Article No. jmbi.1999.3312 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (1999) 294, 829–849

JMB

REVIEW ARTICLE

# Stitching Together RNA Tertiary Architectures

**Thomas Hermann* and Dinshaw J. Patel***

*Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center New York, NY 10021, USA*

*Corresponding authors

The powerful explanatory paradigm of molecular biology requiring form to co-evolve with function has again been proven successful when, over the recent two decades, a wealth of biological functions have been uncovered for RNA. Previously considered as a mere mediator of the genetic code, RNA is now acknowledged as a key player in a wide variety of cellular processes. Along with the discovery of novel biological functions of RNA molecules, a number of RNA three-dimensional structures have been solved which beautifully demonstrate the molecular adaptability which allows RNA to participate as a key player in these functions. A distinct repertoire of molecular motifs provides a basis for the assembly of complex RNA tertiary architectures.



*Keywords:* metal ions and hydration; ribozymes and catalysis; RNA folding and function; RNA motifs; X-ray, NMR, and molecular modeling

## Introduction

The finding of numerous key cellular processes associated with RNA molecules asx active players has changed our view on nucleic acids, and it is now acknowledged that the diversity of RNA functionality is on a par with that of proteins (reviewed by Gesteland & Atkins, 1993; Gesteland *et al*., 1999). RNA is no longer considered solely as a passive transporter of the genetic code, but as an extremely versatile class of molecules actively participating in all steps of gene expression. RNAs that specifically recognize substrate molecules, e.g. aptamers (reviewed by Famulok, 1999; Patel, 1997), and catalyze chemical reactions, e.g. ribozymes (reviewed by Carola & Eckstein, 1999; Lilley, 1999; McKay & Wedekind, 1999), have been discovered both in nature and by *in vitro* selection.

The specific functions of RNA molecules are modulated by their distinct three-dimensional structures. While for over two decades regular helices of double-stranded RNA and the transfer-RNA (tRNA) fold had been the only known examples of RNA three-dimensional structures (reviewed by Saenger 1984), the last years have seen an explosion of novel RNA architectures. Technical improvements in RNA synthesis and structure determination methods have led to the solution of many three-dimensional structures of RNA molecules and complexes between RNA and both proteins and small molecules. Especially the crystal structure of the P4-P6 domain of self-splicing group I introns (Cate *et al*., 1996a) has revealed a treasure of novel RNA interaction motifs (Batey *et al*., 1999).

Even before detailed structural information on RNAs other than tRNA was available, the dominance of hierarchical principles in RNA folding (Brion & Westhof, 1997) has allowed the construction of large RNA structural models (Michel & Westhof, 1990; Lehnert *et al*., 1996) which agree extremely well with the available experimental data (Golden *et al*., 1998).

With a growing number of three-dimensional structures of RNA molecules at hand (reviewed by Masquida & Westhof, 1999; Nowakowski & Tinoco, 1999; Ferré-D'Amaré & Doudna, 1999), we have begun to understand the principles governing the architecture of RNA folds. Here, we discuss basic motifs of RNA structure involved in stabilizing RNA tertiary folds.

### The paradigm of RNA architectures: building handles into helices

In RNA folds, usually more than half of all nucleotides participate in standard Watson-Crick base-pairing. Consecutive stacking of such canonical base-pairs gives rise to *A*-form helices of double-stranded RNA. Close packing of double-stranded helices is the principle governing the basic architecture of all higher-order RNA folds which have been structurally elucidated so far (Figure 1).

Instead of major and minor grooves as in *B*-form double-stranded DNA, *A*-form helices have charac-

E-mail addresses of the corresponding authors: thermann@sbnmr1.ski.mskcc.org; pateld@mskcc.org

0022-2836/99/490829–21 $30.00/0

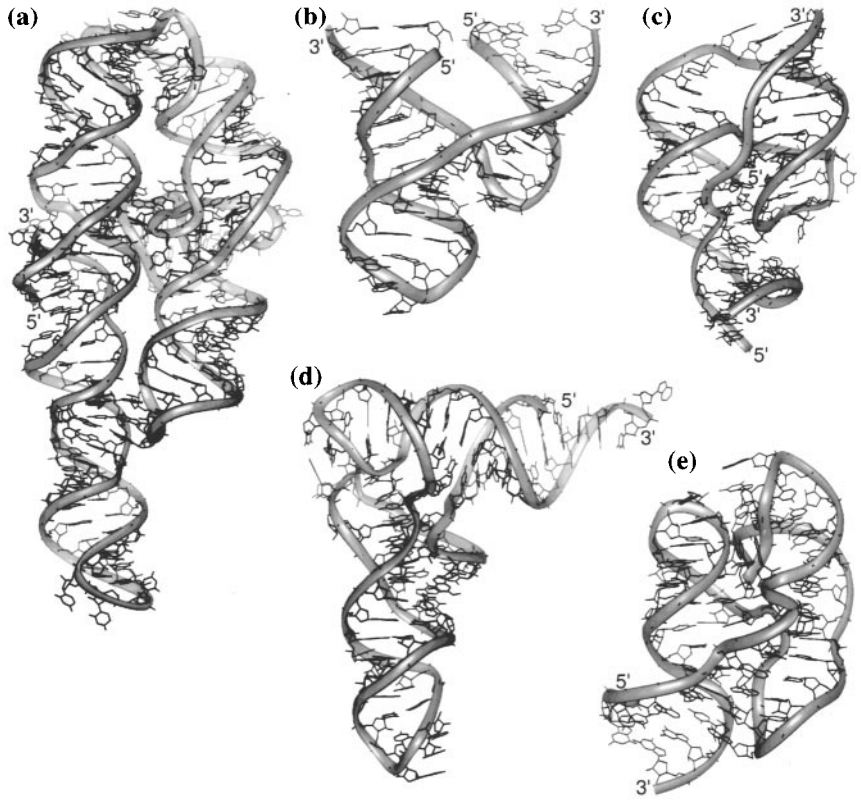


**Figure 1.** The arrangement of double-stranded helices governs the overall architecture of RNA folds. The three-dimensional structures of a number of complex RNA molecules are shown here in order to illustrate the packing of helices along with the intricate folding observed in RNA architectures. (a) The P4-P6 domain of the self-splicing group I intron from *Tetrahymena thermophila* (Cate *et al.*, 1996a). (b) The hammerhead ribozyme (Pley *et al.*, 1994a; Scott *et al.*, 1995). (c) The genomic ribozyme from hepatitis delta virus (HDV) (Ferré-D'Amaré *et al.*, 1998). (d) Phenylalanine-tRNA from yeast (Robertus *et al.*, 1974; Kim *et al.*, 1974). (e) The L11-binding domain of 23 S rRNA from *Escherichia coli* bound to the L11 protein (protein not shown) (Conn *et al.*, 1999; Wimberly *et al.*, 1999).

teristic deep and shallow grooves. Since the discriminatory edges of the bases are buried in the deep groove, regular *A*-form helices exert little potential for specific interactions with other domains (Weeks & Crothers, 1993). In order to build the intricate three-dimensional architectures observed in RNA folds (Westhof & Michel, 1998; Ferré-D'Amaré & Doudna, 1999), specific interaction sites are grafted into helices. Such sites comprise structurally conserved modules which have been found in the three-dimensional architectures of many different RNA molecules. Among these molecular modules are: (1) variations of the Watson-Crick base-pairing scheme, i.e. mismatches; (2) triples and quadruples of interacting bases; (3) platforms with pairing between consecutive bases within one strand; (4) bulged-out residues; (5) alternate cross-stacking between bases in different strands, i.e. ''interdigitation''; and (6) recurring hydrogen-bonding pattern between riboses of consecutive nucleotides in two strands, i.e. ''ribose zipper''.

### *Interactions in the plane: base-pairing in mismatches, triples, quadruples and platforms*

Hydrogen-bond interactions between sets of coplanar bases allow for a large number of mismatch base combinations beyond the canonical A:U, G:C Watson-Crick and G·U wobble pairs (Gautheret & Gutell, 1997). More than 25 years ago, the three-dimensional structure of tRNA, the first and most concisely analyzed RNA fold, revealed the geometry of a few non-canonical base interactions (Robertus *et al.*, 1974; Kim *et al.*, 1974). Since then mismatch pairs have been found widely in RNA architectures (Figure 2).

Inserted into RNA helices, mismatch pairs participate in stacking interactions (Holbrook *et al.*, 1991) and provide recognition sites both by presenting functional groups at the base edges and through distortion of the regular helical backbone. In the loop E of 5 S ribosomal RNA (rRNA), four purine-purine base-pairs interspersed in a regular *A*-form helix give rise to a widened deep groove and a unique hydrogen-bonding surface in the shallow groove permitting recognition by ribosomal proteins (Figure 2(b)) (Szewczak *et al.*, 1993; Szewczak & Moore, 1995; Correll *et al.*, 1997; Dallas & Moore, 1997). The extensive mismatch pairing in loop E is stabilized by metal ions and organized water molecules (Correll *et al.*, 1997) indicative of cations and solvent involvement in the maintenance of non-canonical base-pairs (see below) (Leontis & Westhof, 1998a). The loop E module itself has been found as a conserved motif in ribosomal RNAs, RNase P, the hairpin ribozyme and several self-splicing introns (Leontis & Westhof, 1998b).

