6.04
RNA Structures Determined by X-ray Crystallography

JENNIFER A. DOUDNA
Yale University, New Haven, CT, USA

and

JAMIE H. CATE
University of California, Santa Cruz, CA, USA

6.04.1 INTRODUCTION

Many RNA molecules have complex three-dimensional structures under physiological conditions, and the chemical basis for their functional properties cannot be understood unless these structures are known. In recent years it has become practical to determine RNA structures by X-ray crystallography, which can provide high-resolution information not only about RNA conformation but also about RNA interactions with ligands such as metal ions. The study of RNA by X-ray crystallography has become technically feasible due to the development of methods for producing

milligram quantities of virtually any RNA molecule, and for crystallizing RNA and producing heavy atom derivatives of RNA crystals. This chapter discusses these methods, and reviews the RNA crystal structures that are currently known.

6.04.2 CRYSTALLIZATION OF RNA

6.04.2.1 Production of Homogeneous RNA by *In Vitro* Transcription

In vitro transcription with bacteriophage T7 RNA polymerase is the method of choice for obtaining milligram quantities of RNA for crystallization. Its only drawback is that transcription with this enzyme results in molecules that are heterogeneous at their 3′-termini and, depending on template sequence, may also be heterogeneous at their 5′-termini.[1,2] Transcripts containing these extra residues cannot be removed from RNA preparations by preparative purification techniques when chain lengths exceed ~50 nucleotides.

Terminal heterogeneity can be removed from transcripts using ribozymes in *cis* and *trans* geometries. When included as part of an RNA transcript, hammerhead, hairpin, and hepatitis delta virus ribozyme sequences will self-cleave during or after transcription to produce RNA with defined termini. This method has been used by several groups to obtain RNA samples suitable for structure determination.[3,4]

6.04.2.2 Chemical Synthesis

Short oligoribonucleotides are often conveniently prepared using automated, solid-phase DNA synthesis machines. Chemically protected ribonucleoside phosphoramidites are sequentially coupled to a protected nucleoside attached at its 3′ end to a solid support such as controlled-pore glass or polystyrene. When synthesis of the sequence is complete, base hydrolysis is used to cleave its linkage to the solid support, releasing a 2′-*O*-silyl protected oligomer. Silyl protecting groups are removed using tetrabutylammonium fluoride (TBAF) or similar chemical reagents. Improvements in the 2′-OH protecting groups and deprotection methods, as well as development of effective oligonucleotide purification methods, have made the chemical synthesis of RNA oligonucleotides of up to 40–50 nucleotides routine (for more information, see Chapter 6.06). With extreme care, RNA oligomers of up to 80 nucleotides in length can be produced in milligram quantities by solid-phase synthesis.[5,6] In practice, however, chemical synthesis of RNA for crystallization is practical for oligonucleotides 30 nucleotides or less in length.

One advantage of this approach is that nucleotide analogues are readily incorporated into synthetic oligoribonucleotides at specific sites. This is useful for the production of heavy atom derivatives (see below) and for investigating ligand–RNA interactions. Enzymatic ligation of short synthetic RNAs and longer RNA molecules prepared by *in vitro* transcription can be used to produce chimeric molecules that contain modified bases at specific locations[7] (see Chapter 6.14).

6.04.2.3 Purification of RNA for Crystallization

Prior to crystallization experiments, contaminating salts and chemical reagents must be removed from RNA samples. This is usually accomplished using ion exchange or reverse phase chromatography, for RNA molecules up to ~40 nucleotides long, and by denaturing polyacrylamide gel electrophoresis for larger RNAs. Following purification, the RNA is dialyzed extensively into a low-salt buffer, and often it is then annealed by heating to 60–90 °C and slow cooling in the presence of 1–10 mM magnesium ion.

6.04.2.4 Establishing the Suitability of RNA Preparations for Crystallization

Whenever possible, purified RNA samples are tested for biological activity prior to crystallization. In the case of a tRNA, this might involve assaying for charging by tRNA synthetase, or, in the case of ribozymes, measuring catalytic activity. Once the activity of a sample is confirmed, it is tested for conformational homogeneity (polydispersity) using native polyacrylamide gel electrophoresis, size

exclusion chromatography, or dynamic light scattering.[8] The first two techniques can evaluate polydispersity only under low ionic strength conditions, while light scattering allows the determination of conformational homogeneity of RNA in solutions containing a variety of electrolytes and additives.

6.04.2.5 Sparse Matrix Approaches to RNA Crystallization

The crystallization of macromolecules is a trial and error process, and it is usually necessary to screen a wide range of conditions to find any that are conducive to crystal nucleation and growth. In the case of RNA, additional factors may complicate crystallization, such as the source and purity of material and the inherent instability of RNA. Furthermore, since some RNA molecules adopt several different conformations in solution, conditions that favor a single conformer must be found and used for crystallization.

To facilitate the search for crystallization conditions, sets of precipitating solutions have been developed that are biased towards conditions that have generated RNA crystals in the past.[9–11] These sets are applied to RNA using approaches based on the incomplete factorial and sparse matrix methods developed for protein crystallization. Satisfactory crystals of RNA duplexes, the hammerhead ribozyme, the P4–P6 domain of the *Tetrahymena* group I intron, and a fragment of 5S ribosomal RNA have all been obtained this way.

6.04.3 HEAVY ATOM DERIVATIVES OF RNA CRYSTALS

Once satisfactory crystals of a macromolecule are obtained, the phases for structure factors must be determined so that an electron density map can be calculated. For new structures this is usually achieved by making heavy atom derivatives of crystals, measuring diffraction intensities, and calculating phases based on the positions of the heavy atom(s).[12,13] Heavy atom derivatives of tRNA crystals were produced by soaking lanthanides into crystals, or by reacting crystals with osmium pyridine.[14] For the hammerhead ribozyme, crystal derivatives were prepared by lanthanide soaks and by covalent modification of the RNA with bromine. Covalent modification with bromine or iodine has also been used to solve the structures of short RNA duplexes and the loop E fragment of 5S ribosomal RNA. The crystal structure of the P4–P6 domain of the *Tetrahymena* ribozyme was solved by osmium hexammine substitution of magnesium binding sites in the major groove of the RNA.

6.04.4 DUPLEX STRUCTURES

The A-form helix is the structural unit from which complex, three-dimensional RNA structures are built. Isolated RNA helices often crystallize readily, and their structures can be solved using molecular replacement or covalent modification of the RNA. The high resolution (>2 Å) of some of the structures that have resulted has allowed a detailed look at metal ion binding sites, non-Watson–Crick base pairings, base bulging, helix packing in crystal lattices, and hydration.

