of molecule 1 that displays flipped-out and stacked residues G24, A25, and G26. (B) Stereo drawing of molecule 2 with the same residues flipped-out but G24 unstacked from A25 and G26, thus leading to a sharp kink in the ribose-phosphate backbone at the scissile bond (marked with an arrow). The base planes of residue C23, G24, A25, G26, and A45 are filled, the 2′-oxygen of C23 is highlighted with a sphere, metal ions [Ba(II) and Mg(II)] are drawn as black spheres, and terminal residues are labeled.

promises insight into the essential features of metal-ion–assisted catalysis. Moreover, because the lead ion is a strong X-ray scatterer, one would expect a crystal structure of an active form of the ribozyme to shed light on the roles of metal ions in catalysis and structure stabilization.

The crystal structure of a leadzyme was recently determined to 2.7 Å resolution (108). In the crystallographic RNA construct the asymmetric internal loop was flanked by two tetramer duplexes, and both strands featured single 5′-overhanging cytosines to promote crystal packing. The asymmetric unit of the crystals consists of two independent leadzyme molecules (Figure 8). Both display a continuous base-stacked helical structure of the short ribozyme strand that comprises the AG-loop segment. In the substrate strand, the purines G24, A25, and G26 of the internal loop are extrahelical (Figure 8). However, the trinucleotide bulges of the two molecules differ in some important aspects. In one molecule the three bases form a continuous stack (Figure 8A), while the base of G24 in the second molecule is unstacked from the other two (Figure 8B). Moreover, the ribose of C23 flips into a C2′-endo conformation as a result of the unstacking whereas the corresponding sugar in the first molecule assumes a standard C3′-endo pucker. In the phosphoryl transfer reaction catalyzed by the leadzyme, the phosphate of G24 is attacked by the 2′-oxyanion of C23. Therefore, it is intriguing that the unstacking of G24 leads to a sharp kink of the backbone at the scissile bond. While a substantial rearrangement of molecule 1 would be necessary to bring about a catalytically competent in-line conformation of O2′ and scissile P-O5′ bond, a slight twist of the phosphodiester bond and a change of the ribose conformation to the C3′-endo pucker in molecule 2 would generate a near in-line arrangement. Thus, the sugar-phosphate backbone geometry of the active form of the leadzyme differs from the conformational properties of the backbone that are associated with strand cleavage catalyzed by the hammerhead ribozyme. There, the ribose of the residue that carries the attacking O2′ has to undergo a conformational change away from the helical C3′-endo conformation. It is tempting to speculate that the two conformations of the molecules trapped in the leadzyme crystal are somehow representative of ground-state and pre-catalytic forms of the RNA.

The crystal structure of the leadzyme clearly pointed to the inherent flexibility of RNA as the fundamental basis of its catalytic abilities. Beyond the conformation of the RNA, the structure also revealed several ion coordination sites (108). Two hydrated magnesium ions and a barium ion were located in the major groove of molecule 1 (Figure 8A). No magnesium ions were coordinated to molecule 2, but soaking experiments furnished three barium sites (one of them minor and equivalent to the barium site in molecule 1) and a single lead site. The latter is identical with one of the two major barium sites in molecule 2 and also matches a magnesium site in molecule 1. Because this site is far removed from the scissile bond, the metal ions coordinating there likely play a structure-stabilizing rather than a catalytic role. The third barium ion coordination site is in close vicinity of the scissile bond and modeling a lead ion in its place suggested a plausible mechanism for the lead-dependent cleavage reaction (Figure 8B). In this model the lead ion coordinates to the 2′-hydroxyl group of C23 and could abstract its proton to produce the 2′-oxyanion. The conformational change of the ribose-phosphate backbone toward the in-line geometry would then shift a non-bridging oxygen of the attacked phosphate group into the coordination sphere of the lead ion, thus allowing it to dissipate the negative charge in the trigo-

nal-bipyramidal transition state. However, the fact that the lead ion also coordinates to N1 of A45 and that A in this location can be replaced by U, C, or even an abasic residue without drastically impairing activity is somewhat unsettling. Therefore, it appears that the structural work has not yet removed all controversies that surround the mechanism of this small catalytic RNA.

2.6 The 5S rRNA domain

The crystal structures of an *E. coli* 5S rRNA domain that comprises helices I and IV and loop E at 3 Å resolution and of a loop E dodecamer at 1.5 Å resolution have demonstrated the effects of non-canonical base pairs on RNA topology and the roles of magnesium ions and water in structure formation and stabilization (20). The structure of the helix I and IV/loop E domain, termed fragment I, revealed a more or less linear helical stack of almost 100 Å length, a dimension that is roughly equal to the radius of the *E. coli* ribosome. Helix IV, located at one end, and loop E, located in the center, stack in an approximately colinear fashion. However, loop E and helix I, which is located at the other end of the helical stack and connected to loop E via a single-stranded linker, are not coaxially stacked. The flexible linker allows helix I to rotate and translate relative to loop E such that the terminal residues of both are oriented in a way that would permit docking of the stem composed of residues 12 to 69, which are missing in fragment I.

Loop E and helix IV are rich in non-canonical secondary structure, including a 3–base-pair motif that occurs three times in this portion of 5S rRNA. The motif consists of a standard Watson-Crick G-C base pair followed by a sheared A•G base pair and a reverse Hoogsteen A•U base pair. A particular feature of this motif is the stacking of the adenine of the

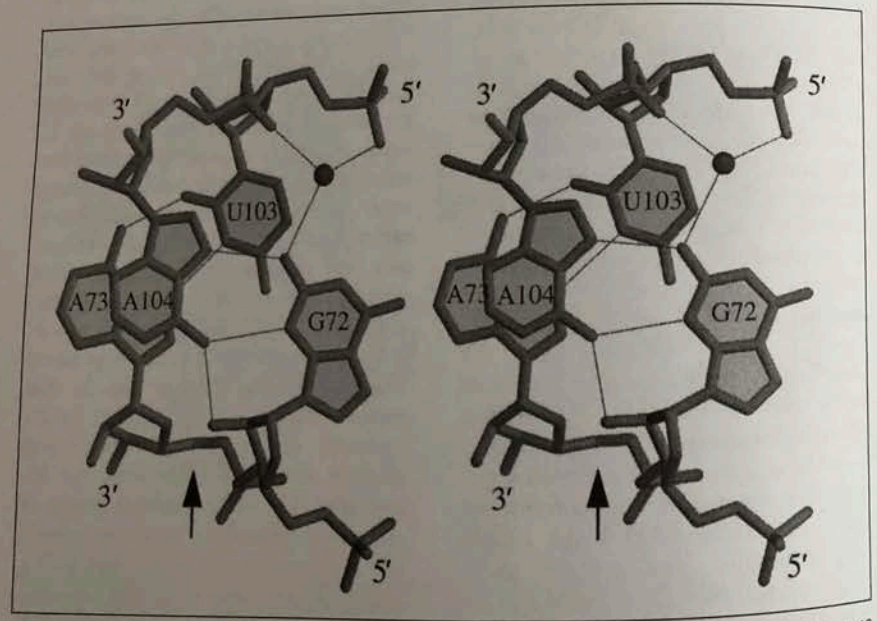


Figure 9. Stereo drawing of the cross-strand purine stack between adenosines A73 and A104 in the loop E of *E. coli* 5S rRNA. Hydrogen bonds are drawn as thin solid lines. The arrow marks the kink in the backbone of the G72-A73 strand in the region of the O5′-C5′ and C5′-C4′ bonds, and the dark sphere represents a water molecule.