Among homopurine base-pairs, G·A mismatches stand out as they are the most common non-canonical structural motifs in RNA molecules (Gautheret *et al.*, 1994). The sheared configuration of G·A pairs has been found as a conserved building block in the three-dimensional structures of many RNAs (Heus *et al.*, 1997), among them the GNRA loops (Heus & Pardi, 1991) (N is any nucleotide; R is a purine; see below), the GNRA-

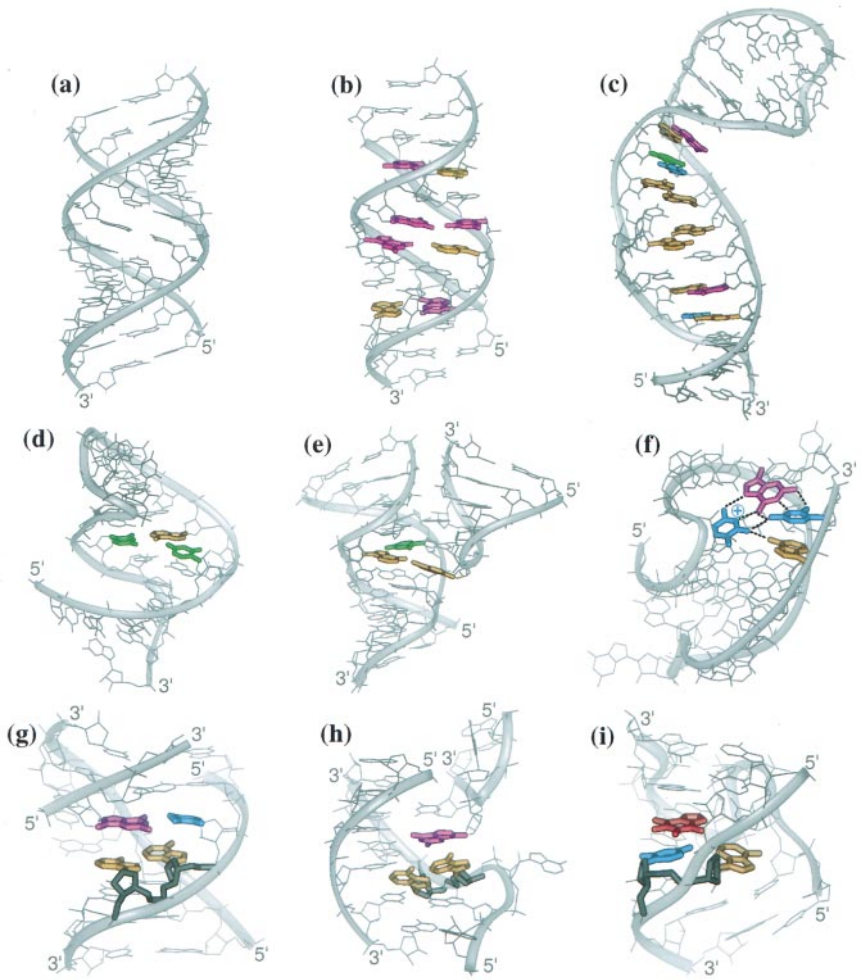


**Figure 2.** (a) Hydrogen-bonding interactions between coplanar bases yield Watson-Crick pairs, the building blocks of plain *A*-form RNA helices, and beyond, a wide variety of mismatches, multiples and platforms. (b) Mismatches dispersed among Watson-Crick pairs can give rise to helical stretches with distorted grooves, as observed in the loop E of 5 S ribosomal RNA (Correll *et al.*, 1997) and, (c) consecutively stacked, they can zipper up extended non-complementary regions, as in the loop B of the hairpin ribozyme (Butcher *et al.*, 1999). (d) Base triples widen RNA grooves for the interaction with proteins, as in the BIV TAR RNA (Ye *et al.*, 1995), and (e) they participate in the tertiary interaction between tetraloops and helical loop-receptors, as in the group I ribozymes (Cate *et al.*, 1996a). (f) A base quadruple has been found stabilizing the loop of an RNA pseudoknot (Su *et al.*, 1999). (g) Interactions between consecutive coplanar bases within one strand give rise to platforms, first observed for adenosine in group I ribozymes (Cate *et al.*, 1996b). Adenosine platforms provide on both sides stacking surfaces for other bases, as in (b) the group I ribozyme (G) and in the L11-binding domain of 23 S rRNA (Conn *et al.*, 1999), and (i) for planar substrates, as in the theophylline aptamer RNA (Zimmermann *et al.*, 1997). In this and all following Figures, bases are marked by color coding as follows: A, orange; G, magenta; U, green; C, blue. The theophylline substrate in (i) is in red.

like pentaloop of *box*B RNA in the complexes with peptides (Cai *et al.*, 1998; Legault *et al.*, 1998b), and the architectures of several ribozymes (Cate *et al.*, 1996a; Pley *et al.*, 1994a; Scott *et al.*, 1995; Butcher *et al.*, 1999) and ribosomal RNA fragments (Szewczak *et al.*, 1993; Szewczak & Moore, 1995; Correll *et al.*, 1998; Conn *et al.*, 1999; Wimberly *et al.*, 1999).

Double-stranded motifs other than *A*-form helices are created by consecutive stacking of mismatch pairs providing open structures as docking surfaces in RNA tertiary interactions which is illustrated by the loop B of the hairpin ribozyme (Figure 2(c)) (Butcher *et al.*, 1999). A peculiar case of consecutively stacked mismatches occurs in the crystal packing of the lead-dependent ribozyme (Wedekind & McKay, 1999), in which homo-purine pairs are formed between bases within two parallel-oriented strands making up a short parallel helix.

The variety of partners for mismatch base-pairing is extended by the potential of A and C nucleobases for pH-dependent protonation providing additional hydrogen donor sites, as have been observed, for example, in $C \cdot A^+$ (Cai & Tinoco, 1996; Hoogstraten *et al.*, 1998; Jang *et al.*, 1998; Pan *et al.*, 1998) and $G \cdot A^+$ mismatches (Pan *et al.*, 1999).

Watson-Crick pairs and mismatches can participate in triple and quadruple interactions in sets of coplanar bases. Base triples have already been found in the early crystal structures of tRNAs (Robertus *et al.*, 1974; Kim *et al.*, 1974). Depending on the topological connection of the strand which provides the third base docking to the pair, triples can lead to widened grooves within double-stranded architectures, or they mediate the tertiary interaction of a third strand. An RNA deep groove opened by a U·A:U triple is the common theme of the binding sites for argininamide in the human immunodeficiency virus type 2 (HIV-2) trans-activating region (TAR) RNA (Brodsky & Williamson, 1997), the bovine immunodeficiency virus (BIV) Tat-peptide in TAR RNA (Figure 2(d)) (Ye *et al.*, 1995; Puglisi *et al.*, 1995) and for the HIV-1 Rev-peptide in an RNA aptamer (Ye *et al.*, 1996). Docking of adenosine to the shallow-groove edge of a canonical G:C pair results in an A·C:G triple

which has been recurrently observed in strand junctions of RNA folds such as the P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a), hepatitis delta virus (HDV) ribozyme (Ferré-D'Amaré *et al.*, 1998), a turnip yellow mosaic virus pseudoknot (Kolk *et al.*, 1998), the L11-protein-binding domain of 23 S rRNA (Conn *et al.*, 1999; Wimberly *et al.*, 1999) and a ribosomal frameshifting pseudoknot (Su *et al.*, 1999). A particular strand junction mediated by triples is the tertiary interaction between a tetraloop and a tetraloop receptor seen in the crystal structures of hammerhead (Pley *et al.*, 1994b) and group I (Cate *et al.*, 1996a) ribozymes (Figure 2(e)). The tetraloop-receptor motif, first proposed from sequence comparisons and modeling (Michel & Westhof, 1990), is probably the most fundamental building block mediating RNA tertiary interactions (Abramovitz & Pyle, 1997) (see below).

Stacks of consecutive base triples give rise to triple helices. Depending on the direction of approach of the third base to the Watson-Crick pair, a distinction can be made between major and minor groove, respectively, deep and shallow groove triplexes. A shallow groove RNA triplex comprising six layers of triples has been discovered in the crystal structure of a ribosomal frameshifting pseudoknot (Su *et al.*, 1999), providing the first example of an extended triple interaction in the RNA shallow groove. Shorter shallow groove triplexes mediate the tertiary contacts of GNRA tetraloops with receptor duplexes (see below). Several RNA structures with the third strand in the deep groove are known (Klinck *et al.*, 1995; Holland & Hoffman, 1996), a feature also common to DNA triplexes (De los Santos *et al.*, 1989; Rajagopal & Feigon, 1989).

The GAAA tetraloop-receptor motif of the P4-P6 domain of group I ribozymes also provides an example of a G·A·C:G quadruple in which four bases in a plane associate by hydrogen bonding (Cate *et al.*, 1996a). An A·A·C·A quadruple has been proposed to bridge the P4-P6 and P3-P9 domains of group I ribozymes (Golden *et al.*, 1998). Due to the low resolution of the crystal structure of the intact group I ribozyme, however, a conclusive proof of this quadruple is still lacking. The only other quadruple in RNA proven so far is an $A \cdot C^{+} \cdot G{:}C$ interaction in the loop of pseudoknots in an HDV ribozyme (Ferré-D'Amaré *et al.*, 1998; Hilbers *et al.*, 1998) and a ribosomal frameshifting signal (Figure 2(f)) (Su *et al.*, 1999). This quadruple is organized by a protonated C which forms hydrogen bonds with each of the three other bases. In RNAs, the $A \cdot C^{+} \cdot G{:}C$ quadruple in the pseudoknots along with $C^{+} \cdot G{:}C$ triples in a mutant TAR RNA (Brodsky *et al.*, 1998) and in a tentative RNA triple helix (Klinck *et al.*, 1995) are the only examples for mismatches with protonated cytosine bases yet. Alignments involving protonated cytosine bases have been seen before in multi-stranded DNA structures (see, for example, De los Santos *et al.*, 1989; Rajagopal & Feigon, 1989; Gehring *et al.*, 1993).

Side-by-side base-pairing between consecutive nucleotides within one strand of RNA gives rise to platform motifs, first observed for vicinal adenosine bases in the P4-P6 domain of group I ribozymes (Figure 2(g)) (Cate *et al.*, 1996b). In adenosine platforms, the rise of the RNA backbone is reduced, thus creating a notch which presents the paired adenosine bases as a stacking surface for other bases. Stabilization of adenosine platforms requires the presence of monovalent metal ions, preferably potassium, which bind to a specific pocket immediately below the two adenosine bases (see below) (Basu *et al.*, 1998) where, in all adenosine platforms, a non-canonical base-pair is located. One of the three adenosine platforms in the P4-P6 domain participates in the formation of a tetraloop receptor motif (see below). An adenosine platform similar to the one in the P4-P6 domain has been found in the L11-binding domain of 23 S rRNA (Figure 2(h)) (Conn *et al.*, 1999; Wimberly *et al.*, 1999). In a theophylline aptamer RNA, the backbone of consecutive A and C residues adopts the conformation of the adenosine platform giving rise to an A-C platform which provides a stacking surface for the theophylline substrate (Figure 2(i)) (Zimmermann *et al.*, 1997). This function of the A-C platform resembles the role of a G·U:A triple serving as the substrate binding surface in an aptamer RNA for flavin mononucleotide (FMN) (Fan *et al.*, 1996).

### *Out of the plane: cross-strand stacking of bases*

The stacking interactions between the bases, mediated by their π-electron system, can be considered as the nucleic acid's analogue of the hydrophobic core in proteins. Stacking of bases, contributing significantly to the stability of RNA architectures, occurs predominantly between consecutive residues within one strand (reviewed by Saenger, 1984). In a number of RNA three-dimensional structures, cross-strand stacking of bases belonging to different strands is also observed (Figure 3(a)-(d)).