6.04.4.1 Metal Ion Interactions in RNA Duplexes

Most structured RNAs require divalent metal ions for folding, and ribozymes generally need them for catalysis.[15] Several divalent ions have been located in the hammerhead ribozyme crystal structures (see below), but their functional significance remains unclear. One divalent metal ion binding site seen in hammerhead structures has also been found in a duplex containing sheared G·A and asymmetric A·A base pairs.[16] The site occurs at a C·G pair followed by the sheared G·A pair. Interestingly, tandem G·A mismatches have been found near the active sites of a lead-dependent ribozyme and an RNA ligase ribozyme, and they occur frequently in ribosomal RNA.[17–19] Thus, this motif may turn out to be a common way to position divalent metal ions within an RNA structure.

6.04.4.2 Noncanonical Base Pairs

Noncanonical or mismatch base pairs are common in RNA, and internal loops in rRNA often contain a high proportion of adenosines.[20] How are these nucleotides arranged, and how do they alter helical geometry? One example has been seen in a symmetric duplex, which includes a 5′-GAAA-3′ bulge surrounded by Watson–Crick pairs.[16] In this structure, tandem asymmetric A · A base pairs are sandwiched between sheared G · A pairs. Another common motif in rRNA involves tandem U · U pairs.[20] The three known duplex structures that contain this motif demonstrate that its structure varies depending on flanking sequences.[21–23] Two of these duplexes contain U · U wobble pairs,[21,23] but interestingly, the number of hydrogen bonds between the U · U pairs depends on the flanking base pairs. This result is consistent with effects seen in thermodynamic studies in solution,[24] but crystal packing forces may also affect the base pair geometry. In the third example, the U · U tandems form at the end of a duplex in an intermolecular contact.[22] These U · U tandems form unusual Hoogsteen pairs in which the N-3—H and O-4 of one uridine hydrogen bond to the O-4 and C-5—H of the other. It still is not clear whether tandem Hoogsteen U · U pairs like this can form in the middle of a duplex region, but they certainly might occur at the end of a helix. More importantly, the structure provides clear examples (at 1.4 Å resolution) of CH—O hydrogen bonds in base pairs and provides a model for U · Ψ base pairs in RNA.

6.04.4.3 RNA Packing and Hydration

RNA packing and hydration play important roles in RNA function, as highlighted in experiments involving large entropic contributions to ΔG.[25,26] While deceptively simple in form, the A-form helix can be greatly distorted, as seen in a structure of an RN–DNA chimeric duplex with a single looped-out adenosine.[27] In addition, the conformation of the extruded adenosine sheds some light on why the backbone of looped-out bases is often susceptible to magnesium-induced hydrolysis. Two high-resolution structures reveal in detail the pattern of hydration of G–C base pairs,[22,28,29] while a lower resolution structure sheds new light on the hydration of A–U pairs.[30,31] In these structures, the backbone plays key roles in the observed hydration patterns: the 2′-OH and the *pro*-R_p phosphate oxygen. As 2′-OH groups play important roles in RNA packing, exemplified by the ribose zipper (see Section 6.04.7), the heavy involvement of the 2′-OH group in hydration is a major factor to consider in thermodynamic studies of RNA–RNA interactions.

6.04.5 TRANSFER RNA

The first RNA molecule to be solved by X-ray crystallography that is large enough to have a tertiary structure was transfer RNA (tRNA). They were first because tRNAs are quite small and so abundant that they are readily purified from cells in adequate quantities. Modern methods for RNA production were not required. The crystal structure of $tRNA^{Phe}$, which was determined independently by three groups, became the basis for much of the RNA structural and functional biology that was done for the next 20 years.

tRNA plays a crucial role in protein biosynthesis. It is an adaptor molecule, one end of which interacts with amino acids and the other of which interacts with messenger RNA. Unlike normal double-stranded DNA, tRNA contains short helical elements interspersed with loops and its secondary structure is often drawn as a "cloverleaf" (Figure 1(a)). On the acceptor end, it carries an amino acid that corresponds to the genetic code triplet in its anticodon loop. The anticodon loop forms base pairs with messenger RNA on the ribosome, which then catalyzes peptide bond formation between the amino acid covalently bonded to one tRNA and the growing peptide chain covalently attached to a second one. Our understanding of tRNA structure and function has been reviewed in far more detail than is appropriate here;[32] however, some experiments regarding nucleotide modifications in tRNA deserve mention.

In the crystal structures of $tRNA^{Phe}$, the anticodon loop is ~70 Å away from the acceptor end of the molecule where the amino acid is attached. The four helical stems in that tRNA form an L-shaped molecule, each arm of which consists of a stack of two helices. A network of tertiary interactions between the D and T loops stabilize the assembly (Figure 1(b)),[33,34] a pattern that is conserved in other tRNA crystal structures.[35,36] While some tRNAs fold properly in the absence of divalent ions, Mg^{2+} stabilizes the tertiary structure of all of them, and the binding sites of some of the Mg^{2+} ions involved have been inferred from tRNA crystal structures.[37,38]

Although tRNAs synthesized *in vitro* from the four naturally occurring nucleotides are active,

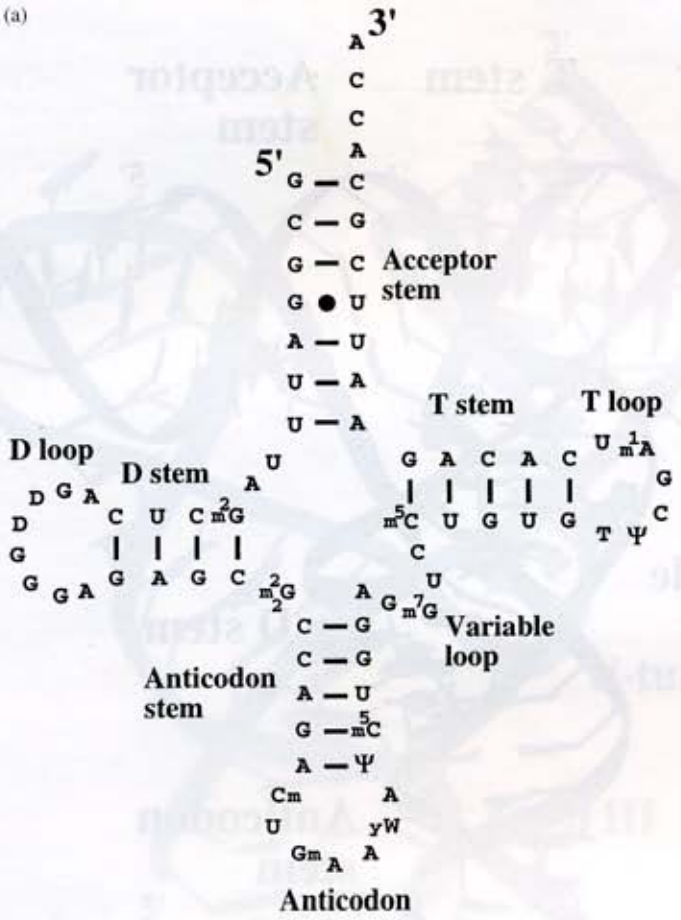

Figure 1 Structure of yeast phenylalanyl tRNA. (a) The secondary structure shows numerous naturally occurring modified nucleosides within the conserved tRNA fold. (b) Ribbon diagram of the crystal structure. Two sets of stacked helices form the L-shaped structure: the T-stem and the acceptor stem are in purple, and the D-stem and anticodon stem are in green.