G•A base pair from one strand on the adenine of the A•U base pair from the other (Figure 9). Due to this stacking mode between purines from opposite strands the motif was termed "cross-strand purine stack" (Table 2) (Reference 20). Besides stacking, this arrangement involves intra- and inter-strand hydrogen bonding that is in part mediated by a 2′-hydroxyl group and a water molecule. A further characteristic of the cross-strand stacking motif is a severe kink in the backbone of the adenosine that participates in the reverse Hoogsteen base pair (Figure 9). Loop E contains two cross-strand A stacks and helix IV features a cross-strand G stack. Similar motifs are also present in the structures of other RNAs (Table 2). Combinations of such cross-strand purine stacks can modulate the widths of the minor and major grooves in RNA. For example, the major groove of loop E is contracted by up to 6 Å relative to a standard A-form major groove and the minor groove is expanded by about 2 Å. Conversely, the major groove in helix IV is increased by about 7 Å relative to the groove in a canonical RNA duplex and thus is about as wide as the major groove of a B-form duplex. In the case of loop E, the narrowing of the major groove may in part be due to the

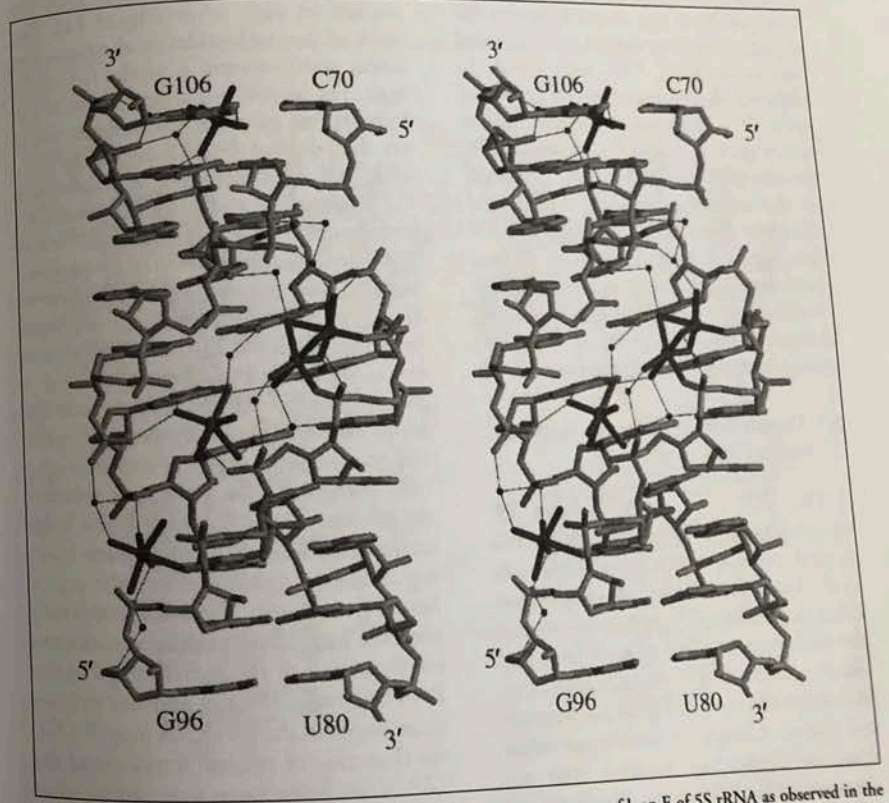


Figure 10. Stereo drawing of the magnesium ion zipper in the major groove of loop E of 5S rRNA as observed in the 1.5 Å resolution crystal structure of that fragment (20). Magnesium ions and their hydration shells are highlighted in black, water molecules are drawn as filled circles, and hydrogen bonds between hydrated ions or water molecules and RNA atoms are drawn as thin solid lines.

coordination of magnesium ions at that site (Figure 10). The metal ions engage in outer- and inner-sphere interactions in the major groove of loop E, an arrangement that was referred to as a "metal zipper", owing to the location of the ions in the center of the groove, with contacts being made to base functions and non-bridging phosphate oxygens from both strands.

The topological changes brought about by the cross-strand purine stacks in combination with three non-canonical base pairs that separate them in loop E define part of a distinctive surface that provides a recognition site for the ribosomal protein L25 (20). The wider minor groove of loop E facilitates access to an unusual array of hydrogen bond donors and acceptors displayed by the three central non-canonical base pairs (G•U, A•G, and G•G). In addition, the expanded major groove of helix IV is positioned adjacent to the minor groove of loop E and valuable sequence information can easily be probed by the interacting protein. The crystal structures of portions of *E. coli* 5S rRNA have provided a nice illustration of how metal ions and non-canonical secondary structural motifs can mediate specific changes in the topology of an RNA that permit its recognition by proteins.

2.7 The HIV *trans*-activation response region (TAR RNA)

The TAR RNA of HIV-1 is a 59-nucleotide stem-loop motif at the 5′-ends of viral transcripts that plays a vital role in the regulation of HIV gene expression (34,51). It constitutes the binding site for the viral *trans*-activator protein Tat, a regulatory factor that stimulates transcription from the so-called viral long-terminal repeat. Compared with most other retroviral regulatory proteins that act through DNA elements, Tat-mediated regulation thus involves a specific *cis*-acting RNA regulatory domain. The Tat recognition elements include a UCU trinucleotide bulge, of which the first residue (U23) appears most crucial, and two base pairs to the 5′- and 3′-sides of the bulge. NMR solution–structures of the complex between TAR RNA and Tat peptide from the bovine immunodeficiency virus are consistent with a recognition of the major groove region of TAR that is adjacent to the trinucleotide bulge by a β-hairpin motif (82,113).

The crystal structure of an apical portion of the TAR stem-loop that comprises all elements necessary for transcriptional activation by Tat in vivo was determined to near-atomic resolution (46). All three bases of the UCU bulge are extrahelical and U23 and C24 are stacked on each other (Figure 11). For most of the nucleotides in the stem, the ribose conformation is of the C3′-endo type. The looped-out arrangement of the bulge region goes along with sugar puckers that deviate from the A-form C3′-endo type. For example, U23 adopts a C2′-exo pucker and U25 adopts a C2′-endo pucker. The bulge conformation is stabilized by the coordination of three calcium ions, one of which is coordinated to phosphate oxygens of all three bulged nucleotides (Figure 11). Another calcium ion is engaged in an interaction with a pyrimidine base (U25), which provides some evidence that the observed metal ion coordination mode is sequence-specific. However, it is uncertain whether the calcium ions maintain only the bulge conformation or whether metal ion binding serves the stabilization of the crystal lattice in addition. Like the previously observed magnesium-stabilized corkscrew conformation of the A-rich bulge in the P4–P6 domain (14,17) and the magnesium zipper in the 5S rRNA loop E (20), the clustering of calcium ions around the TAR RNA bulge once more emphasizes the influence of metal-ion binding on RNA structure and stability.

Does the crystal structure provide some

insight concerning the interaction of TAR with the Tat protein? The major groove in the vicinity of the bulge is widened, thus promoting coordination of a fourth calcium ion in the groove (Figure 11) and possibly allowing side chains of the Tat protein to probe the floor of the groove and to engage in specific interactions with base functions. However, mutation of U23 leads to a reduction of Tat affinity both in vitro and in vivo (51). These data cannot be reconciled with the orientation of U23 in the TAR crystal structure; the base moiety is directed away from the major groove, preventing a contact with the protein that presumably binds within the major groove. In this respect the crystal structure differs significantly from those derived from NMR solution experiments (1,81). In both of those structures, the base of residue U23 is directed into the major groove. It is possible that the relatively high concentration of calcium ions in the crystallization experiments (approximately 100 mM) may have furnished a bulge structure that differs from the one in the peptide-RNA complex in vivo at much lower concentrations of divalent metal ions. Presently, there is only limited information on the influence of metal ions on the TAR RNA bulge conformation in solution and how metal ion coordination affects Tat activity. A more definitive answer to open questions regarding TAR-

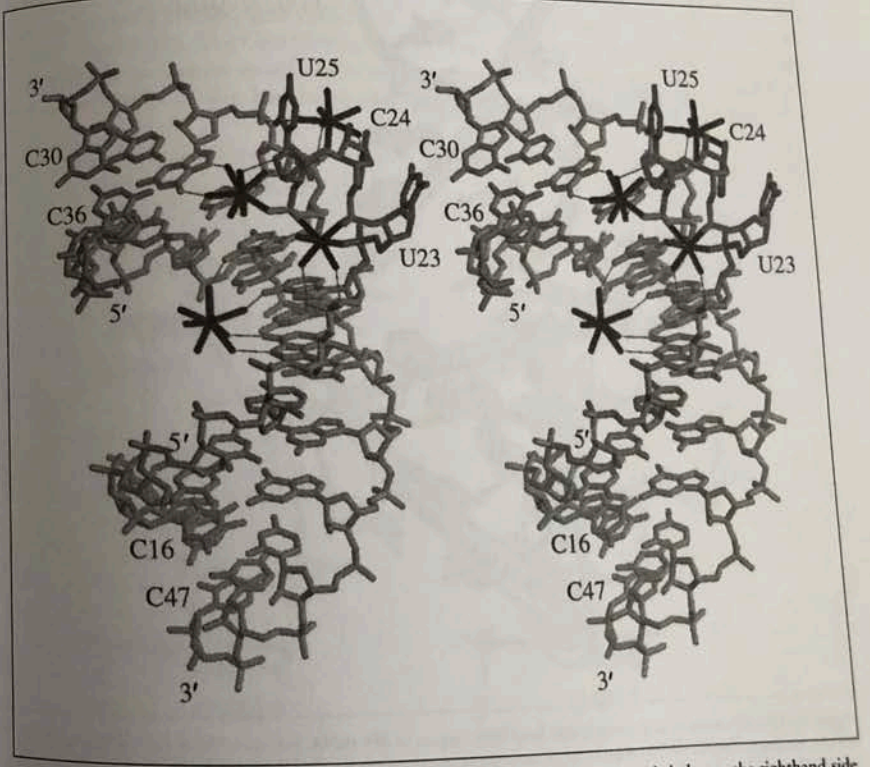


Figure 11. Stereo drawing of the stem region of TAR RNA with the UCU-trinucleotide bulge on the righthand side. Calcium ions and their hydration shells are highlighted in black, the bases of bulged residues are labeled, and hydrogen bonds between ions and RNA are drawn as thin solid lines.