Again, the tRNA fold has provided the first example of cross-strand stacking. In the ''corner'' of the L-shaped tRNA, four interleaved purines participate in the junction of the T and D-loops (Figure 3(c)) (Robertus *et al.*, 1974; Kim *et al.*, 1974). Recently, when other examples of such stacks of interdigitated bases have been discovered in RNA architectures, the term ''base zipper'' was coined, yielding to the apparent need for a taxonomy of RNA structural motifs. Base zippers have been found in the structures of a theophylline aptamer RNA (Figure 3(d)) (Zimmermann *et al.*, 1997), a tetraloop receptor RNA (Butcher *et al.*, 1997) and the loop B of the hairpin ribozyme (Butcher *et al.*, 1999). Interdigitation contributes significantly to the stability and compactness of the three-dimensional structures of these RNAs, with the exception

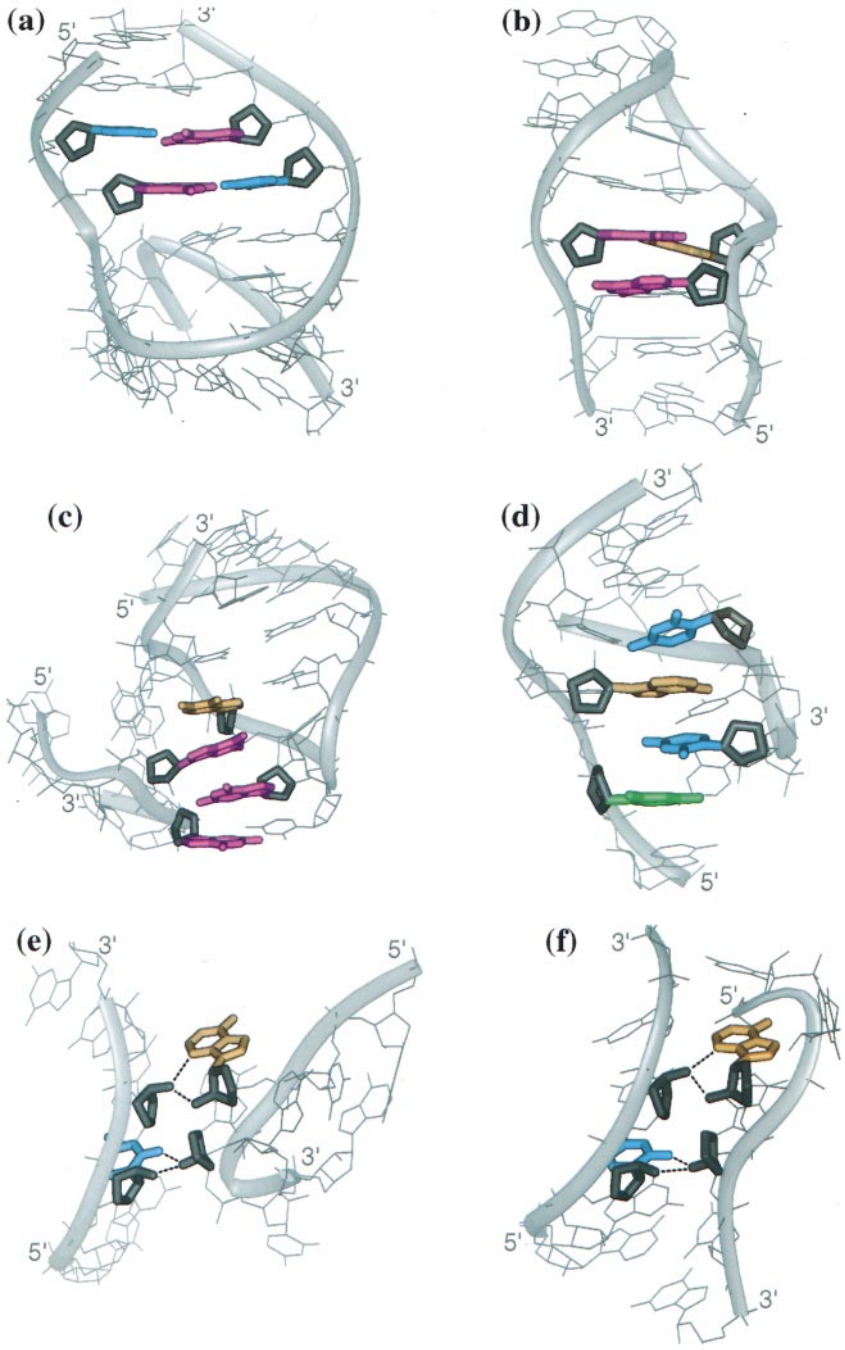


**Figure 3.** Tight interlocking of RNA strands is provided by (a)-(d) cross-strand stacking of bases and (e), (f) ribose zippers. (a) Cross-strand stacking of bases belonging to different strands is observed sometimes even for Watson-Crick pairs as in the *ColE1* RNA loop dimer where the guanosine bases of two consecutive canonical G:C pairs are stacking on top of each other (Lee & Crothers, 1998). (b) Extensive cross-strand stacking occurs between two guanosine bases one of which participates in a sheared G·A pair in the sarcin/ricin loop of 28 S rRNA (Szewczak & Moore, 1995; Correll *et al.*, 1998). (c) Interleaved cross-strand stacking gives rise to base zipper structures as in the junction of the T and D-loops in tRNA (Robertus *et al.*, 1974; Kim *et al.*, 1974) and (d) in the theophylline aptamer RNA (Zimmermann *et al.*, 1997). A regular hydrogen-bonding pattern between the 2′-OH groups of consecutive ribose moieties in the backbone of two different RNA strands is the signature of the ribose zipper structural motif. The local geometry of the ribose zippers found in the crystal structures of the (e) P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a) and of the (f) HDV ribozyme (Ferré-D'Amaré *et al.*, 1998) is strikingly conserved, in contrast to the dissimilar overall structure of the two strands which are stitched together.

of the tetraloop receptor in which the base zipper forms an open and exposed stretch. This, however, correlates well with the function of the receptor RNA to establish tertiary contacts with a tetraloop which forms hydrogen bonds with residues in the base zipper (see below). Upon docking of the tetraloop, a local conformational rearrangement has to take place in order to convert the base zipper, as in the free receptor, into an adenosine platform observed in the docked state (see below).

The cross-strand stacking of purines in the base zipper of the hairpin ribozyme B loop resembles the interactions observed in other RNA architectures. Cross-strand stacked purines occur in the sarcin/ricin loop of eucaryal 28 S rRNA (Figure 3(b)) (Szewczak *et al.*, 1993; Szewczak & Moore, 1995; Correll *et al.*, 1998), the loop E of 5 S rRNA (Wimberly *et al.*, 1993; Correll *et al.*, 1997; Dallas & Moore, 1997), the aptamer RNA for FMN (Fan *et al.*, 1996), the bulge-helix-bulge splice site RNA (Diener & Moore, 1998) and an RNA containing the binding site of signal recognition particle (Schmitz *et al.*, 1999). In both ribosomal RNA fragments, one of the guanosine bases involved in the cross-strand stacking participates also in a sheared G·A pair.

Even for canonical G:C base-pairs, cross-strand stacking between purines is possible, exceeding the

small purine overlap seen in regular YpR steps (where Y is a pyrimidine and R is a purine), as has been discovered in the structure of the *ColE1* inverted loop RNA dimer (Lee & Crothers, 1998). At the junction of an *A*-form helix and a loop involved in base-pairing with a second RNA monomer, the closing G:C pair of the helix is retracted towards the shallow groove so as to allow cross-strand stacking with an adjacent C:G pair in the helix (Figure 3(a)).

### *Backbone interactions: ribose zippers*

Regular hydrogen-bonding patterns occur between the bases and in some cases also between the backbone riboses in RNA architectures. In the densely packed core regions of RNA folds where the backbones of two strands come close, ribose zipper structures have been discovered which are held together by hydrogen bonds involving the 2′-OH groups of the sugar moieties (Figure 3(e) and (f)). This type of interaction is unique to RNA since DNA lacks 2′-OH groups.

Hydrogen bonding between 2′-OH groups of consecutive riboses was first noted in intermolecular crystal packing contacts of RNA molecules in a hammerhead ribozyme (Pley *et al*., 1994b) and a double-stranded RNA dodecamer (Schindelin *et al*., 1995). Later, intramolecular ribose zipper motifs have been found in the three-dimensional folds of the P4-P6 domain of group I ribozymes (Figure 3(e)) (Cate *et al*., 1996a), the HDV ribozyme (Figure 3(f)) (Ferré-D'Amaré *et al*., 1998) and the L11-binding domain of 23 S rRNA (Conn *et al*., 1999; Wimberly *et al*., 1999). A ribose zipper has been suggested connecting the two helical domains of the hairpin ribozyme (Earnshaw *et al*., 1997).

In all cases, the local geometry of the hydrogen-bonding network is conserved, despite the dissimilarity of the conformations of the RNA strands which are held together. The 2′-OH group of the ribose in one strand forms two hydrogen bonds to both the 2′-OH and either a pyrimidine O2 atom or a purine N3 atom of a nucleotide in the opposite strand. Two such ribose pairs make up the known ribose zippers; however, longer zippers comprising more pairs could be envisaged to exist in large RNA folds.

## Assembling modules into RNA architectures: packing and connecting helices

Double-stranded helices pack together in space to give rise to the intricate three-dimensional architectures observed in RNA folds (Strobel & Doudna, 1997). The helix modules are connected by distinctly structured single strands and tertiary contacts. The connecting regions themselves can form RNA structural motifs such as turns, loops, pseudoknots and loop-receptor complexes, which are found conserved in different RNA architectures.

### *Filling space with cylinders: stacking and docking of RNA helices*

Close packing of double-stranded helices which governs the architecture of RNA folds relies on two principles of assembling cylinders, namely end-to-end stacking and side-by-side docking. The preference of RNA helices for end-to-end stacking has been known for a long time as it dominates the arrangement of duplex fragments into long quasi-continuous helices in crystals (Holbrook & Kim, 1997). Likewise, helix stacking dominates the shape of RNA folds by defining the length of the principal helical modules. tRNA comprises four helices which, by pairwise stacking, assemble into the L-shaped tRNA fold (Robertus *et al*., 1974; Kim *et al*., 1974). Extensive stacking of helices occurs in the group I ribozymes, illustrated by crystal structures which reveal two pairs of coaxially stacked helices in the P4-P6 domain (Cate *et al*., 1996a) and no less than four such end-to-end stacked helices in the P3-P9 domain (Golden *et al*., 1998). In addition to tRNA and group I ribozymes, all other RNA folds for which the three-dimensional structures have been solved to date contain coaxial stacks of helices, namely one in the hammerhead ribozyme (Pley *et al*., 1994a; Scott *et al*., 1995), and two each in the HDV ribozyme (Ferré-D'Amaré *et al*., 1998) and the L11-binding domain of 23 S rRNA (Conn *et al*., 1999; Wimberly *et al*., 1999).