tRNAs purified from cells always contain at least a few modified nucleotides. Modifications have been found in all of the stems and loops of tRNA.[32] One class of modifications, in the D and T loops, for example, optimizes the tertiary folding of tRNA.[33,34] Other modifications in the acceptor arm (acceptor stem and T stem) play a role in helping the protein-synthesizing machinery distinguish between initiator tRNAs and elongator tRNAs.[36] In the anticodon, modifications often play key roles in specifying codon recognition.[39]

Two advances in the chemical analysis of tRNAs have accelerated our understanding of these modifications. First, mass spectroscopic analysis of nucleotides in natural tRNAs has greatly expanded our knowledge of the kinds of modifications that occur in tRNA. In addition, the extent of modification in different types of organisms can be quickly assessed in the same way. For example, tRNAs from bacteria that grow in cold habitats have a higher abundance of dihydrouridine, which may increase conformational flexibility,[40] while thermophiles have modifications that may stabilize conformation.[41] Second, the chemical synthesis of tRNA-length RNAs may allow for milligram quantities of tRNA to be made that contain single-site modifications.[6] For example, both intra- and interhelical disulfide cross-links have already been incorporated into tRNA for the purpose of biophysical studies.[42] Combined with our increased knowledge of how tRNA interacts with amino-acyl tRNA synthetases and elongation factors, the ability to make designed modifications in tRNA, natural or otherwise, opens many new areas for biochemical and biophysical study.

6.04.6 THE HAMMERHEAD RIBOZYME

The hammerhead motif is a self-cleaving RNA sequence found in small RNAs that are plant pathogens. They make it possible for the multimeric genomes produced by rolling circle replication

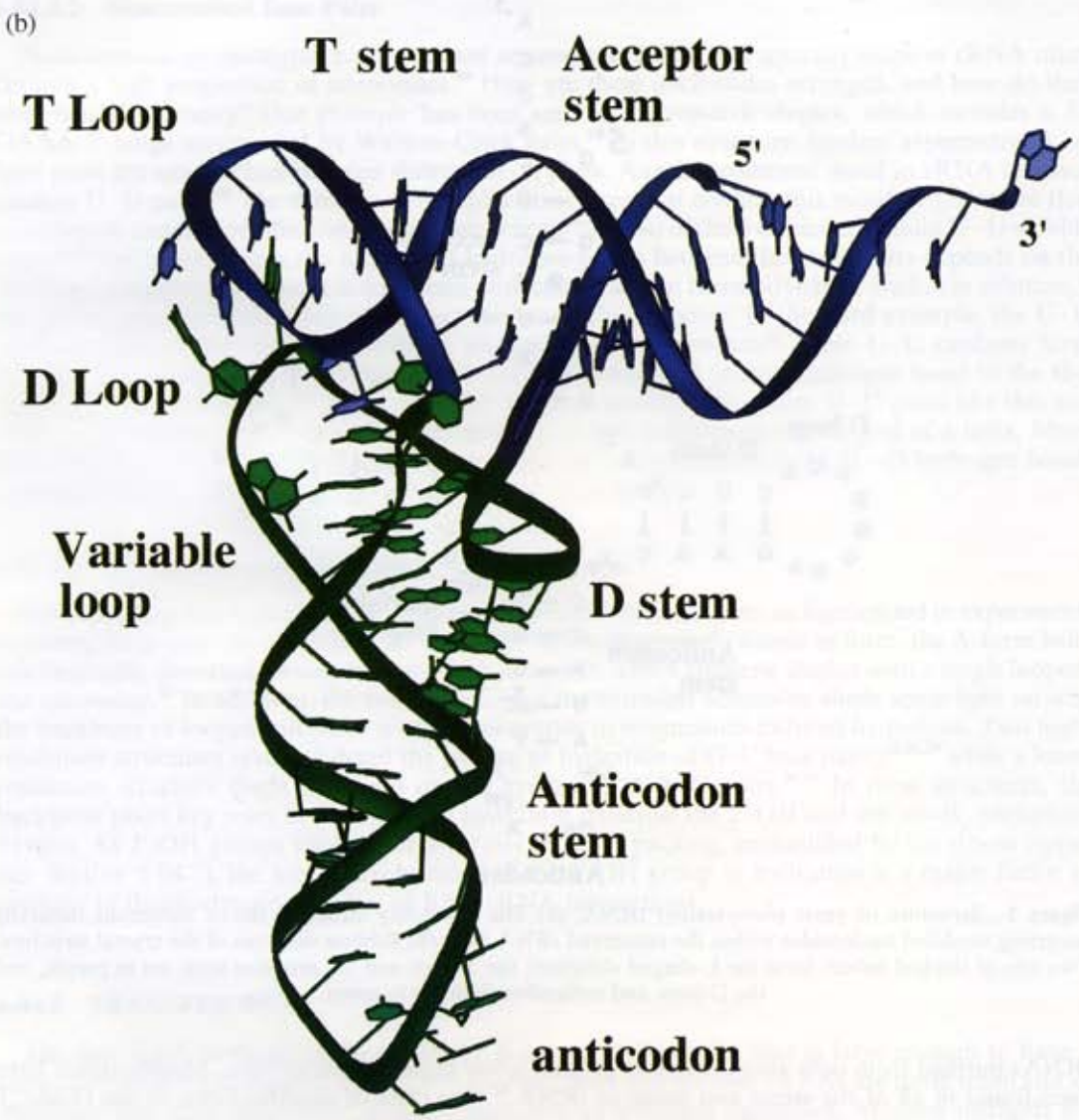

Figure 1 (continued)

to cleave into unit length molecules.[43] Unlike ribozymes such as self-splicing introns and the catalytic RNA subunit of ribonuclease P, the hammerhead domain is small, consisting of three helices that adjoin a core of phylogenetically conserved nucleotides (Figure 2). Cleavage occurs via nucleophilic attack of the 2′-hydroxyl of a specific nucleotide within the core on its adjacent phosphodiester bond to produce a 2′,3′-cyclic phosphate and a 5′-hydroxyl terminus.[44] Normally a single-turnover catalyst, the hammerhead is readily made into a multiple-turnover enzyme by separating the strand containing the cleavage site from the rest of the core.[44,45] Divided molecules like this have proven useful for crystallization because they allow replacement of the substrate strand with an all-DNA strand, or with an RNA strand modified at the cleavage site by a 2′-*O*-methyl group, neither of which can be cleaved. The crystal structures of these hammerhead–inhibitor complexes have revealed the overall geometry of the ribozyme but have raised almost as many new questions concerning the catalytic mechanism as they have answered.