Tat recognition might come from a future high-resolution structure of the RNA-peptide complex.

2.8 The sarcin/ricin loop of 28S rRNA

The sarcin/ricin loop (SRL) constitutes a highly conserved domain of 28S ribosomal RNA and serves an important function in protein synthesis by interacting with the elongation factors EF-Tu and EF-G. Its name refers to the recognition of particular elements within the loop by the ribotoxins sarcin and ricin, ribonucleases that cleave the phosphodiester backbone on the 3′-side of G4325

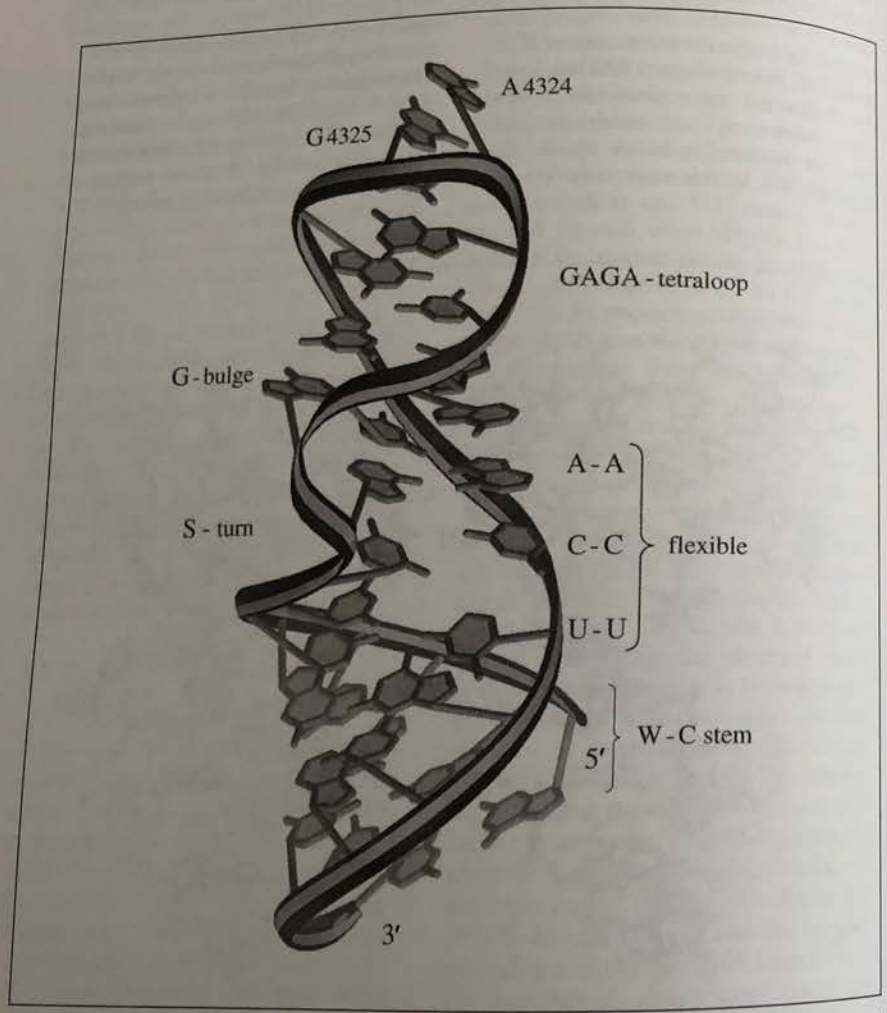


Figure 12. Overall structure of the sarcin/ricin loop (SRL) region of 28S rRNA. Residues A4324 and G4325 are labeled (the numbering is for 28S rRNA from the rat). The sarcin ribonuclease cleaves the phosphodiester backbone on the 3′-side of G4325, and ricin depurinates A4324. Domains of the SRL structure comprise the Watson-Crick base paired stem base (bottom), the flexible region (base pairs U4316•C4332, C4317•4331, and A4318•A4330, middle), the G-bulged cross-strand A stack (top), the S-turn (foreground, left), and the GAGA (4323–4326) hairpin tetraloop.

and depurinate A4324, respectively (Figure 12), thereby inactivating ribosomes (29,43). It is noteworthy that 12 out of 17 residues of the SRL are almost universally conserved. Moreover, the SRL is the only rRNA region that is recognized by both elongation factors (64). Binding of EF-G to an SRL oligoribonucleotide is just one order of magnitude weaker than binding of the elongation factor to the entire *E. coli* ribosome (66).

The crystal structure of a 29-nucleotide fragment containing the SRL region from rat offers a detailed look at the conformational properties of this key site in ribosomal RNA (21). The stem-loop structure is composed of a Watson-Crick paired base, a central flexible region that in turn stacks with a G-bulged cross-strand A stack, and the GAGA tetraloop, thus comprising A4324 and G4325 at the top (Figure 12). The G4319 bulge is embedded in an S-turn motif, featuring two sharp turns of the backbone and inversions of the standard C3′-endo ribose conformations (Figure 13, Table 2). Thus, the SRL RNA is a perfect example of the modular nature of RNA structure; basic building blocks are varied conformationally and assembled in multiple ways. The cross-strand A-stack seen here resembles similar arrangements in the hammerhead ribozyme, the 5S rRNA, and the P4–P6 domain of a group I intron (Table 2) (Reference 21).

An interesting feature of the flexible region is the presence of two sheared pyrimidine•pyrimidine base pairs (U•C and C•C, Figure 12) that are each stabilized by only one direct hydrogen bond. The slightly wider major groove in the flexible region and the adjacent cross-strand A-stack with the G4319 bulge offer unique recognition sites for the interacting proteins. In addition, the S-turn displays closely spaced phosphate groups that could serve as a recognition site for sarcin that appears to employ a

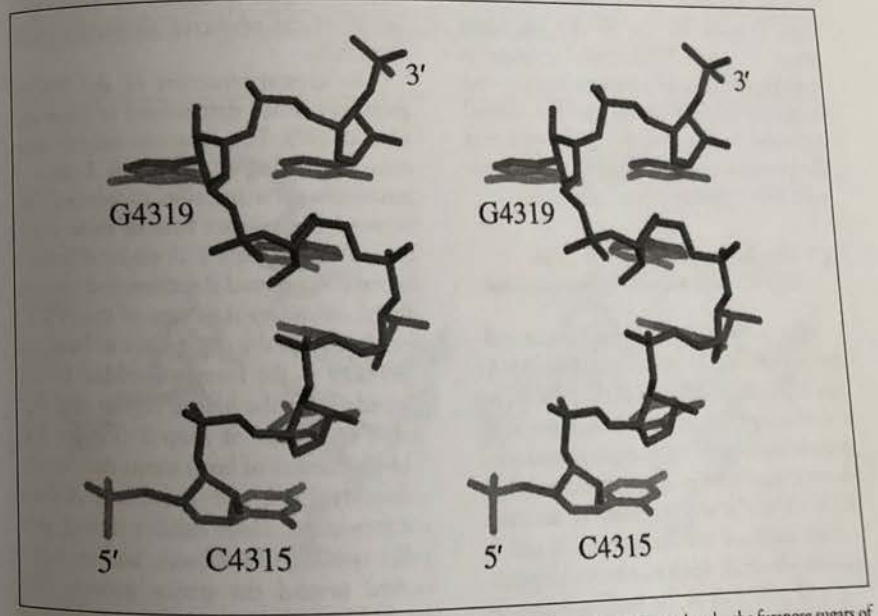


Figure 13. Stereo drawing of S-turn structure. Its main features are adoption of C2′-endo puckers by the furanose sugars of residues A4318 and G4319, flipping of the G4319 furanose ring such that its backside is exposed, thus forming a hydrophobic patch with the adjacent C5 and C8 atoms, and close spacing of the phosphate groups of U4316, A4317, and G4319.

ruler to select its cleavage site (Figure 13) (Reference 21). Another unusual property of the S-turn region is the formation of a hydrophobic patch near the bulged G4319 that arises as a consequence of the backbone distortions (Figure 13). In addition, the 2′-hydroxyl groups of residues A4318 and G4319 project into the major groove, in contrast to the situation in a canonical A-form helix where the 2′-hydroxyl groups are arranged in the minor groove. Thus, there are plenty of opportunities for proteins to specifically interact with the SRL rRNA region.