Whereas helix stacking mediated by the inherently flat surface of terminal base-pairs leaves little variety in the configuration of the resulting assembly, docking of helices *via* their sides comes in many ways. And so do the molecular handles that mediate helix docking in RNA architectures. In the P4-P6 domain of group I ribozymes, two helical stacks are docked side-by-side, locked in place by a number of RNA interaction motifs which are grafted into the helices (Cate *et al*., 1996a). Interactions between the unperturbed helical parts involve 2′-OH groups of the RNA backbone. Such 2′-OH-mediated contacts of helix sides have been seen in the crystal packing of double-stranded RNA helices (Holbrook *et al*., 1991; Schindelin *et al*., 1995; Baeyens *et al*., 1995; Lietzke *et al*., 1996).

Roughly parallel packing of helices occurs in the three-dimensional structures of the HDV ribozyme (Ferré-D'Amaré *et al*., 1998) and the L11-binding domain of 23 S rRNA (Conn *et al*., 1999; Wimberly *et al*., 1999) with extensive intermingling of helices and connecting regions. In the HDV ribozyme, compact packing is additionally enforced through formation of a pair of nested pseudoknots (see below) (Ferré-D'Amaré *et al*., 1998). The compactness of the HDV ribozyme is attested by the strong cohesion forces within the core of this RNA fold which allow for its catalytic activity even under highly denaturing conditions (reviewed by Been & Wickham, 1997).

### Single-stranded connections: bulges, turns and loops

The focus of the description of RNA structure has traditionally been the formation of base-pairs, drawing the line between base-paired double-stranded and unpaired single-stranded regions. In RNA architectures, unpaired regions can be as small as a single bulged nucleotide within a helix or as large as a long single-stranded stretch of nucleotides connecting distant helix ends. A wide variety of conformations of nucleotides in single-stranded regions is known, reflecting the diversity of three-dimensional structures of such loops and junctions. In some cases, however, the conformational pattern of adjacent nucleotides are conserved characterizing distinct classes of motifs such as tetraloops and turns (Figure 4).

Bulged-out residues can be found in helices and, given a wider definition, in loops and other compact RNA motifs. Single unpaired bases within an RNA double-strand are, depending on the structural context, either stacked into the helix (Figure 4(a)) (Borer *et al.*, 1995; Varani *et al.*, 1999) or looped-out with uninterrupted stacking of the flanking base-pairs (Figure 4(b)) (Portmann *et al.*, 1996). Intermolecular interactions affect the conformation of bulges, as has been observed for an unpaired adenosine which is looped out into a protein pocket in the complex between bacteriophage MS2 coat protein and a bulge-containing RNA hairpin (Valegard *et al.*, 1994; Convery *et al.*, 1998), but intercalated in the free RNA (Borer *et al.*, 1995).

Looping-out of single residues is frequently found in hairpin and internal loops which are zippered up into compact structures by stacking and formation of mismatch pairs (see above). Such ''bulges in a loop'' play key structural roles by introducing flexible hinges in the RNA backbone, providing recognition sites for proteins and other domains in large RNAs and acting as ''flaps'' which close over substrate binding cavities. Depending on their structural function, flipped-out residues are conserved when the base is involved in specific interactions, or they are non-conserved when the bulge serves as a mere source of backbone flexibility.

In an FMN aptamer RNA, a non-conserved bulged-out base within a mismatch-paired internal loop allows the formation of an adjacent base triple functioning as a stacking platform for the FMN substrate (Fan *et al.*, 1996). A uridine residue, flipped into the shallow groove, increases backbone flexibility contributing to the widening of the deep groove where Tat protein binds to BIV TAR RNA (Puglisi *et al.*, 1995; Ye *et al.*, 1995). A conserved single bulged adenosine base within a hairpin loop has been discussed as an interaction site for proteins or RNA tertiary contacts in spliced leader RNA (Greenbaum *et al.*, 1996). In the L11-binding domain of 23 S rRNA, conserved bulged-out residues introduce substantial distortions into regular RNA helices allowing intimate packing and tertiary interactions (Conn *et al.*, 1999; Wimberly *et al.*, 1999).

Larger RNA bulges with several flipped-out residues are seldom unstructured but display different mechanisms of conformational stabilisation such as intermolecular base-pairing, stacking and metal ion binding. In the bulge-helix-bulge splice site RNA (Diener & Moore, 1998) and the lead-dependent ribozyme (Wedekind *et al.*, 1999), three-nucleotide bulges are stabilized by extrahelical base stacking. The three flipped-out residues in the lead-dependent ribozyme form intermolecular base-pairs in the crystal packing and are involved in metal ion binding. The eminent role metal ions play in bulge stabilization, illustrated by extensive cation binding in the UCU bulge of HIV-1 TAR RNA (Ippolito & Steitz, 1998) and the A-rich bulge of the P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a) will be discussed below.

The reduced conformational restrictions in single-stranded regions of RNA allow for reversing the direction of the backbone. Several structural families of such turns are known, classified by the nucleotide preceding the turning phosphate (U and C-turns), or by the appearance of the backbone (S-turn) (Figure 4(c)-(e)). U-turns are stabilized by specific hydrogen bonding and base-phosphate stacking interactions across the bend (Figure 4(c)). Turn stabilization involves, however, only few functional groups of the bases, leaving free many specific sites for tertiary interactions.

U-turns, following a UNR consensus (N is any nucleotide; R is a purine), were first identified in the anticodon loop and T-loop in tRNA (Quigley & Rich, 1976). Later they have been found in other RNAs, such as the hammerhead ribozyme (Pley *et al.*, 1994a; Scott *et al.*, 1995), a hairpin loop (Huang *et al.*, 1996) and pseudoknots (Kolk *et al.*, 1998). The L11-binding domain of 23 S rRNA contains no less than three U-turns along with one S-turn (see below) within a stretch of 58 nucleotides, indicative for the compactness of this RNA fold (Conn *et al.*, 1999; Wimberly *et al.*, 1999). It has recently been pointed out that certain tetraloops (see below) form a distinct class of U-turns despite the turning phosphate belonging to a G in these loops (Jucker & Pardi, 1995). U-turn-like conformations have also been observed for the loop regions folding around the antibiotic substrate in the solution structures of different aminoglycoside RNA aptamers (Jiang *et al.*, 1997, 1999; Jiang & Patel, 1998). In the crystal structure of a ribosomal frameshifting pseudoknot, a C-turn has been described which shares some structural features with the U-turn (Su *et al.*, 1999).

The term ''reversed U-turn'' has been coined for a motif found in an isolated RNA hairpin of the HDV ribozyme (Kolk *et al.*, 1997). In the reversed U-turn, the U base directly follows the turning phosphate instead of preceding it, and the stacking between bases and phosphate groups is changed (Figure 4(d)).