In three dimensions, the hammerhead is shaped like a wishbone or γ, with stems I and II forming the arms, and stem III and the core forming the base.[46,47] This fold is seen in both of the inhibitor complexes solved so far despite differences in RNA backbone connectivities, substrate strand identities, crystallization conditions, and crystal packing. Whereas the three stems are all A-form helices, the structure of the central core is created, in part, by noncanonical pairings of the phylo-

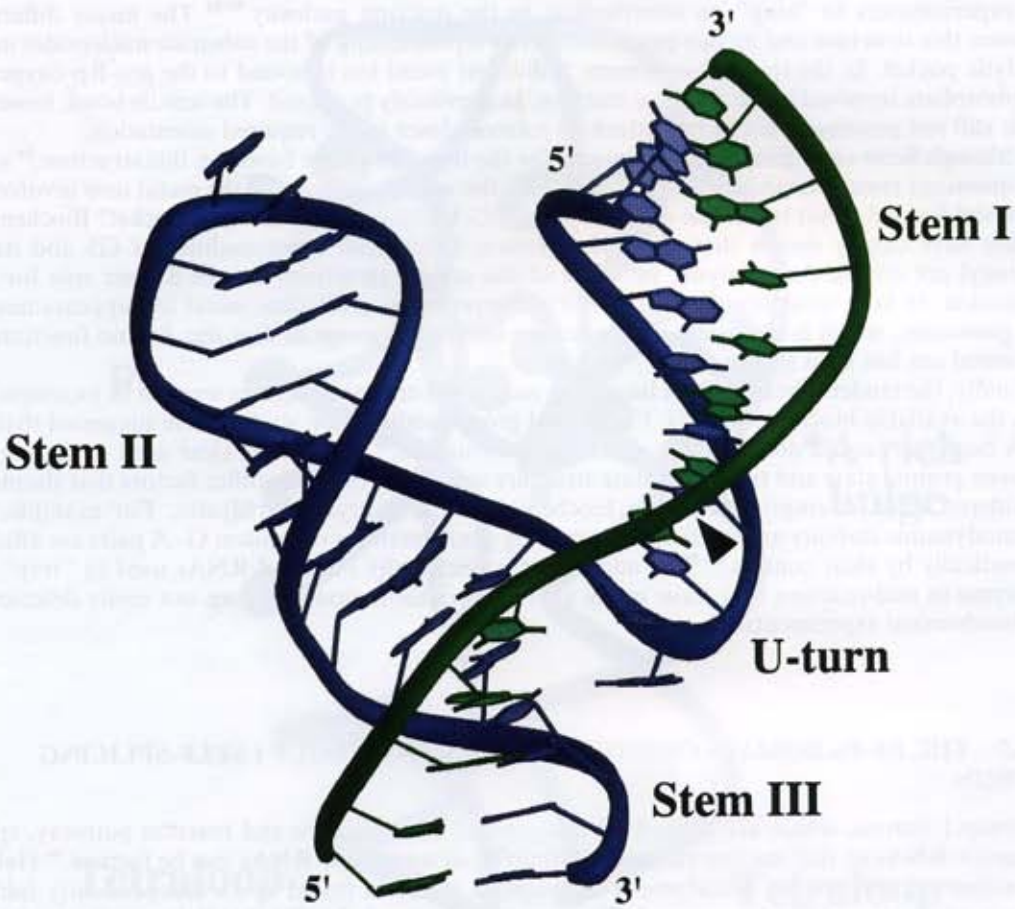

Figure 2 Crystal structure of the hammerhead ribozyme. The substrate strand is in green; the cleavage site is indicated by a black arrow within the U-turn.

genetically conserved nucleotides. Stems II and III sandwich two sheared G · A base pairs and an A–U base pair to form one long pseudocontinuous helix from which stem I and the catalytic site emanate. The highly conserved sequence CUGA between stems I and II forms a tight turn nearly identical in conformation to the uridine turn previously seen in the X-ray crystal structure of yeast phenylalanine transfer RNA.[34,48] The cytosine at the cleavage site between stems I and III is positioned near the CUGA cleft by interactions with the C and A of that sequence. This proximity led Klug and co-workers[47] to propose that the uridine turn, called domain I by McKay's group,[46] constitutes the catalytic pocket of the ribozyme.

Since the hammerhead-inhibitor complexes do not position the scissile bond correctly for the in-line nucleophilic attack that is believed to be part of the catalytic mechanism, these crystal structures probably represent the ground state of the ribozyme.[46,47] This has led to speculation that the hammerhead ribozyme may have to undergo a conformational change in order for cleavage to occur. Eckstein and co-workers[49] have used fluorescence resonance energy transfer (FRET) data to build a three-dimensional model of the hammerhead ribozyme that is similar to the X-ray models, except that the helical groove of stem I facing stem II differs. To distinguish between the solution and X-ray models, an elegant set of disulfide cross-linking experiments was carried out.[50] When stems I and II are cross-linked in conformations that exclude either the FRET or X-ray models, only the ribozyme cross-linked in a manner consistent with the X-ray structures is active. In addition, gel electrophoresis and transient electric birefringence have shown that the three stems are roughly co-planar and do not rearrange significantly after cleavage.[51,52] On the basis of these data, it is unlikely that the cleavage reaction requires a large change in conformation of the ribozyme.

More recent crystal structures of a hammerhead ribozyme complexed with a cleavable substrate have provided new insight into the rearrangements that occur in the catalytic pocket. Unlike the previous hammerhead ribozymes, the construct examined is active in the crystal lattice, allowing

the experimenters to "trap" an intermediate in the reaction pathway.[53,54] The major difference between this structure and its two predecessors is a repositioning of the substrate nucleotides in the catalytic pocket. In the trapped structures, a divalent metal ion is bound to the *pro*-Rp oxygen of the phosphate involved in the cleavage reaction, as previously proposed. The scissile bond, however, while still not positioned for in-line attack, is rotated closer to the required orientation.

Although Scott *et al.* propose a new model for the transition state based on this structure,[54] some key questions remain unanswered. First, what are the actual positions of the metal ions involved in catalysis? Second, what is the role of G5 in the CUGA U-turn in the catalytic pocket? Biochemical studies have clearly shown that all of the Watson–Crick base functionalities of G5 and its 2′-hydroxyl are critical for catalysis, yet none of the crystal structures reveals a clear role for this nucleotide. In later structures,[55] weak density interpreted as a divalent metal ion appears next to this guanosine, which is consistent with uranium-induced cleavage at that site, but no function for this metal ion has been shown.