The ricin-SRL interaction mode differs from the sarcin one because the GAGA tetraloop that contains the depurination site is sufficient for ricin recognition and depurination. The base hydrogen bond donors and acceptors of the loop residues are exposed and can be easily probed by ricin. By comparison, the EF-G and EF-Tu binding sites comprise both the loop regions and the bulged G4319 near the S-turn. These factors might employ the specific recognition features of the S-turn described above. The SRL RNA crystal structure is a treasure chest of novel secondary and tertiary structural motifs and has offered a glimpse at potential recognition features for proteins known to selectively interact with this region of ribosomal RNA.

2.9 The beet western yellow virus (BWYV) frameshifting pseudoknot

Certain folding motifs within the coding regions of messenger RNA (mRNA) can direct the translation of the message by the ribosome. In many viruses, a so-called pseudoknot motif six to eight nucleotides downstream from a "slippery sequence" X XXY YYN (the reading frame in the consensus sequence is indicated and X and Y can be identical) leads to ribosomal pausing and a -1 shift of the reading frame to yield XXX YYY N, thus avoiding a stop codon (30,39). In a pseudoknot, a single-stranded region folds back to pair with residues in a hairpin loop. The stem at the 5′ end is called stem 1, the second one at the 3′ end is stem 2, and the single-stranded loops, loop 1 and loop 2, cross the major groove of stem 2 and the minor groove of stem 1, respectively (Figure 14) (References 77,109).

Pseudoknot-mediated ribosomal frameshifting is not restricted to animal and plant viruses, but has also been observed in yeast and bacterial systems (4). In the case of BWYV, a plant luteovirus, frameshifting regulates the production of an RNA-dependent RNA polymerase and alters the reading frame from G GGA AAC to GGG AAA C (36). Frameshifting in BWYV is pseudoknot-dependent and although pausing was believed to be one way by which frameshifting could be stimulated, it is not sufficient as tested by replacement of the pseudoknot with simple stem-hairpin loop structures (4). In addition, it was found that several luteoviruses display a conserved sequence AACAAA that is located in the loop 2 region of the predicted secondary structures (62).

The crystal structure of the BWYV pseudoknot was determined at 1.6 Å resolution (97). The structure reveals noncoaxial stacking at the stem 1–stem 2 junction and a rotation of around 50° between the bottom base of stem 2 and the top base of stem 1. A major difference between the crystal structure and the predicted secondary structure of the BWYV pseudoknot is the disruption of base pair A25-U13 in the former. Residue U13 is looped out of the helical region and A25 stacks onto A24 of loop 2 (Figure 14). The geometries of both stems deviate significantly from the canonical A-form conformation. Both exhibit strong propeller twisting of the bases, and stem 1 is curved toward the major groove. The gradual bend as well as the narrowing of its minor groove may be consequences of the formation of the extended minor-

groove RNA triplex between that stem and loop 2 (Figure 14). The placement of the loop 2 bases in the minor groove involves six layers of consecutive interactions and a contact between a loop 2 base and a 2′-OH of the stem 1 minor groove in each of them. Thus, the minor groove triplex observed here is RNA-specific. Particular conformational features of the minor groove triplex and the adjacent

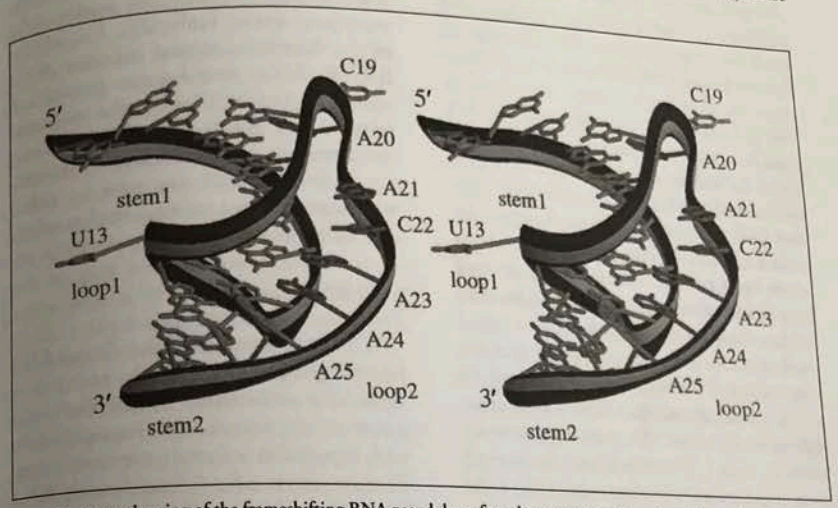


Figure 14. Stereo drawing of the frameshifting RNA pseudoknot from beet western yellow virus. The path of the ribose-phosphate backbone is traced by a ribbon. Residues of the third strand in the minor groove are highlighted and labeled, as are residues A25 and the bulged U13 at the junction between stem 1 and stem 2.

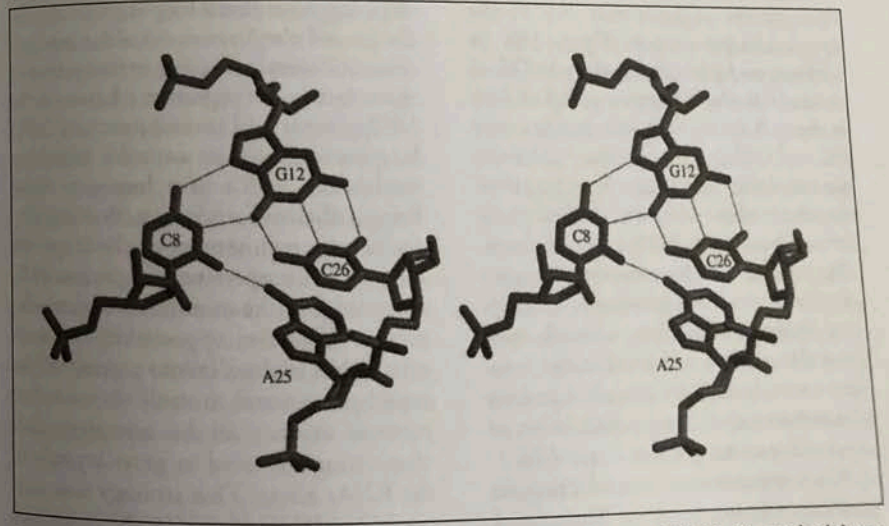


Figure 15. Stereo drawing of the quadruple-base interactions involving base pair G12-C26, C8 of loop 1 that is inserted into the major groove of stem 2, and A25, which does not pair with its predicted base mate U13. The particular arrangement of the C8 base relative to the G12–C26 base pair suggests protonation of its N3 nitrogen.

regions are the extruded orientation of G19, the first base of loop 2, and the anchor-like function of A20 that spans the minor groove of stem 1 and interacts with both strands under formation of seven hydrogen bonds. The next two residues, A21 and C22, are stacked and both form contacts to G16 in stem 1. Similarly, A23 interacts with C15 from one strand of stem 1 and contacts both the O2 and O2′ atoms through its N6 and N1 atoms, respectively. In the next layer, A24 shifts relative to A23 and now forms hydrogen bonds to the N2, N3, and O2′ atoms of G7 from the opposite strand of stem 1. This arrangement sheds some light on the conserved sequence of loop 2 in luteoviruses and is stabilized by no less than 16 hydrogen bonds.

A short loop 1, two nucleotides in the case of the BWYV, is a feature shared by many pseudoknots (97). The nucleotides C8 and A9 cross the major groove of stem 2, with C8 involved in a quadruple base interaction with base pair G12-C26 and A25 (Figure 15). Thus, C8 is inserted deep into the major groove and bridges the G12 and C26 bases with its Watson-Crick face. This particular arrangement suggests that N3 of the cytosine is protonated (Figure 15). In addition to hydrogen bonding to O2 of residue C8, the N6 amino group of A24 is engaged in two hydrogen bonds to the O2 and O2′ atoms of residue C14 from the next layer. This causes the slight tilt of the adenine base relative to the best plane defined by the C8, G12, and C26 bases. The position of C8 in the major groove of stem 2 and the formation of an intricate hydrogen bonding network may drive the rotation and bend at the junction between stem 1 and stem 2 and are also consistent with the conservation of this residue in many luteoviruses (62).