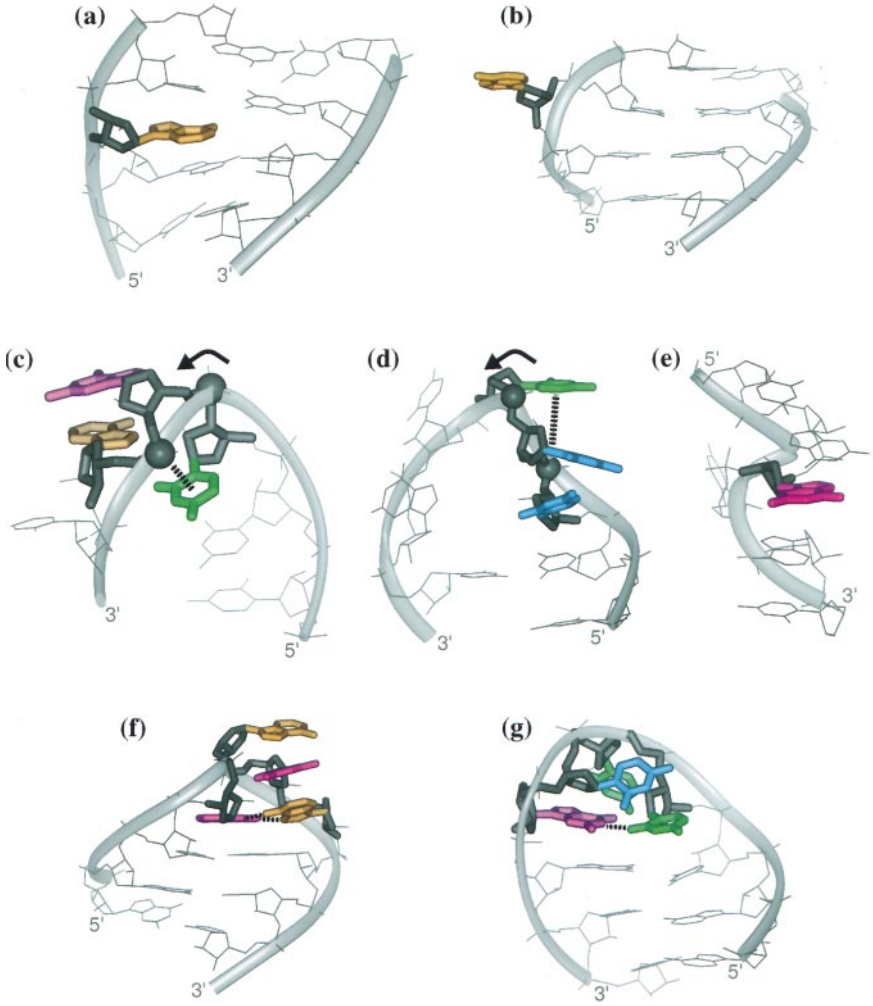


**Figure 4.** Unpaired regions in RNA comprise single residues, termed bulges, and whole stretches of several consecutive bases not involved in either Watson-Crick or non-canonical base-pairs. (a) Single unpaired bases can be stacked inside an RNA helix as has been found for an adenosine in the stem of a 24 nt hairpin loop (Borer *et al.*, 1995). (b) Alternatively, unpaired residues can loop out without interrupting the stacking of the flanking base-pairs as in the crystal structure of an *A*-form duplex containing an adenosine bulge (Portmann *et al.*, 1996). (c)-(e) Single-stranded regions comprising several nucleotides allow for turn formation causing abrupt changes in backbone direction. (c) In U-turns, as the one constituting part of the active site of the hammerhead ribozyme (Pley *et al.*, 1994a; Scott *et al.*, 1995), a sharp bend is introduced at a phosphate (arrow) following on a uridine base which stacks on the phosphate (dashes) of a third nucleotide. (d) In a hairpin of the HDV ribozyme, a reversed U-turn has been described in which the succession of turning (arrow) and stacking (dashes) phosphate is inverted, both preceding the stacking uridine base (Kolk *et al.*, 1997). (e) In S-turns, a locally reversed backbone is caused by a sugar moiety with an inverted pucker (C2′-*endo*) which introduces an S-shaped kink in the RNA strand as has been observed, for example, in the RRE RNA (Battiste *et al.*, 1996). (f) and (g) Tetraloops are the most frequently found single strands capping one end of a duplex. Among tetraloops, consensus sequences of GNRA, as in the (f) GAGA hairpin (Jucker *et al.*, 1996), and UNCG, as in the (g) UUCG P1 hairpin of group I ribozymes (Allain & Varani, 1995), stand out for their high stability due to base-pairing between the first and last residue and extensive stacking interactions. In (f) GNRA loops as well as in (c), (d) U-turns, several of the stacking bases in the turn are available for tertiary contacts with other RNA domains.

In contrast to U-turns which globally reverse the RNA strand direction, S-turns introduce a locally reversed backbone, maintaining an overall linear conformation of the RNA (Figure 4(e)). A flipped sugar moiety pointing with its 5′-end towards the 3′ direction of the RNA strand is central to S-turns. They have been found in several ribosomal RNA fragments (Szewczak *et al.*, 1993; Szewczak & Moore, 1995; Wimberly *et al.*, 1993, 1999; Correll *et al.*, 1998; Conn *et al.*, 1999), HIV Rev-response element (RRE) (Battiste *et al.*, 1996), a Rev-binding aptamer RNA (Ye *et al.*, 1996), and a theophylline aptamer (Zimmermann *et al.*, 1997).

Single-stranded regions connecting the two strands of a duplex stem are termed hairpin loops, constituting common building blocks of RNA architectures. Phylogenetic analyses have revealed that hairpin loops comprising four residues (tetra-

loops) are the most frequent loops in large RNAs such as self-splicing introns, rRNA, RNaseP and viral genomes (Tuerk *et al.*, 1988; James *et al.*, 1988; Woese *et al.*, 1990; Michel & Westhof, 1990). The evolutionary success of tetraloops has been attributed to their conformational stability and their ability to participate in tertiary contacts with other RNA motifs (Michel & Westhof, 1990, 1996; Abramovitz & Pyle, 1997). The most extensively studied GNRA (N is any nucleotide, R is A or G) and UNCG tetraloops achieve their stability through the formation of closing sheared G·A and G·U pairs, stacking interactions of the two remaining unpaired residues and a network of hydrogen bonds (Figure 4(f) and (g)) (Cheong *et al.*, 1990; Heus & Pardi, 1991; Jucker *et al.*, 1996). The extraordinary rigidity of GNRA tetraloops is attested by the finding that the conformation of a GAAA loop in an RNA hairpin observed by NMR (Heus & Pardi, 1991) is virtually identical with that found in the crystal structure of the P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a). Larger loops may adopt the favourable GNRA conformation by excluding the additional nucleotides from the loop scaffold. The extrusion of one base has been seen in the pentanucleotide loop of the *boxB* RNA hairpin which forms a GNRA-like structure in the complexes between the RNA and phage N proteins (Cai *et al.*, 1998; Legault *et al.*, 1998b). The excluded base is available for interactions with proteins.

In larger hairpins, a considerable content of purines often allows extensive mismatch base interactions leaving only few unpaired residues. Examples are provided by the sarcin/ricin loop of 28 S rRNA (Szewczak *et al.*, 1993; Szewczak & Moore, 1995) and the spliced leader RNA hairpin in which only three of 11 single-stranded loop residues are pointing out in the solvent while eight bases are involved in stacked mismatch pairs (Greenbaum *et al.*, 1996).

Unpaired residues may be non-conserved, such as the second nucleotide in GNRA tetraloops, indicating a role of the sugar backbone rather than of the base for the loop structure. In other hairpin loop fragments, however, unpaired residues without base contacts inside the loop are conserved, pointing to a participation of the loop in either a tertiary contact or protein binding. This is seen in the solution structure of the isolated P loop of 23 S rRNA which is closed by a non-canonical G·C pair and contains three highly conserved splayed-out guanosine bases which, in the full 23 S rRNA, are involved in tertiary contacts (Viani Puglisi *et al.*, 1997).

### Pseudoknots

A pseudoknot may be considered as consisting of two fused hairpins in which the loops mutually provide bases for the formation of two stems. Classically, pseudoknots are described as elements of RNA tertiary structure derived from base-pairing between a hairpin loop and a complementary single-stranded sequence close to the hairpin (Pleij *et al.*, 1985; Puglisi *et al.*, 1988). The two base-paired stems are stacked upon each other at the junction (Hilbers *et al.*, 1998), giving rise to an extended architecture (Figure 5(a)). Helix arrangement in pseudoknots is subject to some conformational flexibility as both linear and bent geometries of the stems have been found in different pseudoknot structures along with either coaxial or non-coaxial stacking. In non-coaxially stacked pseudoknots, considerable rotation between the adjacent stems, along with shifting and tilting of the helix axes occurs, giving rise to an overall kinked geometry, as has been observed in a ribosomal frameshifting pseudoknot (Su *et al.*, 1999). In some pseudoknots, the two helices are separated by intervening nucleotides which can stack between the stems, introducing a bent junction (Chen *et al.*, 1996). The kinked shape plays an important role in the biological role of frameshifting pseudoknots (Chen *et al.*, 1996; Su *et al.*, 1999).

The connecting loops in pseudoknots cross the grooves of the helices so as to position loop 1 in the deep groove of stem 2 and loop 2 in the shallow groove of stem 1. In order to accommodate the single strands, the geometry of the stems may deviate significantly from standard *A*-form RNA helices (Su *et al.*, 1999). The single-stranded loops can, thus, participate in tertiary contacts in the RNA grooves, attested by a triple-strand interaction of a single base discovered in a classical pseudoknot (Kolk *et al.*, 1998) and triplex formation over six layers in a ribosomal frameshifting viral pseudoknot (see above) (Su *et al.*, 1999). Distortions of the *A*-conformation in pseudoknot helices can be also induced by connecting loops which are too short to cross the grooves of regular stems (Pleij *et al.*, 1985). In the crystal structure of a ribosomal frameshifting pseudoknot, a two-nucleotide loop is accommodated by overtwisting of the first base-pair in the adjacent stem (Su *et al.*, 1999).

The tight packing of stems and single-stranded regions in pseudoknots leads to highly compact folds which are stabilized by metal ion binding (see below) and extensive tertiary hydrogen bonding giving rise to characteristic RNA structural motifs such as base quadruples (Su *et al.*, 1999) and ribose zippers (Ferré-D'Amaré *et al.*, 1998) (see above). Two nested pseudonots are responsible for the extraordinary stability of the HDV ribozyme, the three-dimensional fold of which beautifully demonstrates RNA's capacity to form intricate, compact architectures (Figure 5(b)) (Ferré-D'Amaré *et al.*, 1998).

### Interactions of loops: kissing and docking

Loops are stabilized by interactions often involving only few functional groups of the bases which are, thus, available for tertiary contacts with other RNA domains. Such tertiary interactions comprise loop-to-loop ''kissing'' complexes and loop-receptor complexes in which a loop docks into a double-