Finally, the tandem sheared G · A base pairs seen in the crystal structures seem to be incompatible with the available biochemical data. Functional group modification studies have suggested that the G · A base pairs could not be in the sheared conformation.[56] Besides the clear need to distinguish between ground state and transition state structure stabilities, there are other factors that should be considered when attempting to relate biochemical data to crystal structures. For example, the thermodynamic stability and even the base pairing conformations of tandem G · A pairs are affected dramatically by their context.[57–59] In addition, the chemically modified RNAs used to "trap" the ribozyme in mid-reaction may have many alternate conformations that are not easily detected in the biochemical experiments.[60]

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

Group I introns, which are defined by a conserved catalytic core and reaction pathway, splice precursor RNAs so that mature ribosomal, transfer, or messenger RNAs can be formed.[61] Half of the conserved core in the *Tetrahymena thermophila* intron is found in an independently folding domain consisting of the base-paired (P) regions P4 through P6 (P4–P6).[62] By itself, the P4–P6 domain folds into a structure whose chemical protection pattern is very similar to that seen for the P4–P6 region of the intact intron.[62,63] The crystal structure of this domain, a 160-nucleotide RNA, has revealed several new aspects of RNA secondary and tertiary folding, and provides the first example of a kind of helical packing that is thought to occur in large ribozymes and RNA–protein complexes.[64]

In the 2.8 Å crystal structure of the P4–P6 domain, a sharp bend allows stacked helices of the conserved core to pack alongside helices of an extension (helices P5a, P5b, and P5c, or P5abc) that is important for folding and catalytic efficiency (Figure 3).[64,65] Two specific sets of tertiary interactions clamp the two halves of the domain together: an adenosine-rich corkscrew plugs into the minor groove of helix P4, and a GAAA tetraloop binds to a conserved 11-nucleotide internal loop,[66] termed the tetraloop receptor. The A-rich bulge coordinates two magnesium ions via its phosphate oxygens, allowing the backbone to invert and the bases to flip out. The adenosines make numerous tertiary contacts that connect the core helices to the helices in the P5abc extension. From biochemical evidence, these interactions are crucial to the stability of the entire domain.[62–64] The other half of the clamp, equally important to the packing of helices P5abc against the core (although not to the folding of the P5abc region itself), involves a GAAA tetraloop in the same conformation as seen previously.[67,68] The tetraloop receptor, a motif seen in many RNAs, has a widened minor groove that enables it to dock with the tetraloop in a highly specific manner.

The ribose 2′-hydroxyl group is involved in a common motif that occurs in both clamp interactions between the helical stacks. Pairs of riboses form an interhelical "ribose zipper"—a major component of the packing interactions (Figure 4). McKay and co-workers[68] also observed packing that involves pairs of 2′-hydroxyl contacts between a GAAA loop and the stem II minor groove of another hammerhead molecule in the crystal lattice. In a group II intron, riboses likely to be involved in a ribose zipper each contribute 2 kcal mol^{-1} of binding energy via their 2′-OH groups.[69] The number of ribose zippers seen so far suggests that this is a common way to pack RNA helices together.

One unexpected motif seen in the P4–P6 domain structure mediates both intramolecular and intermolecular interactions. At three separate locations in the 160-nucleotide domain, adjacent adenosines in the sequence lie side by side and form a pseudo-base pair within a helix (Figure 5).[70]

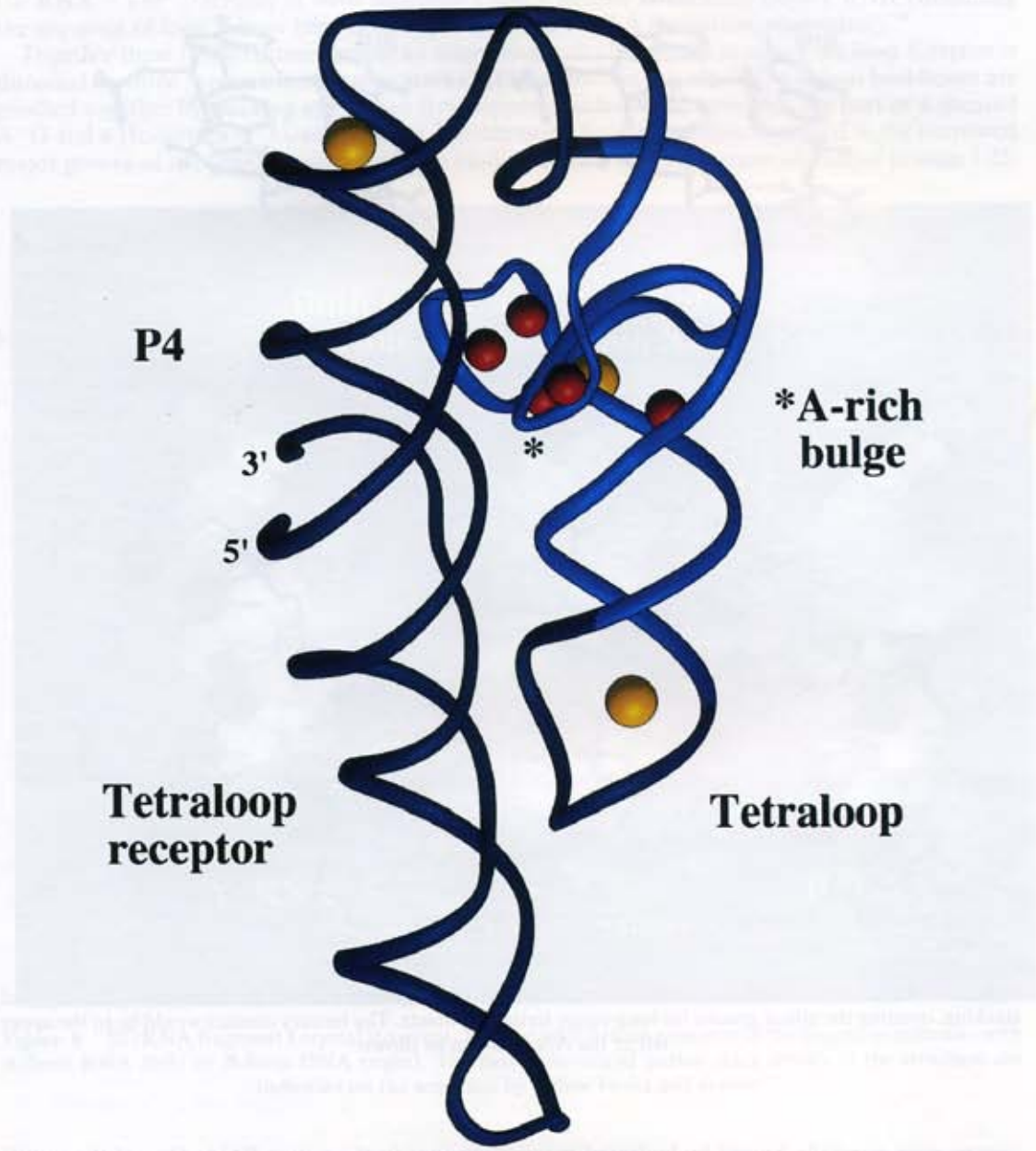

Figure 3 Crystal structure of the P4–P6 domain of the *Tetrahymena* group I intron. In red, magnesium ions in the metal ion core; in gold, osmium hexammine binding sites in the major groove; in light blue, the P5abc helices. The tertiary contacts between the helical stacks are indicated: the A-rich bulge docks into the minor groove of P4, and the tetraloop docks into the minor groove of the tetraloop receptor.