The mechanisms of frameshifting are presently not understood. The crystal structure of the BWYV pseudoknot has defined several unusual features of this particular RNA folding motif, including the minor groove triplex, bending and rotation at the junction between stem 1 and stem 2, the quadruple interaction of loop 1, the anchoring interaction of A20 from loop 2 in the stem 1 major groove, and stabilizing interactions mediated by numerous water molecules. Knowledge of the three-dimensional structure of the BWYV RNA pseudoknot provides a starting point for probing the structural features that trigger frameshifting. Thus, the functional significance of the observed interactions can now be tested with extensive mutational studies of the BWYV pseudoknot.

2.10 RNA aptamers

The coat protein of the *E. coli* RNA bacteriophage MS2 binds to a short stem-loop structure within the MS2 genome to control RNA encapsidation and repression of replicase translation. The structure of a 19-nucleotide RNA fragment soaked into MS2 capsid crystals revealed 13 ordered nucleotides, including the tetra-nucleotide loop and a bulged adenosine (105). The structural data suggested that a loop nucleotide and the looped out A are essential for recognition. In contrast, a recent investigation of the structures of complexes between the MS2 protein and several mutated RNA hairpins is consistent with the four loop nucleotides and a two base-pair stem being sufficient for binding, thus arguing against the requirement of the unpaired nucleotide for operator recognition (42). Knowledge of the minimum binding elements offers the opportunity to soak other RNA molecules into crystals of the capsid protein and to study their conformations as part of the complex, thus eliminating the need to grow crystals of the RNAs alone. This strategy was successfully used to determine the structures of small in vitro selected RNA motifs (19,86). However, the relatively small size

of the RNA molecules that can be soaked into the crystals and bound on the protein surface as well as the limited resolution of the structures (typically around 2.8 Å) are clear drawbacks of this system.

2.11 The 10–23 deoxyribozyme

The 10–23 deoxyribozyme emerged from an in vitro selection experiment aimed at identifying DNA molecules that could cleave a given target RNA under simulated physiological conditions (88). Other criteria that the DNA molecules had to meet were the ability to recognize the RNA substrate via Watson-Crick base pairing that would allow recognition of any sequence by changing the sequence of the recognition region, a reasonable catalytic efficiency that matched or surpassed that of ribozymes, and an overall size of less than 50 nucleotides. The name of the deoxyribozyme refers to the selection and cloning process: This particular DNA was obtained after 10 rounds of selection and corresponds to the 23rd clone from that round (88). The 10–23 enzyme displays a catalytic efficiency (k_{cat}/K_m) of 10^9 $M^{-1}min^{-1}$ and thus exceeds that of all other known catalytic nucleic acids. Unfortunately, its crystal structure did not allow any conclusions as to the structural origins of the cleavage specificity. Instead of the expected 1:1 complex between DNA enzyme and RNA substrate, two DNA strands and two RNA strands form an 82-nucleotide RNA-DNA complex in the crystal (74). Although the structure is unlikely to represent the active conformation, the observed four-way junction can serve as an interesting model for a Holliday junction that is formed as an intermediate during genetic recombination.

3. Conclusions and Outlook

A 20-year period with no advances beyond the crystal structures of transfer RNAs and their complexes with the corresponding aminoacyl synthetases has given way to a time of increased activity in the X-ray crystallographic structure determination of RNA over the last five years. The growing database of RNA structures includes four types of ribozymes, ribosomal RNA fragments, RNA domains that have key roles in the transcription and translation of viral genes, and a few RNA aptamers. Not only are there more crystal structures of RNA molecules, but the size of the successfully crystallized fragments has increased dramatically as well. Thus, the group I intron and the P4–P6 domain of the same molecule are both more than twice the size of transfer RNA. If the past years are indeed an indicator of the things to come, we can look forward to a flurry of interesting new structures of RNA molecules that exhibit diverse functions (72,99,114) and high degrees of complexity (7).

4. Techniques

4.1 RNA synthesis and purification

There are two basic methods to produce RNAs suitable for structure determination by X-ray crystallography. In vitro transcription using the DNA-dependent T7 RNA polymerase furnishes fragments of any length (63,112). For relatively short RNAs the method of choice is chemical synthesis, usually by the solid phase phosphoramidite technique (35). Thus far, chemical synthesis of oligoribonucleotides longer than 35 residues on a micromole scale was rather problematic and typically yielded limited amounts of material. The presence of the additional 2′-hydroxyl group in the furanose sugar renders more complicated chemical synthesis of RNA compared with DNA. Currently, protecting the 2′-OH moiety with the tertiary butyl dimethyl silyl

(TBDMS) group represents the most common strategy (90). This route is well suited for the synthesis of shorter oligonucleotides (110) and diffraction-quality crystals can be grown from the resulting material (25,56,104,107). An alternative strategy that makes use of the triisopropylsilyloxymethyl functionality (TOM) to protect the 2′-OH group was recently introduced (111). This new method allows production of RNAs as long as 100 residues and thus makes accessible to chemical synthesis a large number of biologically interesting RNA motifs.

After deprotection and cleavage from the solid support, RNA oligonucleotides are normally purified via trityl-on reverse phase HPLC (C-18 column) and ion-exchange chromatography. The purified sample is lyophilized prior to the preparation of an aqueous stock solution that is stored at -20°C. Column chromatography is not well suited for the purification of longer fragments (approximately 50 residues and more). Such samples are normally purified by denaturing polyacrylamide gel electrophoresis and are desalted following elution from the gel.

4.2 Crystallization

Many nucleic acid fragments can be crystallized in the presence of either magnesium chloride or polyamines (e.g., spermine hydrochloride), sodium cacodylate buffer, and 2-methyl-2,4-pentanediol (MPD) precipitant. These conditions were initially used to grow crystals of the yeast phenylalanine transfer RNA suitable for high-resolution diffraction data collection (45). Recently, these consensus nucleic acid crystallization conditions were compiled into a 24-condition crystallization screen (10). In addition to such standard conditions, a range of polyethylene glycols were shown to promote crystallization of short RNA fragments (6). Sparse matrix crystallization screens allow rapid sampling of a great variety of different conditions. Originally, such a strategy was used to facilitate the growth of protein crystals (47), but the same screen was shown to be useful for growing RNA crystals as well (25). Two sparse matrix crystallization screens were created specially for use with RNA fragments. The first is based on a compilation of the crystallization conditions that resulted in the growth of tRNA crystals and was successfully tested with a variety of RNA molecules ranging in length from 12 to 208 nucleotides (23). An alternative screen was created for growing crystals of a hammerhead ribozyme (91). The two screens differ considerably in their compositions, and the second employs mostly combinations of polyethylene glycols and monovalent salts. Despite all of these novel methods, 2-methyl-2,4-pentanediol (MPD) continues to be an extremely useful precipitant for the crystallization of nucleic acid molecules. Of the RNA motifs listed in Table 1, five were crystallized from MPD, three from high gradients of sulfate (lithium or ammonium), and for another three, polyethylene glycol (PEG 2000–8000) was the magic ingredient.

Although a large number of crystallization conditions can now be tested in a short time and with relatively small sample amounts, one should bear in mind that growth of diffraction-quality RNA crystals often requires that the length and sequence of the tested fragment be varied. Instead of simple "helix engineering", it may be advantageous to promote crystal growth by building regions into the crystallization construct that can mediate intermolecular contacts (79). This strategy was initially proposed and successfully applied in the case of the HDV ribozyme and a fragment of a group II intron RNA, both of which featured tetraloop and tetraloop receptor motifs (32). A further method to establish crystallization is the use of mutagenized RNAs that contain a protein binding site. For example,

an HDV ribozyme with a high-affinity binding site for the basic RNA binding domain (RBD) of the U1A spliceosomal protein was co-crystallized with the protein (33). Not only does the complex generate a host of novel lattice contacts compared with the RNA alone, but the protein portion can be of great help in the structure determination process (see section below). Another example is the co-crystallization of a series of RNA aptamers with the bacteriophage MS2 coat protein (19,86). Although some of the aptamers did not contain all the components of the MS2 consensus operator, binding occured in the same location on the protein. Thus, crystals of the MS2 protein could serve as a scaffold to carry out structure determinations of small nucleic acid recognition motifs without crystals of the RNAs alone.

4.3 Structure determination

There are four basic techniques used to solve the phase problem with crystals of macromolecules: Multiple Isomorphous Replacement (MIR), Multiwavelength Anomalous Diffraction (MAD), a combination of the two (MIRAS), Molecular Replacement (MR), and Direct Methods (DM). For a recent review of these techniques and their applications see References 12 and 13. The MR method is very powerful in cases where a good model structure is available. All of the RNA structures described in this chapter display novel tertiary structural motifs and it is therefore not surprising that none of them were determined by MR.