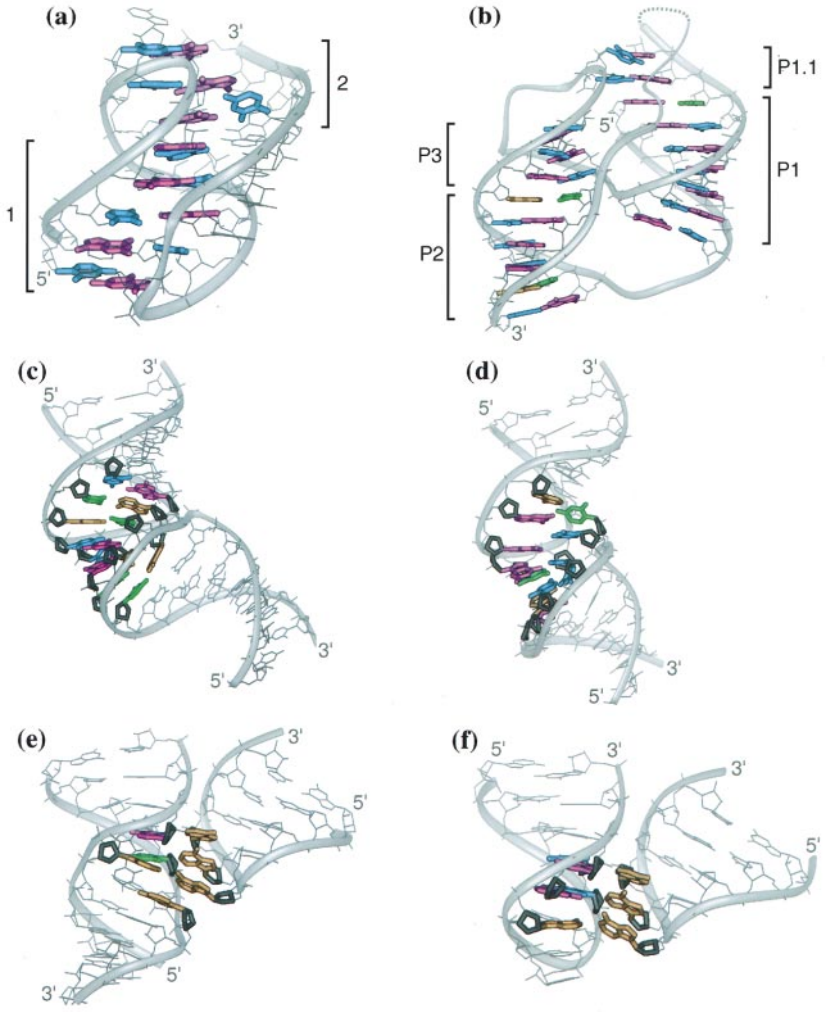


**Figure 5.** Compact folding of RNA architectures depends on conserved tertiary interaction motifs such as (a), (b) pseudoknots and (c)-(f) long-range contacts involving loops. Stacking between two stems (1 and 2, indicated by parentheses) and tight packing of the connecting single-stranded regions into the helix grooves stabilize pseudoknots as illustrated by the crystal structures of a (a) frameshifting viral pseudoknot (Su *et al*., 1999) and (b) the HDV ribozyme (Ferré-D'Amaré *et al*., 1998). In the two nested pseudoknots of the HDV ribozyme, comprising the stem pairs P1-P2 and P1.1-P3 (parentheses), the order of stacking deviates from that of classical pseudoknots. Complementary loops can form kissing complexes in which all single-stranded nucleotides are involved in base-pairing giving rise to quasicontinuous helix modules as have been observed for the complexes of the (c) *ColE1* inverted loop RNA (Lee & Crothers, 1998) and the (d) HIV-2 TAR hairpin (Chang & Tinoco, 1997). In GNRA tetraloops, the bases are free to form tertiary interactions with other RNA motifs such as conserved tetraloop (e) receptors and (f) helices. (e) In the P4-P6 domain of group I ribozymes, a GAAA loop docks into a receptor module (Cate *et al*., 1996a) which provides stabilizing contacts by stacking, mismatch-pairing and hydrogen bonding to the backbone. (f) An intermolecular contact in crystals of a hammerhead ribozyme provides an example for a GAAA tetraloop interacting with the shallow groove of a helix (Pley *et al*., 1994b).

stranded receptor motif. An interaction between two loops has already been found in tRNA, namely between the D and T-loops, which, however, are stabilized by additional tertiary contacts of the flanking regions (Quigley & Rich, 1976).

Kissing complexes are formed by base-pairing between the single-stranded residues of sequence-complementary loops. To date, structural data are available for two RNA kissing contacts, namely of the *ColE1* inverted loop RNA (Marino *et al*., 1995; Lee & Crothers, 1998) and the HIV-2 TAR hairpin (Chang & Tinoco 1997) in complex with their respective complementary loops (Figure 5(c)-(d)). These complexes represent transient structures which, *in vivo*, precede the formation of extended sense-antisense RNA duplexes. The kissing complexes are, however, stable RNA architectures which are specifically recognized by proteins (Eguchi & Tomizawa, 1990).

In both the TAR and *ColE1* kissing complexes, all loop residues participate in intermolecular base-pairing, creating double helices of, respectively, six and seven Watson-Crick pairs. The duplex formed by the kissing loops stacks on both sides coaxially with the flanking hairpin helices. The overall structure of the kissing complexes resembles a bent quasicontinuous helix. The extensive base-pairing between the loops forces several of the backbone phosphate groups into close proximity giving rise to ''phosphate clusters'' (Lee & Crothers, 1998) which are stabilized by metal cations (see below) (Jossinet *et al*., 1999).

GNRA tetraloops (see above) are frequently found in large RNAs such as rRNA, self-splicing introns and RNase P, with these motifs playing crucial roles in proper folding by providing long-range tertiary interactions with appropriate docking sites (Costa & Michel, 1995). Such contacts, first suggested based on phylogenetic analyses and modeling studies (Michel & Westhof, 1990; Jaeger *et al*., 1994), have later been confirmed by crystal structures of ribozymes in which GAAA tetraloops dock into either a helix or a tetraloop receptor.

In the crystal packing of a hammerhead ribozyme, intermolecular contacts are formed between the three adenosine bases of a GAAA tetraloop in

one RNA molecule and the shallow groove edges of consecutive G:C base-pairs in a second molecule (Figure 5(f)) (Pley *et al.*, 1994b). In addition to hydrogen bonds between the bases, interactions involve 2′-OH groups of the residues in both the tetraloop and the helix. A similar contact between a tetraloop and the shallow groove of an RNA duplex has been identified for the L9 GAAA loop docking into the P5 stem in the group I ribozymes (Golden *et al.*, 1998).

Also in group I ribozymes, but in the P4-P6 domain, a tertiary interaction has been observed between a GAAA tetraloop and a specific RNA receptor module (Figure 5(e)) (Cate *et al.*, 1996a). The tetraloop receptor motif comprises 11 conserved nucleotides (Costa & Michel, 1995), five of which are located in an internal loop zippering up into a compact duplex-like structure with an opened shallow groove and nearly coaxial alignment of the flanking helices. The key feature of the tetraloop receptor is a notch resulting from a reduced backbone rise due to base-pairing between two consecutive adenosine bases in an adenosine platform (see above). The platform opens up the shallow groove and serves as a stacking surface for the pivotal adenine of the GAAA tetraloop which docks into the notch within the receptor motif. As in the tetraloop-helix contact, hydrogen bonds involving both bases and 2′-OH groups of the backbone contribute to the stability of the tetraloop-receptor interaction motif. The tetraloop receptor itself is stabilized by a potassium ion specifically bound within a pocket below the adenosine platform (see below) (Basu *et al.*, 1998).

A peculiar example of an internal loop docking into a receptor structure is provided by the hairpin ribozyme. This catalytic RNA motif consists essentially of two connected duplexes each carrying a large internal loop. The loop B, comprising 16 nucleotides, zippers up by mismatch formation into an open duplex with an extremely widened shallow groove (Figure 2(c)) (Butcher *et al.*, 1999) which serves as a docking surface for loop A (Cai & Tinoco, 1996; Earnshaw *et al.*, 1997).

## Adding glue to RNA folds: metal ions and water

RNA is a polyanion which binds cations and water in order to acquire its biologically functional three-dimensional structure. In the absence of metal ions, all known complex RNA folds form most of their secondary structure but little, if any, tertiary structure. In many crystal structures of RNA molecules, metal ions have been found at specific binding sites (reviewed by Feig & Uhlenbeck, 1999). Some catalytic RNAs, termed ribozymes, use metal ions in their active sites (reviewed by Pyle, 1993; Carola & Eckstein, 1999; McKay & Wedekind, 1999). Evidence for stable, conserved water molecules participating in RNA structures comes from high-resolution crystallographic analyses. Both metal ions and water can, thus, be considered as intrinsic parts of RNA three-dimensional architecture (Westhof, 1988).

### *Di- and monovalent cations*

At physiological pH, RNA molecules are associated with a number of counter ions corresponding to that of the negatively charged nucleotides. While the majority of cations are delocalized, contributing to non-specific counter ion condensation, some ions occupy specific ion binding sites precisely defined by the three-dimensional folding of the RNA chain (Laing *et al.*, 1994). These metal ion binding sites (Figure 6) are positioned within local minima in the electrostatic field created by the charges on the RNA atoms (Hermann & Westhof, 1998a). The important role of metal ions for RNA three-dimensional folding was recognized early when it was found that cations are necessary for the stabilization of the native structure of tRNA (Fresco *et al.*, 1966). Crystal structure analysis has later revealed the positions of four $Mg^{2+}$ bound to $tRNA^{Phe}$ in loop regions (Figure 6(a) and (b)) (Holbrook *et al.*, 1977; Jack *et al.*, 1977; Quigley *et al.*, 1978).

Specific binding of metal ions to RNA can occur either directly through inner-sphere coordination (Figure 6(a)), or can be mediated by the hydration shell *via* outer-sphere contacts (Figure 6(b)). The interaction between RNA and directly bound metal ions is governed by a competition between electrostatic and hydration forces, both of which are decreasing functions of ionic radius and increasing with ionic charge (Draper & Misra, 1998). Small divalent ions such as $Mg^{2+}$ bind in inner-sphere coordination almost exclusively to the negatively charged phosphate groups of RNA, while interactions with the hydration shell of $Mg^{2+}$ also involve other, less electronegative sites such as purine N7 and carbonyl oxygen atoms of the bases.

Metal ions stabilize the close approach of phosphate groups at the interface between RNA strands, as has frequently been observed in crystal packing (Holbrook & Kim, 1997) and in strand junctions associated with complex folds such as the P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a) and the L11-binding domain of 23 S rRNA (Wimberly *et al.*, 1999). Specific pockets for the binding of divalent metal ions are formed in loops and bulges where the RNA backbone folds back on itself bringing phosphate groups in close proximity. Such metal-stabilized, folded backbone structures have been found in the turn connecting the acceptor stem and the D-loop in tRNA (Figure 6(b)) (Holbrook *et al.*, 1977; Jack *et al.*, 1977; Quigley *et al.*, 1978), in the A-rich bulge of the P4-P6 domain of group I ribozymes (Figure 6(c)) (Cate *et al.*, 1996a), in the UCU bulge in HIV TAR (Figure 6(d)) (Ippolito & Steitz, 1998), in the L11-binding domain of 23 S rRNA (Wimberly *et al.*, 1999) and in kissing loop complexes (Lee & Crothers, 1998; Jossinet *et al.*, 1999).