This AA platform opens the minor groove for base stacking or base pairing with nucleotides from a noncontiguous RNA strand.[70] The platform motif has a distinctive chemical modification signature which may make it possible to detect it in other RNAs chemically.[62,70] The ability of this motif to facilitate higher order folding provides at least one explanation for the abundance of adenosine residues in internal loops of many RNAs.

Many of the contacts that stabilize the P4–P6 domain structure, as well as its packing in the crystal lattice, involve the wide and shallow minor groove as opposed to the deep and narrow major groove. However, osmium hexammine, the compound used to determine the RNA structure by multiwavelength anomalous diffraction, binds at three locations in the major groove where non-standard base pairs create pockets of negative electrostatic potential.[71] In two cases, the heavy atoms

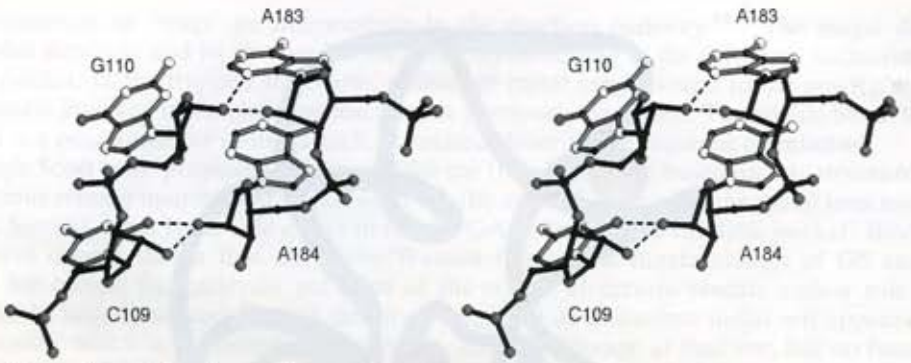

Figure 4 Stereoview of the ribose zipper. Pairs of nucleotides from different strands form a network of directed hydrogen bonds via 6 2′OH groups and the minor-groove functionalities of the bases.

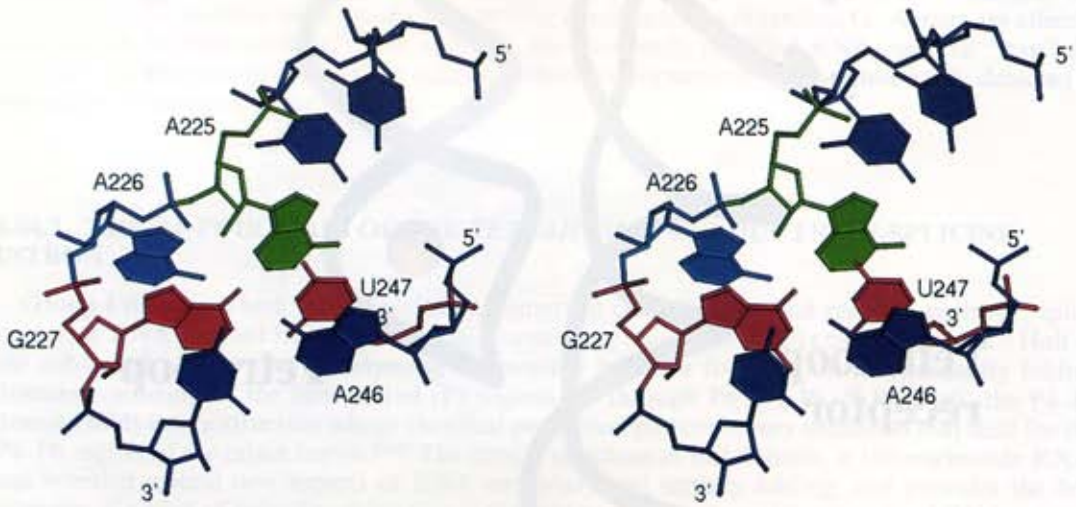

Figure 5 Stereoview of an AA platform. Viewed from the major groove, the AA platform allows cross-strand stacking, opening the minor groove for long-range tertiary contacts. The tertiary contact would be to the upper left of the AA platform as shown.

occupy sites normally bound by hydrated magnesium ions in the native RNA. One of the motifs involved, tandem G–U wobble pairs, occurs frequently in ribosomal RNAs,[72] suggesting a mechanism for metal binding in the ribosome.

6.04.8 5S RIBOSOMAL RNA FRAGMENT

E. coli ribosomal 5S RNA (5S rRNA), which contains 120 nucleotides, forms part of the 50S ribosomal subunit and binds three proteins—L25, L18, and L5. Like all structured RNAs, it has internal non-Watson–Crick base paired regions of loops. One of these, loop E, adopts its biologically functional structure only in the presence of millimolar magnesium ion concentrations. Mild nuclease digestion of 5S rRNA yields a 62 nucleotide fragment I that includes helices I and IV and loop E. The ribosomal protein L25 binds to both 5S rRNA and fragment I and protects helix IV and loop E from chemical modification.

Crystals of fragment I were originally obtained in 1983, but the structure proved difficult to solve due to limited diffraction resolution (~4 Å) and the lack of suitable heavy atom derivatives. Heavy atom derivatives were ultimately obtained by incorporation of chemically modified nucleotides into

the RNA.[73] The structures of both fragment I and a smaller dodecamer duplex RNA containing the sequence of loop E have been determined at 3 Å and 1.5 Å resolution, respectively.[74]

Together these two structures reveal an interesting helical molecule in which the loop E region is distorted by three "cross-strand purine stacks" (Figure 6).[74] In this motif, the helical backbones are pinched together by stacking of A bases from opposite sides of the helix that are part of a sheared A · G and a Hoogsteen A · U, respectively. Furthermore, four magnesium ions bind in the narrowed major groove of the helix, creating a unique binding surface for the cognate ribosomal protein L25.