Direct Methods are completely model-independent, but require very high-resolution diffraction data (typically <1.2 Å). The resolution of the diffraction data for certain RNA crystals approaches that limit. For example, data with 100% completeness to a resolution of 1.25 Å are now available in our laboratory for crystals of the BWYV pseudoknot. Similarly, the crystal structure of the TAR RNA was reported at a resolution of 1.3 Å (46). However, determination of such structures via DM has not been reported thus far, although there have been numerous reports of the use of DM for determining the structures of peptide and protein crystals. Therefore, it appears that nucleic acid structures pose a bigger challenge for this technique; it is interesting that even crystals of oligonucleotides (8 mers to 12 mers), where data to 1 Å were collected, have so far resisted structure determination by the DM technique (Egli, M. and Weeks, C.M., personal communication). However, the structure of a chemically modified DNA decamer duplex now has been solved by DM based on 0.83 Å resolution data (Egli, M. and Weeks, C.M., personal communication).

While the initial crystal structures of $tRNA^{Phe}$ were solved by MIR, preceded by many protein structure determinations, the MAD technique now offers a powerful alternative and is rapidly gaining popularity. Application of the MIR technique requires at least two derivatized crystals that need to be highly isomorphous with the native ones. The search for heavy atom derivatives can be tedious and time consuming. However, the pioneering work on the crystal structures of the transfer RNA has led to the identification of a number of compounds that appear particularly suited for the preparation of derivatives, for example the lanthanides (45,53). Some of these continue to be very useful, in particular Sm(III) (chloride or acetate), which furnished derivative crystals for both a hammerhead ribozyme (78) and the *T. thermophila* group I intron (41). Another salt that has emerged as a very useful agent for derivatizing RNA crystals is osmium hexammine triflate. Osmium hexammine tends to replace weakly bound magnesium ions and was observed to have a preference for non-canonical 5′-GU-3′/3′-UG-5′ and 5′-GG-3′/3′-UU-5′ steps where it is

tucked between phosphate oxygens on one side and major groove hydrogen bond acceptors of G and U on the other (16). Osmium hexammine was used for derivative preparation with crystals of both the group I intron P4–P6 domain (14) and the BWYV pseudoknot (97). In cases where the RNA fragment was chemically synthesized, specific U or C residues can easily be replaced by 5Br-U or 5Br-C, respectively. This strategy was applied with several of the recently solved RNA structures, including two hammerhead ribozymes (78,92) (5I-U), the BWYV (97) (one derivative), the 5S rRNA domain (20) (one derivative), the SRL region of 28S rRNA (21) (two derivatives), the TAR RNA (46) (two derivatives), and the 10–23 deoxyribozyme (74) (5Br-T and 5I-T, six derivatives).

Contrary to MIR that is dependent on the availability of multiple derivatives, phasing with the MAD technique is possible, in principle, with a single one (44). However, since diffraction data need to be collected at multiple wavelengths this method requires a synchrotron source. The wavelength at synchrotrons can be tuned relatively easily, and with more synchrotron sources and beamlines dedicated to macromolecular crystallography coming on-line worldwide, MAD is on the way to becoming the method of choice for phasing. Several of the above mentioned elements are strong anomalous scatterers and have absorption edges in the wavelength range available at synchrotrons. In addition to bromine that was used for MAD phasing with a domain from 5S rRNA (20), the Sarcin/Ricin Loop (21), TAR RNA (46), Hg(II) [5S rRNA (20)], and Os(III) [P4–P6 domain of the group I intron (14) and BWYV pseudoknot (97)] were used as anomalous scatterers for MAD structure determination with RNA crystals. The most common anomalous scatterer to determine protein structure is selenium, which is incorporated into the protein in the form of Se-Met (22). Replacement of Met by Se-Met often furnishes highly isomorphous derivative crystals. In the case of the HDV ribozyme, which was crystallized in the presence of the RNA binding domain of U1A, derivative crystals were generated by incorporation of Se-Met into the protein portion (33).

References

1. **Aboul-Ela, F., J. Karn, and G. Varani.** 1995. The structure of the human immunodeficiency-virus type-1 TAR RNA reveals principles of RNA recognition by tat protein. J. Mol. Biol. *253*:313-332.
2. **Allain, F.H.T. and G. Varani.** 1995. Structure of the P1 helix from group-I self-splicing introns. J. Mol. Biol. *250*:333-353.
3. **Altman, S., L. Kirsebom, and S. Talbot.** 1993. Recent studies of ribonuclease P. FASEB J. *7*:7-14.
4. **Atkins, J.F. and R.F. Gesteland.** 1999. Intricacies of ribosomal frameshifting. Nature Struct. Biol. *6*:206-207.
5. **Bacher, J.M. and A.D. Ellington.** 1998. Nucleic acid selection as a tool for drug discovery. Drug Discovery Today *3*:265-273.
6. **Baeyens, K.J., J. Jancarik, and S.R. Holbrook.** 1994. Use of low-molecular-weight polyethylene glycol in the crystallization of RNA oligomers. Acta Crystallogr. D *50*:764-767.
7. **Ban, N., B. Freeborn, P. Nissen, P. Penczek, R.A. Grassucci, R. Sweet, J. Frank, P.B. Moore, and T.A. Steitz.** 1998. A 9 Ångstrom resolution X-ray crystallographic map of the large ribosomal subunit. Cell *93*:1105-1115.
8. **Basu, S., R.P. Rambo, J. Strauss-Soukup, J.H. Cate, A.R. Ferré-D'Amaré, S.A. Strobel, and J.A. Doudna.** 1998. A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. Nature Struct. Biol. *5*:986-992.
9. **Been, M.D. and G.S. Wickham.** 1997. Self-cleaving ribozymes of hepatitis delta virus RNA. Europ. J. Biochem. *247*:741-753.
10. **Berger, I., C.H. Kang, N. Sinha, M. Wolters, and A. Rich.** 1996. A highly efficient 24-condition matrix for the crystallization of nucleic acid fragments. Acta Crystallogr. D *52*:465-468.
11. **Biou, V., A. Yaremchuk, M. Tukalo, and S. Cusack.** 1994. The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with $tRNA^{Ser}$. Science *263*:1404-1410.
12. **Carter, Jr., C.W. and R.M. Sweet (Eds.).** 1997. Macromolecular crystallography, part A. Methods Enzymol. Vol. 276, Academic Press Inc., New York, NY.
13. **Carter, Jr., C.W. and R.M. Sweet (Eds.).** 1997. Macromolecular crystallography, part B. Methods Enzymol. Vol. 277, Academic Press Inc., New York, NY.
14. **Cate, J.H., A.R. Gooding, E. Podell, K. Zhou, B.L. Golden, C.E. Kundrot, T.R. Cech, and J.A. Doudna.** 1996. Crystal structure of a group I ribozyme domain: principles of RNA packing. Science *273*:1678-1685.
15. **Cate, J.H., A.R. Gooding, E. Podell, K. Zhou, B.L. Golden, A.A. Szewczak, C.E. Kundrot, T.R. Cech, and J.A. Doudna.** 1996. RNA tertiary structure mediation by adenosine platforms. Science *273*:1696-1699.
16. **Cate, J.H. and J.A. Doudna.** 1996. Metal-binding sites in the major groove of a large ribozyme domain. Structure *4*:1221-1229.
17. **Cate, J.H., R.L. Hanna, and J.A. Doudna.** 1997. A magnesium core at the heart of a ribozyme domain. Nature Struct. Biol. *4*:553-558.
18. **Cech, T.R.** 1990. Self-splicing of group I introns. Annu. Rev. Biochem. *59*:543-568.
19. **Convery, M.A., S. Rowsell, N.J. Stonehouse, A.D. Ellington, I. Hirao, J.B. Murray, D.S. Peabody, S.E.V. Phillips, and P.G. Stockley.** 1998. Crystal structure of an RNA aptamer-protein complex at 2.8 Å resolution. Nature Struct. Biol. *5*:133-139.
20. **Correll, C.C., B. Freeborn, P.B. Moore, and T.A. Steitz.** 1997. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. Cell *91*:705-712.
21. **Correll, C.C., A. Munishkin, Y.-L. Chan, Z. Ren, I.G. Wool, and T.A. Steitz.** 1998. Crystal structure of the ribosomal RNA domain essential for binding elongation factors. Proc. Natl. Acad. Sci. USA *95*:13436-13441.
22. **Doublié, S.** 1997. Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. *276*:523-530.
23. **Doudna, J.A., C. Grosshans, A. Gooding, and C.E. Kundrot.** 1993. Crystallization of ribozymes and small RNA motifs by a sparse matrix approach. Proc. Natl. Acad. Sci. USA *90*:7829-7833.
24. **Doudna, J.A.** 1998. The hammerhead swings into action. Curr. Biol. *8*:R495-R497.
25. **Egli, M., S. Portmann, D. Tracz, D. Workman, and N. Usman.** 1995. Crystallization and preliminary X-ray diffraction analysis of double-helical RNA octamers. Acta Crystallogr. D *51*:1065-1070.
26. **Egli, M. and I. Berger.** 1997. The role of backbone oxygen atoms in the organization of nucleic acid tertiary structure: Zippers, networks, clamps and C-H···O hydrogen bonds. Chem. Europ. J. *3*:1400-1404.
27. **Ekland, E.H., J.W. Szostak, and D.P. Bartel.** 1995. Structurally complex and highly active RNA ligases derived from random RNA sequences. Science *269*:364-370.
28. **Ekland, E.H. and D.P. Bartel.** 1996. RNA-catalyzed RNA polymerization using nucleoside triphosphates. Nature *382*:373-376.
29. **Endo, Y. and I.G. Wool.** 1982. The site of action of α-sarcin on eukaryotic ribosomes. J. Biol. Chem. *257*:9054-9060.
30. **Farabaugh, P.J.** 1996. Programmed translational frameshifting. Microbiol. Rev. *60*:103-134.
31. **Feig, A.L., W.G. Scott, and O.C. Uhlenbeck.** 1998. Inhibition of the hammerhead ribozyme cleavage reaction by the site-specific binding of Tb(III). Science *279*:81-84.
32. **Ferré-D'Amaré, A.R., K. Zhou, and J.M. Doudna.** 1998. A general module for RNA crystallization. J. Mol. Biol. *279*:621-631.
33. **Ferré-D'Amaré, A.R., K. Zhou, and J.M. Doudna.** 1998. Crystal structure of a hepatitis delta virus ribozyme. Nature *395*:567-574.
34. **Frankel, A.D.** 1994. Using peptides to study RNA-protein recognition, p. 221-247. *In* K. Nagai and I.W. Mattaj (Eds.), RNA-Protein Interactions. Oxford University Press, Oxford, UK.
35. **Gait, M. J. (Ed.).** 1984. Oligonucleotide Synthesis: A Practical Approach. IRL Press Ltd., Oxford, UK.
36. **Garcia, A., J. Vanduin, and C.W.A. Pleij.** 1993. Differential response to frameshifting signals in eukaryotic and prokaryotic translational systems. Nucleic Acids Res. *21*:401-406.
37. **Gaur, R.K. and G. Krupp.** 1993. Modification interference approach to detect ribose moieties important for the optimal activity of a ribozyme. Nucleic Acids Res. *21*:21-26.
38. **Gesteland, R.F. and J.F. Atkins (Eds.).** 1993. The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. **Gesteland, R.F. and J.F. Atkins.** 1996. Recoding: dynamic reprogramming of translation. Annu. Rev. Biochem. *65*:741-768.
40. **Gold, L., P. Allen, J. Binkley, D. Brown, D. Schneider, S.R. Eddy, C. Tuerk, L. Green, S. MacDougal, and D. Tasset.** 1993. RNA: the shape of things to come, p. 497-509. *In* R.F. Gesteland and J.F. Atkins (Eds.), The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
41. **Golden, B.L., A.R. Gooding, E.R. Podell, and T.R. Cech.** 1998. A preorganized active site in the crystal structure of the Tetrahymena ribozyme. Science *282*:259-264.
42. **Grahn, E., N.J. Stonehouse, J.B. Murray, S. van den Worm, K. Valegård, K. Fridborg, P.G. Stockley, and L. Liljas.** 1999. Crystallographic studies of RNA hairpins in complexes with recombinant MS2 capsids: Implications for binding requirements. RNA *5*:131-138.
43. **Hausner, T.P., J. Atmadja, and K.H. Nierhaus.** 1987. Evidence that the G2661 region of 23S rRNA is located at the ribosomal binding site of both elongation factors. Biochimie *69*:911-923.
44. **Helliwell, J.R.** 1992. Macromolecular Crystallography with Synchrotron Radiation. Cambridge University Press, Cambridge, UK.
45. **Holbrook, S.R. and S.-H. Kim.** 1985. Crystallization and heavy-atom derivatives of polynucleotides. Methods Enzymol. *114*:167-176.
46. **Ippolito, J.A. and T.A. Steitz.** 1998. A 1.3-Å resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal-ion dependent bulge conformation. Proc. Natl. Acad. Sci. USA *95*:9819-9824.
47. **Jancarik, J. and S.-H. Kim.** 1991. Sparse matrix sampling: a screening method for crystallization of