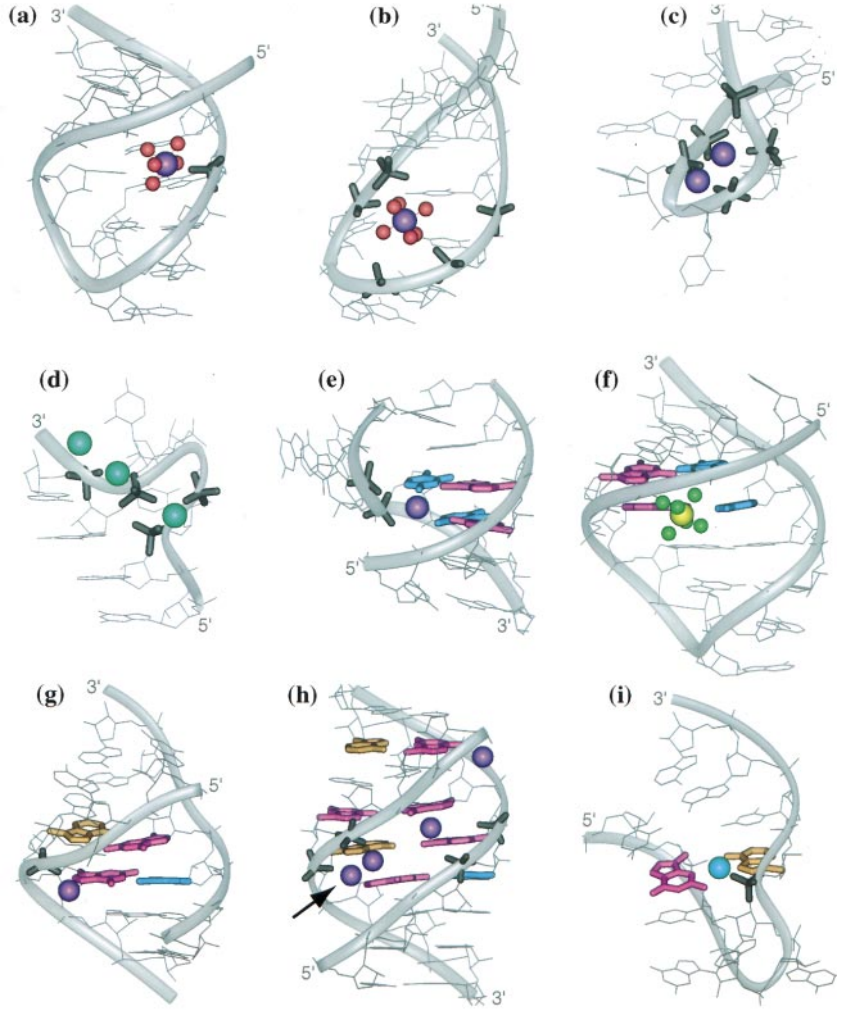


**Figure 6.** Metal ions are an integral part of RNA architecture where they occupy specific binding pockets, contributing to both formation and stabilization of RNA structural motifs. Metal cations are bound either directly, (a) *via* inner-shell coordination, or (b) by outer-shell contacts mediated by hydration water, as shown here for two $Mg^{2+}$ in the (a) anticodon loop and the (b) turn between the acceptor stem and D loop of $tRNA^{Phe}$ (Holbrook *et al.*, 1977; Jack *et al.*, 1977; Quigley *et al.*, 1978). In narrow turns such as the (c) A-rich bulge in the P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a) and the (d) UCU bulge in HIV-1 TAR RNA (Ippolito & Steitz, 1998), phosphate groups are brought in close proximity requiring metal cations which compensate the electrostatic repulsion. Metal ions can simultaneously interact with phosphate groups and bases of RNA, tethering the backbone of single-stranded regions with the groove of duplexes, as in the (e) lead-dependent ribozyme (Wedekind & McKay, 1999). Some non-canonical base-pairs bind metal ions at structurally conserved locations, as illustrated by an (f) osmium(III) hexammine found at a $Mg^{2+}$-binding site at a G·U tandem pair in the P4-P6 domain of group I ribozymes (Cate *et al.*, 1996a) and a (g) divalent cation coordinated at the deep groove edge of a sheared G·A pair in an RNA duplex (Baeyens *et al.*, 1996). (h) A stack of non-canonical base-pairs is stabilized by several $Mg^{2+}$ bound to the deep groove in the high-salt crystal structure of the loop E of 5 S ribosomal RNA (Correll *et al.*, 1997). (The arrow indicates a pair of binding sites which, due to their close proximity, are likely to be occupied in an alternating fashion.) (i) Specific binding sites for monovalent ions have recently been characterized, among them a sodium ion in a ribosomal frameshifting pseudoknot (Su *et al.*, 1999), tying the loop 2 into the shallow groove of stem 1 (see also Figure 5(a)). Phosphate groups are shown as grey stick tetrahedra, metal ions as spheres colored according to: (a)-(c), (e), (g), (h) $Mg^{2+}$, violet; (d), $Ca^{2+}$ cyan; (f), Os(III), yellow; (i) $Na^{+}$, light blue. Water ligands are shown as red spheres in (a) and (b), ammine in green in (f). In all panels, except for (a), (b) and (f), water ligands of the hydration shell of cations have been omitted for clarity.

The requirement of certain bulges for metal ions to adopt a defined three-dimensional conformation is illustrated by the HIV-1 TAR RNA and the lead-dependent ribozyme for both of which solution structures, solved by NMR at low salt conentrations, and crystal structures, obtained in the presence of high ion concentrations, are available. In HIV-1 TAR RNA, the UCU bulge is flexible in low-salt solution (Aboul-ela *et al.*, 1996), but forms a rigid, metal ion-stabilized structure in the crystal in the presence of high concentrations of $Ca^{2+}$ (Figure 6(d)) (Ippolito & Steitz, 1998). In the crystal structure of the lead-dependent ribozyme, the backbone of the GAG bulge is tethered to the deep groove by a fully hydrated $Mg^{2+}$ which interacts with two phosphate groups of the bulge and two guanosine bases in the adjacent stem (Figure 6(e)) (Wedekind & McKay, 1999), whereas in low-salt solution the bases of the bulge are pointing inside the groove (Hoogstraten *et al.*, 1998).

In the A-rich bulge of group I ribozymes, the backbone follows a corkscrew turn positioning phosphate groups together with $Mg^{2+}$ in a cluster inside the loop while the bases are oriented towards the outside (Cate *et al.*, 1996a). The metal-associated corkscrew geometry of the A-rich bulge provides a most intriguing and beautiful example of the versatility of RNA architectures, since the prevailing organization of nucleotides in RNA modules, with their bases oriented to the core and their backbone lining the surface, is turned inside-out. Among the numerous metal ion binding motifs discovered in the crystal structure of the P4-P6 domain of group I ribozymes, it was primarily

the A-rich bulge which led to the proposal that RNA folds around a ''metal ion core'' (Cate *et al.*, 1997), in analogy to the hydrophobic core of proteins.

In addition to the formally charged phosphate groups of the backbone, electronegative atoms of the bases contribute to specific metal ion binding in RNA architectures. This becomes evident especially for deep groove regions where the electrostatic field of regular helical RNA is perturbed by non-canonical base-pairs. A minimal change of duplex geometry which is introduced by a G·U wobble pair followed by a Y·G pair is sufficient to create specific cation-binding pockets in the deep groove (Cate & Doudna, 1996; Hermann & Westhof, 1998a). The uridine O4 atom, which is not involved in hydrogen bonding, along with the N7 and O6 atoms of guanosine, contribute to cation coordination at the G·U pair. Metal ion binding sites at the 5′-GU-3′/3′-YG-5′ motif have been found in the anticodon stem/loop of $tRNA^{Asp}$ (Westhof *et al.*, 1985) and in the P4-P6 domain of group I ribozymes (Figure 6(f)) (Cate *et al.*, 1996a; Cate & Doudna, 1996; Kieft & Tinoco, 1997; Colmenarejo & Tinoco, 1999). The 5′-GU-3′/3′-YG-5′ motif occurs frequently in large RNAs such as rRNA (Gautheret *et al.*, 1995) and might be of general importance as a metal ion binding site in RNA folds.

Sheared G·A pairs (see above) are another example for non-canonical base-pairs associated with cation binding sites. Divalent metal ions have been found in an RNA duplex (Figure 6(g)) (Baeyens *et al.*, 1996) and in hammerhead ribozymes (Pley *et al.*, 1994a; Scott *et al.*, 1995) at the deep groove edge of guanosine in Watson-Crick C:G pairs which are followed by a sheared G·A pair. At the 5′-CG-3′/3′-GA-5′ motif, cations bind to the N7 atom of the C:G and the phosphate group of the G·A which, due to the backbone conformation introduced at the G·A pair, are positioned to line an electronegative pocket. $Mg^{2+}$ binding at a sheared G·A pair has also been found in the loop E of 5 S rRNA, in which a stack of three non-canonical base-pairs, centered at the G·A, is stabilized by extensive metal ion binding in the deep groove (Figure 6(h)) (Correll *et al.*, 1997). This ''metal ion zipper'', which forces a significant narrowing of the deep groove, might partially form due to the crystallization conditions at high $Mg^{2+}$ concentrations.

While a considerable number of nucleic acid three-dimensional structures containing divalent metal ions are available, examples for monovalent cations specifically bound to either DNA or RNA are scarce, with the exception of G quadruplexes which are stabilized by alkali metal ions between the layers of the G tetrad planes (Kang *et al.*, 1992; Cheong & Moore, 1992; Laughlan *et al.*, 1994; Hud *et al.*, 1996; Kettani *et al.*, 1998). The role of monovalent metal ions as a part of RNA three-dimensional structure has been acknowledged only recently (Wang *et al.*, 1993; Draper & Misra 1998) despite it being known for some time that sodium ions can subsitute for $Mg^{2+}$ in the correct folding of tRNA (Crothers & Cole, 1978). The growing number of high-resolution crystal structures of RNA molecules in combination with novel ''chemogenetic'' approaches (see below) for identifying the binding partners of individual RNA chemical groups (reviewed by Strobel, 1999), have permitted the discovery of specific pockets for monovalent metal ions in RNA architectures.

In the crystal structure of a ribosomal frameshifting pseudoknot, a sodium ion is located at a key position tethering the single-stranded loop 2 into the shallow groove of stem 1 (Figures 5(a) and 6(i)) (Su *et al.*, 1999). The sodium ion binds to a phosphate group of the loop and is positioned in the plane between two purines coordinating through nitrogen atoms to the ion. The tight packing of loops and stems requires the presence of metal cations for the stable folding of pseudoknots (Wyatt *et al.*, 1990; Kolk *et al.*, 1998) which contain binding sites for divalent metal ions (Hermann & Westhof, 1998a; Gonzalez & Tinoco, 1999). A high concentration of monovalent cations can, however, substitute for divalent metals in pseudoknot folding (Wyatt & Tinoco, 1993).

Adenosine platforms, conserved building blocks of RNA architectures (see above), contain a specific binding site for monovalent metal ions as an integral part of their structure (Basu *et al.*, 1998). In the P4-P6 domain of group I ribozymes, monovalent ions have been found immediately below adenosine platforms (Cate *et al.*, 1996b; Basu *et al.*, 1998), stabilizing the intra-strand side-by-side AA pair by coordinating to the phosphate group between the adenosine bases and to bases of the non-canonical pair which stacks below. Since adenosine platforms are a key structural element of the tetraloop receptor module, site-specific binding of monovalent cations is important for the formation of correct tertiary structure in a wide variety of large RNAs which contain the tetraloop-receptor contact.