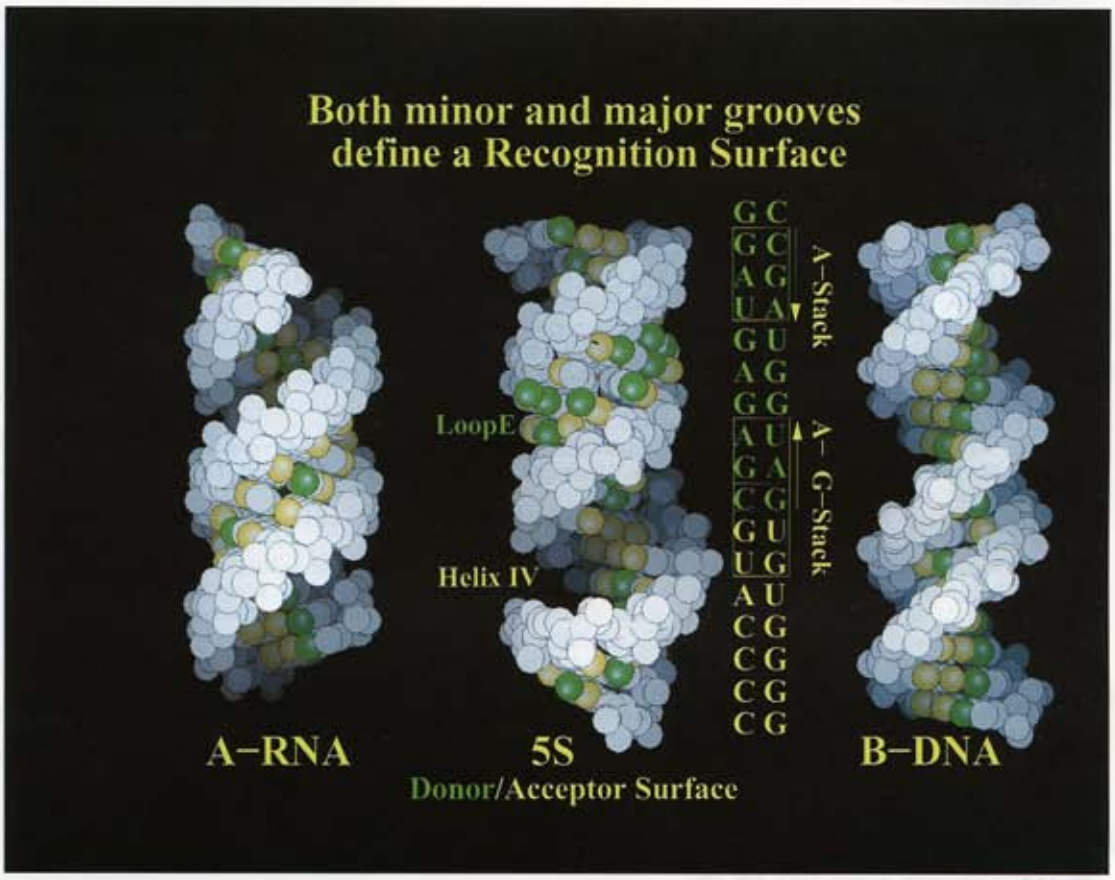

Figure 6 5S rRNA fragment I crystal structure. The irregular helical geometry of the fragment contrasts with A-form RNA (left) or B-form DNA (right). The two cross-strand purine stack motifs in the structure are indicated on the sequence by yellow boxes and arrows.

6.04.9 FUTURE DIRECTIONS

While the crystal structures of several RNAs provide new insights into RNA folding and catalysis, exciting challenges lie ahead. The hammerhead catalytic center is now the best understood of numerous ribozyme active sites; the others remain mysteries. The structures and roles of RNA in ribonucleoprotein particles including telomerase, signal recognition particle, the spliceosome, and the ribosome remain to be tackled by crystallographers. Chemical and biochemical experiments, critical for solving and understanding the hammerhead and P4–P6 domain structures, will be key to structural studies of these other RNAs also.

6.04.10 REFERENCES

1. J. F. Milligan, D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck, *Nucleic Acids Res.*, 1987, **15**, 8783.
2. J. F. Milligan and O. C. Uhlenbeck, *Methods Enzymol.*, 1989, **180**, 51.
3. S. R. Price, N. Ito, C. Oubridge, J. M. Avis, K. Nagai, *J. Mol. Biol.*, 1995, **249**, 398.