proteins. J. Appl. Crystallogr. *24*:409-411.
48.**Jenison, R.D., S.C. Gill, A. Pardi, and B. Polisky.** 1994. High-resolution molecular discrimination by RNA. Science *263*:1425-1429.
49.**Joyce, G.F.** 1994. In vitro evolution of nucleic acids. Curr. Opin. Struct. Biol. *4*:331-336.
50.**Jucker, F.M. and A. Pardi.** 1995. GNRA tetraloops make a U-turn. RNA *1*:219-222.
51.**Karn, J., M.J. Gait, M.J. Churcher, D.A. Mann, I. Mikaélian and C. Pritchard.** 1994. Control of human immunodeficiency virus gene expression by the RNA-binding proteins tat and rev, p. 192-220. *In* K. Nagai and I.W. Mattaj (Eds.), RNA-Protein Interactions. Oxford University Press, Oxford, UK.
52.**Kim, S.-H., F.L. Suddath, G.J. Quigley, A. McPherson, J.L. Sussman, A.H.-J. Wang, N. Seeman, and A. Rich.** 1974. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. Science *185*:435-440.
53.**Kim, S.-H., W.-C. Shin, and R.W. Warrant.** 1985. Heavy metal ion-nucleic acid interaction. Methods Enzymol. *114*:156-167.
54.**Lai, M.M.** 1995. The molecular biology of hepatitis delta virus. Annu. Rev. Biochem. *64*:259-286.
55.**Lehnert, V., L. Jaeger, F. Michel, and E. Westhof.** 1996. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. Chem. and Biol. *3*:993-1009.
56.**Lietzke, S.E., C.L. Barnes, and C.E. Kundrot.** 1995. Crystallization and structure determination of RNA. Curr. Opin. Struct. Biol. *5*:645-649.
57.**Lohse, P.A. and J.W. Szostak.** 1996. Ribozyme-catalyzed amino-acid transfer reactions. Nature *381*:442-444.
58.**Lorsch, J.R. and J.W. Szostak.** 1994. In vitro evolution of new ribozymes with polynucleotide kinase activity. Nature *371*:31-36.
59.**Marino, J.P., R.S. Gregorian, G. Csankovszki, and D.M. Crothers.** 1995. Bent helix formation between RNA hairpins with complementary loops. Science *268*:1448-1454.
60.**McKay, D.B.** 1996. Structure and function of the hammerhead ribozyme: an unfinished story. RNA *2*:395-403.
61.**Michel, F. and E. Westhof.** 1990. Modelling of three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. *216*:585-610.
62.**Miller, W.A., S.P. Dinesh-Kumar, and C.P. Paul.** 1995. Luteovirus gene expression. Crit. Rev. Plant Sci. *14*:179-211.
63.**Milligan, J.F. and O.C. Uhlenbeck.** 1989. Synthesis of small RNAs using T7 RNA-polymerase. Methods Enzymol. *180*:51-62.
64.**Moazed, D., J.M. Robertson, and H.F. Noller.** 1988. Interaction of elongation-factors EF-G and EF-TU with a conserved loop in 23S RNA. Nature *334*:362-364.
65.**Moras, D., M.B. Comarmond, J. Fischer, R. Weiss, and J.C. Thierry.** 1980. Crystal structure of yeast $tRNA^{Asp}$. Nature *288*:669-674.
66.**Munishkin, A. and I.G. Wool.** 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*:12280-12284.
67.**Murphy, F.L. and T.R. Cech.** 1993. An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. Biochemistry *32*:5291-5300.
68.**Murphy, F.L. and T.R. Cech.** 1994. GAAA tetraloop and conserved bulge stabilize the tertiary structure of a group-I intron domain. J. Mol. Biol. *236*:49-63.
69.**Murray, J.B., D.P. Terwey, L. Maloney, A. Karpeisky, N. Usman, L. Beigelman, and W.G. Scott.** 1998. The structural basis of hammerhead ribozyme self-cleavage. Cell *92*:665-673.
70.**Nagai, K. and I.W. Mattaj (Eds.).** 1994. RNA-Protein Interactions. Oxford University Press, Oxford, UK.
71.**Nitta, I., T. Ueda, and K. Watanabe.** 1998. Possible involvement of *Escherichia coli* 23S ribosomal RNA in peptide bond formation. RNA *4*:257-267.
72.**Nix, J.C., A.R. Newhoff, and C. Wilson.** 1999. Preliminary crystallographic characterization of an in vitro evolved biotin-binding RNA pseudoknot. Acta Crystallogr. D *55*:323-325.
73.**Noller, H.F., V. Hoffarth, and L. Zimniak.** 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. Science *256*:1416-1419.
74.**Nowakowski, J., P.J. Shim, G.S. Prasad, C.D. Stout, and G.F. Joyce.** 1999. Crystal structure of an 82-nucleotide RNA-DNA complex formed by the 10-23 DNA enzyme. Nature Struct. Biol. *6*:151-156.
75.**Pan, T. and O.C. Uhlenbeck.** 1992. A small metalloribozyme with a two-step mechanism. Nature *358*:560-563.
76.**Pan, T., B. Dichtl, and O.C. Uhlenbeck.** 1994. Properties of an in vitro selected Pb^{2+} cleavage motif. Biochemistry *33*:9561-9565.
77.**Pleij, C.W.A.** 1990. Pseudoknots: a new motif in the RNA game. Trends Biochem. Sci. *15*:143-147.
78.**Pley, H.W., K.M. Flaherty, and D.B. McKay.** 1994. Three-dimensional structure of a hammerhead ribozyme. Nature *372*:68-74.
79.**Pley, H.W., K.M. Flaherty, and D.B. McKay.** 1994. Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. Nature *372*:111-113.
80.**Puglisi, J.D., J.R. Wyatt, and I. Tinoco, Jr.** 1990. Conformation of an RNA pseudoknot. J. Mol. Biol. *214*:437-453.
81.**Puglisi, J.D., R.Y. Tan, B.J. Calnan, A.D. Frankel, and J.R. Williamson.** 1992. Conformation of the TAR RNA-arginine complex by NMR-spectroscopy. Science *257*:76-80.
82.**Puglisi, J.D., L. Chen, S. Blanchard, and A.D. Frankel.** 1995. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. Science *270*:1200-1203.
83.**Rich, A.** 1977. An overview of protein-nucleic acid interactions, p. 3-11. *In* Nucleic Acid-Protein