4. A. R. Ferre-D'Amare and J. A. Doudna, *Nucleic Acids Res.*, 1996, **24**, 977.
5. K. K. Ogilvie, N. Usman, K. Nicoghosian, and R. J. Cedergren, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 5764.
6. J. T. Goodwin, W. A. Stanick, and G. D. Glick, *J. Org. Chem.*, 1994, **59**, 7941.
7. M. J. Moore and P. A. Sharp, *Science*, 1992, **256**, 992.
8. A. R. Ferre-d'Amare and J. A. Doudna, *Methods Mol. Biol.*, 1997, **74**, 371.
9. J. A. Doudna, C. Grosshans, A. Gooding, and C. E. Kundrot, *Proc. Nat. Acad. Sci. USA*, 1993, **90**, 7829.
10. W. G. Scott, J. T. Finch, R. Grenfell, J. Fogg, T. Smith, M. J. Gait, and A. Klug, *J. Mol. Biol.*, 1995, **250**, 327.
11. B. L. Golden, E. R. Podell, A. R. Gooding, and T. R. Cech, *J. Mol. Biol.*, 1997, **270**, 711.
12. C. W. J. Carter and R. M. Sweet (eds.) "Macromolecular Crystallography Part A. Methods in Enzymology." Academic Press, New York, Vol. 276, 1997.
13. C. W. J. Carter and R. M. Sweet (eds.), "Macromolecular Crystallography Part B. Methods in Enzymology." Academic Press, New York, Vol. 277, 1997.
14. W. Saenger, "Principles of Nucleic Acid Structure," Springer, New York, 1984.
15. A. M. Pyle, *Science*, 1993, **261**, 709.
16. K. J. Baeyens, H. C. De Bondt, A. Pardi, and S. R. Holbrook, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 12 851.
17. E. H. Ekland, J. W. Szostak, and D. P. Bartel, *Science*, 1995, **269**, 364.
18. T. Pan and O. C. Uhlenbeck, *Nature*, 1992, **358**, 560.
19. D. Gautheret, D. Konings, and R. R. Gutell, *J. Mol. Biol.*, 1994, **242**, 1.
20. R. R. Gutell, N. Larsen, and C. R. Woese, *Microbiol. Rev.*, 1994, **58**, 10.
21. K. J. Baeyens, H. L. De Bondt, and S. R. Holbrook, *Nature Struct. Biol.*, 1995, **2**, 56.
22. M. C. Wahl, S. T. Rao, and M. Sundaralingam, *Nature Struct. Biol.*, 1996, **3**, 24.
23. S. E. Lietzke, C. L. Barnes, J. A. Berglund, and C. E. Kundrot, *Structure*, 1996, **4**, 917.
24. M. Wu, J. A. McDowell, and D. H. Turner, *Biochemistry*, 1995, **34**, 3204.
25. Y. Li, P. C. Bevilacqua, D. Mathews, and D. H. Turner, *Biochemistry*, 1995, **34**, 14 394.
26. T. S. McConnell and T. R. Cech, *Biochemistry*, 1995, **34**, 4056.
27. S. Portmann, S. Grimm, C. Workman, N. Usman, and M. Egli, *Chem. Biol.*, 1996, **3**, 173.
28. S. Portmann, N. Usman, and M. Egli, *Biochemistry*, 1995, **34**, 7569.
29. M. Egli, S. Portmann, and N. Usman, *Biochemistry*, 1996, **35**, 8489.
30. M. C. Wahl, C. Ban, S. Sekharudu, B. Ramakrishnan, and M. Sundaralingam, *Acta Crystallogr., Ser. D*, 1996, **52**, 655.
31. A. C. Dock-Bregeon, B. Chevrier, A. Podjarny, J. Johnson, J. S. Debear, G. R. Gough, P. T. Gilham, and D. Moras, *J. Mol. Biol.*, 1989, **209**, 459.
32. D. Söll and U. L. RajBhandary, (eds.), "tRNA: Structure, Biosynthesis and Function." ASM Press; Washington, DC, 1995.
33. F. L. Suddath, G. J. Quigley, A. McPherson, D. Sneden, J. J. Kim, S. H. Kim, and A. Rich, *Nature*, 1974, **248**, 20.
34. J. D. Robertus, J. E. Ladner, J. T. Finch, D. Rhodes, R. S. Brown, B. F. Clark, and A. Klug, *Nature*, 1974, **250**, 546.
35. D. Moras, M. B. Comarand, J. Fischer, R. Weiss, J. C. Thierry, J. P. Ebel, and R. Giege, *Nature*, 1980, **288**, 669.
36. R. Basavappa and P. B. Sigler, *EMBO J.*, 1991, **10**, 3105.
37. A. Jack, J. E. Ladner, D. Rhodes, R. S. Brown, and A. Klug, *J. Mol. Biol.*, 1977, **111**, 315.
38. S. R. Holbrook, J. L. Sussman, R. W. Warrant, G. H. Church, and S. H. Kim, *Nucleic Acids Res.*, 1977, **4**, 2811.
39. S. Yokoyama and S. Nishimura, in "tRNA: Structure, Function and Recognition," eds. D. Soll and U.L. RajBhandary, ASM Press, Washington, DC, 1995, p. 207.
40. J. J. Dalluge, T. Hamamoto, K. Horikoshi, R. Y. Morita, K. O. Steller, and J. A. McCloskey, *J. Bacteriol.*, 1997, **179**, 1918.
41. J. A. Kowalak, J. J. Dalluge, J. A. McCloskey, and K. O. Steller, *Biochemistry*, 1994, **33**, 7869.
42. J. T. Goodwin, S. E. Osborne, E. J. Scholle, and G. D. Glick, *J. Am. Chem. Soc.*, 1996, **118**, 5207.
43. G. A. Prody, J. T. Bakos, J. M. Buzayan, I. R. Schneider, and G. Bruening, *Science*, 1986, **231**, 1577.
44. O. C. Uhlenbeck, *Nature*, 1987, **328**, 596.
45. J. Haseloff and W. L. Gerlach, *Nature*, 1988, **334**, 585.
46. H. W. Pley, K. M. Flaherty, and D. B. McKay, *Nature*, 1994, **372**, 68.
47. W. G. Scott, J. T. Finch, and A. Klug, *Cell*, 1995, **81**, 991.
48. S. H. Kim, G. J. Quigley, F. L. Suddath, A. McPherson, D. Sneden, J. J. Kim, J. Weinzierl, and A. Rich, *Science*, 1973, **179**, 285.
49. T. Tuschl, C. Gohlke, T. M. Jovin, E. Westhof, and F. Eckstein, *Science*, 1994, **266**, 785.
50. S. T. Sigurdsson, T. Tuschl, and F. Eckstein, *RNA*, 1995, **1**, 575.
51. K. M. Amiri and P. J. Hagerman, *Biochemistry*, 1994, **33**, 13 172.
52. K. M. Amiri and P. J. Hagerman, *J. Mol. Biol.*, 1996, **261**, 125.
53. J. B. Murray, D. P. Terwey, L. Maloney, A. Karpeisky, N. Usman, L. Beigelman, and W. G. Scott, *Cell*, 1998, **92**, 665.
54. W. G. Scott, J. B. Murray, J. R. P. Arnold, B. L. Stoddard, and A. Klug, *Science*, 1996, **274**, 2065.
55. A. L. Feig, W. G. Scott, and O. C. Uhlenbeck, *Science*, 1998, **279**, 81.
56. T. Tuschl, M. M. P. Ng, W. Pieken, F. Benseler, and F. Eckstein, *Biochemistry*, 1993, **32**, 11 658.
57. J. SantaLucia, Jr. and D. H. Turner, *Biochemistry*, 1993, **32**, 12 612.
58. A. E. Walter, M. Wu, and D. H. Turner, *Biochemistry*, 1994, **33**, 11 349.
59. M. Wu and D. H. Turner, *Biochemistry*, 1996, **35**, 9677.
60. O. C. Uhlenbeck, *RNA*, 1995, **1**, 4.
61. T. R. Cech, *Annu. Rev. of Biochem.*, 1990, **59**, 543.
62. F. L. Murphy and T. R. Cech, *Biochemistry*, 1993, **32**, 5291.
63. F. L. Murphy and T. R. Cech, *J. Mol. Biol.*, 1994, **236**, 49.
64. J. H. Cate, A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, G. E. Kundrot, T. R. Cech, and J. A. Doudna, *Science*, 1996, **273**, 1678.
65. G. Van der Horst, A. Christian, and T. Inoue, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 184.
66. M. Costa and F. Michel, *EMBO J.*, 1995, **14**, 1276.
67. H. A. Heus and A. Pardi, *Science*, 1991, **253**, 191.
68. H. W. Pley, K. M. Flaherty, and D. B. McKay, *Nature*, 1994, **372**, 111.

69. D. L. Abramovitz, R. A. Friedman, and A. M. Pyle, *Science*, 1996, **271**, 1410.
70. J. H. Cate, A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, A. A. Szewczak, C. E. Kundrot, T. R. Cech, and J. A. Doudna, *Science*, 1996, **273**, 1696.
71. J. H. Cate and J. A. Doudna, *Structure*, 1996, **4**, 1221.
72. D. Gautheret, D. Konings, and R. R. Gutell, *RNA*, 1995, **1**, 807.
73. C. C. Correll, B. Freeborn, P. B. Moore, and T. A. Steitz, *J. Biomol. Struct. Dyn.*, 1997, **15**, 165.
74. C. C. Correll, B. Freeborn, P. B. Moore, and T. A. Steitz, *Cell*, 1997, **91**, 705.