Recognition. Academic Press Inc., New York, NY.
84.**Robertus, J.D., J.E. Ladner, J.R. Finch, D. Rhodes, R.S. Brown, B.F. Clark, and A. Klug.** 1974. Structure of yeast phenylalanine tRNA at 3 Å resolution. Nature *250*:546-551.
85.**Rould, M.A., J.J. Perona, D. Söll, and T.A. Steitz.** 1989. Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. Science *246*:1135-1142.
86.**Rowsell, S., N.J. Stonehouse, M.A. Convery, C.J. Adams, A.D. Ellington, I. Hirao, D.S. Peabody, P.G. Stockley, and S.E.V. Phillips.** 1998. Crystal structures of a series of RNA aptamers complexed to the same protein target. Nature Struct. Biol. *5*:970-975.
87.**Ruff, M., S. Krishnaswamy, M. Boeglin, A. Poterszman, A. Mitschler, A. Podjarny, B. Rees, J.C. Thierry, and D. Moras.** 1991. Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$. Science *252*:1682-1689.
88.**Santoro, S.W. and G.F. Joyce.** 1997. A general purpose RNA-cleaving DNA enzyme. Proc. Natl. Acad. Sci. USA *94*:4262-4266.
89.**Sassanfar, M. and J.W. Szostak.** 1993. An RNA motif that binds ATP. Nature *364*:550-553.
90.**Scaringe, S.A., C. Francklyn, and N. Usman.** 1990. Chemical synthesis of biologically active oligoribonucleotides using β-cyanoethyl protected ribonucleoside phosphoramidites. Nucleic Acids Res. *18*:5433-5441.
91.**Scott, W.G., J.T. Finch, R. Grenfell, J. Fogg, T. Smith, M.J. Gait, and A. Klug.** 1995. Rapid crystallization of chemically synthesized hammerhead RNAs using a double screening procedure. J. Mol. Biol. *250*:327-332.
92.**Scott, W.G., J.T. Finch, and A. Klug.** 1995. The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell *81*:991-1002.
93.**Scott, W.G., J.B. Murray, J.R.P. Arnold, B.L. Stoddard, and A. Klug.** 1996. Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. Science *274*:2065-2069.
94.**Scott, W.G.** 1998. RNA catalysis. Curr. Opin. Struct. Biol. *8*:720-726.
95.**Söll, D. and U. RajBhandari (Eds.).** 1995. tRNA: Structure, biosynthesis and function. American Society for Microbiology, Washington, DC.
96.**Strobel, S.A., L. Ortoleva-Donnelly, S.P. Ryder, J.H. Cate, and E. Moncoeur.** 1998. Complementary sets of noncanonical base pairs mediate RNA helix packing in the group I intron active site. Nature Struct. Biol. *5*:60-65.
97.**Su, L., L. Chen, M. Egli, J.M. Berger, and A. Rich.** 1999. Minor groove RNA triplex in the crystal structure of a ribomomal frameshifting viral pseudoknot. Nature Struct. Biol. *6*:285-292.
98.**Sugimoto, N. and T. Ohmichi.** 1996. Site-specific cleavage reaction catalyzed by leadzyme is enhanced by combined effect of lead and rare earth ions. FEBS Lett. *393*:97-100.
99.**Sussman, D., J.C. Nix, and C. Wilson.** 2000. The structural basis for molecular recognition by the vitamin B_{12} RNA aptamer. Nature Struct. Biol. *7*:53-57.
100.**Symons, R.H.** 1992. Small catalytic RNAs. Annu. Rev. Biochem. *61*:641-671.
101.**Szewcak, A.A., L. Ortoleva-Donnelly, S.P. Ryder, E. Moncoeur, and S.A. Strobel.** 1998. A minor groove RNA triple helix within the catalytic core of a group I intron. Nature Struct. Biol. *5*:1037-1042.
102.**Szostak, J.W. and A.D. Ellington.** 1993. In vitro selection of functional RNA sequences, p. 511-533. *In* R.F. Gesteland and J.F. Atkins (Eds.), The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
103.**Uhlenbeck, O.C.** 1987. A small catalytic oligoribonucleotide. Nature *328*:596-600.
104.**Usman, N., M. Egli, and A. Rich.** 1992. Large scale chemical synthesis, purification and crystallization of RNA-DNA chimeras. Nucleic Acids Res. *20*:6695-6699.
105.**Valegård, K., J.B. Murray, P.G. Stockley, N.J. Stonehouse, and L. Liljas.** 1994. Crystal structure of an RNA bacteriophage coat protein-operator complex. Nature *371*:623-626.
106.**Verma, S., N.K. Vaish, and F. Eckstein.** 1997. Structure-function studies of the hammerhead ribozyme. Curr. Opin. Chem. Biol. *1*:532-536.
107.**Wahl, M.C., B. Ramakrishnan, C. Ban, X. Chen, and M. Sundaralingam.** 1996. RNA - Synthesis, purification and crystallization. Acta Crystallogr. D *52*:668-675.
108.**Wedekind, J.E. and D.B. McKay.** 1999. Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. Nature Struct. Biol. *6*:261-268.
109.**Westhof, E. and L. Jaeger.** 1992. RNA pseudoknots. Curr. Opin. Struct. Biol. *2*:327-333.
110.**Wincott, F., A. DiRenzo, C. Shaffer, S. Grimm, D. Tracz, C. Workman, D. Sweedler, C. Gonzalez, S. Scaringe, and N. Usman.** 1995. Synthesis, deprotection, analysis and purification of RNA and ribozymes. Nucleic Acids Res. *23*:2677-2684.
111.**Wu, X. and S. Pitsch.** 1998. Synthesis and pairing properties of oligoribonucleotide analogues containing a metal-binding site attached to β-*D*-allofuranosyl cytosine. Nucleic Acids Res. *26*:4315-4323.
112.**Wyatt, J.R., M. Chastain, and J.D. Puglisi.** 1991. Synthesis and purification of large amounts of RNA oligonucleotides. BioTechniques *11*:764-769.
113.**Ye, X., R.A. Kumar, and D.J. Patel.** 1995. Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. Chem. and Biol. *2*:827-840.
114.**Yusupov, M., P. Walter, R. Marquet, C. Ehresmann, B. Ehresmann, and P. Dumas.** 1999. Crystallization of the dimerization-initiation site of genomic HIV-1 RNA: preliminary crystallographic characterization. Acta Crystallogr. D *55*:281-284.