In addition to the role they play in RNA three-dimensional folding, metal ions are constituents of the active site of self-cleaving catalytic RNAs such as the self-splicing group I and II introns (Streicher *et al.*, 1996; Deme *et al.*, 1999; Sontheimer *et al.*, 1999), RNaseP (Smith & Pace, 1993), hammerhead ribozyme (Dahm & Uhlenbeck, 1991; Scott & Uhlenbeck 1999) and the lead-dependent ribozyme (Pan & Uhlenbeck, 1992). Crystal structure analyses have revealed positions of putative catalytic metal ions in the hammerhead and lead-dependent ribozymes (Wedekind & McKay, 1999; Pley *et al.*, 1994a; Scott *et al.*, 1995, 1996). In the active site of a group I intron, two metal ions have been positioned in a three-dimensional model of the catalytic core (Streicher *et al.*, 1996) resembling a geometry proposed earlier (Steitz & Steitz, 1993).

The divalent cations suggested to participate in catalysis in both the hammerhead and lead-dependent ribozymes are directly coordinated to the scissile phosphate group. The cations at the active site

of the ribozymes drive the chemistry of the cleavage reaction (reviewed by Kuimelis & McLaughlin, 1998) by providing a metal-bound hydroxide which can deprotonate the 2′-OH group for nucleophilic substitution at the adjacent phosphate. Based on molecular dynamics simulations, it has been suggested that the hydroxide involved in the catalysis of the hammerhead ribozyme (Dahm *et al.*, 1993) is bridging between two $Mg^{2+}$ in proximity of the cleavage site (Hermann *et al.*, 1997).

### *Ordered water in RNA structures*

Stable hydration patterns in surface depressions, pockets and the grooves of duplexes contribute to the stability of RNA architectures (Figure 7). Metal binding sites in RNA folds (see above) inevitably contain restricted water molecules of the hydration shell around the cations. High-resolution X-ray crystallography permits the identification of the positions of tightly bound water molecules in RNA architectures. Organized networks of metal hydration water have been found in crystal structures of the loop E of 5 S rRNA (Correll *et al.*, 1997), HIV-1 TAR RNA (Ippolito & Steitz, 1998), a ribosomal frameshifting pseudoknot (Su *et al.*, 1999) and tRNAs (Westhof, 1988).

In tRNAs, a regular hydration pattern has been observed around the backbone, in double-stranded regions, where water molecules bridge successive phosphate groups of each strand (Westhof *et al.*, 1988). Similar stable arrangements of water are seen in *A*-form RNA duplexes in which solvent molecules form transversal bridges between the strands across the shallow groove and regular longitudinal hydration motifs in the deep groove (Figure 7(b)) (Egli *et al.*, 1996).

The preferred binding of water at phosphate groups is due to both hydrogen bonding to the anionic OP atoms and the electronegative potential around these sites. The effect of electrostatical restriction of water molecules gives rise to the crystallographic ''water sites'' in the anticodon loop of tRNAs (Westhof *et al.*, 1985, 1988), which have been shown to coincide with electronegative minima in the electrostatic field around the RNA (Hermann & Westhof, 1998a). Like metal ions, water molecules occupy electronegative pockets, thereby stabilizing the folding of the RNA chain. Extensive association of tertiary contacts with ordered solvent has been found in the crystal structures of tRNAs (Westhof *et al.*, 1985, 1988). Organized networks of restricted water molecules are

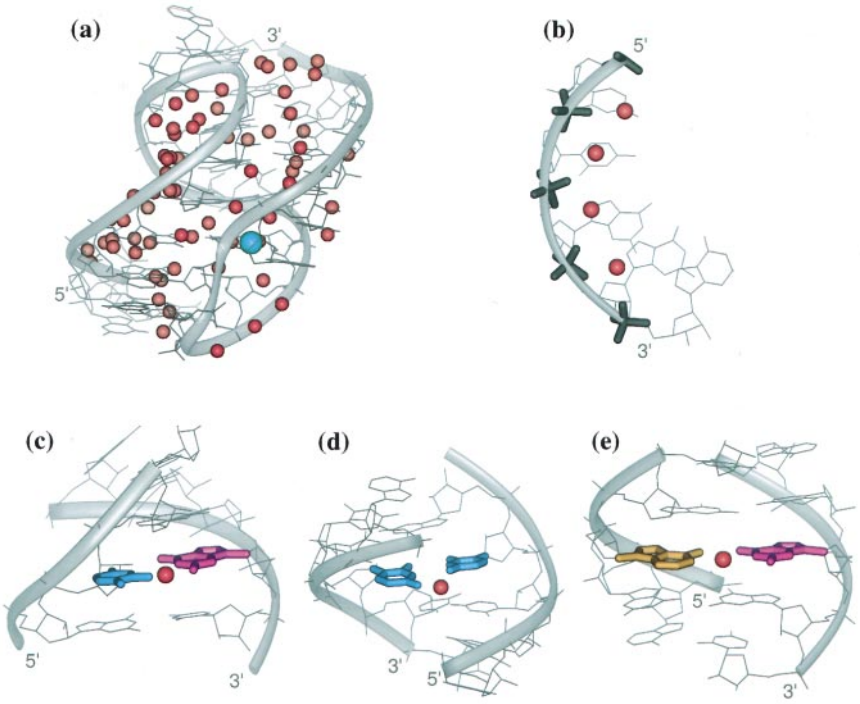


**Figure 7.** Water molecules forming stabilizing hydrogen bond networks are an intrinsic part of RNA architectures. (a) The interpenetration of RNA three-dimensional structure by locally restricted water molecules is beautifully demonstrated by the crystal structure of a frameshifting ribosomal pseudoknot (Su *et al.*, 1999). (b) Regular hydration patterns are associated with the backbone of *A*-form RNA in which water molecules bridge successive phosphate groups as in the crystal structure of an RNA duplex (Egli *et al.*, 1996). Many mismatch base-pairs are stabilized by water molecules linking free polar groups in both bases with hydrogen bonds. Such water bridges are located, for example, at the shallow groove edges of G·U and sheared G·A pairs, as seen in (c) tRNA (Westhof *et al.*, 1988) and the (e) loop E of 5 S rRNA (Correll *et al.*, 1997). (d) At the deep groove edge, a water molecule is found stabilizing a U·U mismatch in an RNA duplex (Baeyens *et al.*, 1995). Water molecules are shown as red spheres, and a $Na^{+}$ in (a) is colored in light blue.

located at the surface of a C-turn (see above) and mediate the tertiary interaction of loop 1 in the deep groove of stem 2 in a ribosomal frameshifting pseudoknot (Figure 7(a)) (Su *et al.*, 1999).

Some of the stable hydration patterns discussed above involve the bases in hydrogen-bonded pairs which are surrounded by systematic arrangements of water (Auffinger & Westhof, 1998). Single bridging water molecules form hydrogen bonds between unoccupied polar groups of the bases in many mismatch pairs (Correll *et al.*, 1997; Leontis & Westhof, 1998a). In G·U pairs, the U is retracted towards the deep groove, creating a depression on the shallow groove surface which is a conserved site of specific hydration with the occupying water molecule bridging the two bases of the wobble pair (Westhof, 1988). Such precisely defined water molecules at G·U pairs have been found in the crystal structures of tRNAs (Figure 7(c)) (Westhof *et al.*, 1988), RNA duplexes (Mueller *et al.*, 1999) and the loop E of 5 S rRNA (Correll *et al.*, 1997).

In contrast to G·U wobble pairs for which the bridging water molecule is located in the shallow groove, U·U mismatches are stabilized by a water molecule in the deep groove, forming hydrogen bonds between the O4 carbonyl atoms of the uridine bases (Figure 7(d)) (Baeyens *et al.*, 1995). Conserved water binding sites linking the bases in the shallow groove are also found at G·A mismatch pairs (Figure 7(e)) (Leonard *et al.*, 1994; Correll *et al.*, 1997). In the crystal structure of the loop E of 5 S rRNA, three consecutive mismatches, centered at a sheared G·A, are stabilized by water molecules alternatingly bridging the shallow and deep groove edges of the base-pairs (Correll *et al.*, 1997). These non-canonical base-pairs are connected by single hydrogen bonds and, thus, require water molecules in order to obtain a stable base-pairing scheme.

In addition to their function for stabilizing RNA three-dimensional structure, water molecules play an important role in recognition processes. In the interactions between RNA architectures and both proteins and small molecules, water molecules provide plasticity at the intermolecular surfaces insuring specific recognition and high-affinity binding.

## The shape of things to come: some future prospects of understanding RNA architecture and function

Analyses of the many three-dimensional structures of RNA molecules becoming available over the last decade have answered important questions about the principles that govern the assembly of RNA architectures (reviewed by Batey *et al.*, 1999; Moore, 1999; this review). While our understanding of RNA structure-function relationships is now deeper beyond the surface once scratched when the first crystal structures of tRNAs had been solved, some of the most exciting challenges of the RNA world still await solutions.

Among these challenges are the three-dimensional structures of large RNAs and RNA-protein complexes, especially those of self-splicing group II introns, RNase P and the multicomponent assemblies of the ribosome and the splicing machinery. Working models of the RNA component of RNase P (Massire *et al.*, 1998) and 16 S rRNA (Malhotra & Harvey, 1994; Mueller & Brimacombe, 1997; Masquida *et al.*, 1997) are available. It will be interesting to see if these models are in similar good agreement with experimental structures yet to be determined, as was the model for the group I ribozyme (Golden *et al.*, 1998).

Despite being ephemeral approximations of molecular architecture, structural models will continue being important amplifiers for our thinking about structure-function relationships in the RNA world, especially with the aim of designing new experiments. The construction of RNA molecular models will greatly benefit from chemogenetic techniques which allow the probing of single chemical groups for their contribution to the function of an RNA fold (reviewed by Strobel, 1999).

The elucidation of the crystal structures of the complete ribosomal subunits has been a challenge of longstanding interest (reviewed by Yonath & Franceschi, 1998; Moore, 1998). Significant progress towards these goals has recently been achieved for both the isolated small and large subunits and the complete functional ribosome. For the isolated subunits, electron density maps of around 5 Å resolution are available (Clemons *et al.*, 1999; Ban *et al.*, 1999). The crystal structure of a 70 S ribosome functional complex containing also tRNA and a piece of mRNA has been refined to 7.8 Å (Cate *et al.*, 1999). Since the crystals of the subunits and the functional ribosome diffract to higher resolution, one can anticipate an exponential increase in our understanding of RNA-RNA and RNA-protein interactions in the forseeable future.

Static three-dimensional architecture alone can provide only partial insights into the function of molecular machines, such as the ribosome, which require dynamic flexibility in order to self-assemble and accomplish their tasks in the cell. The importance of flexibility for RNA function is highlighted by the simplest RNA molecular machines, namely the ribozymes which catalyze chemical reactions in the absence of proteins. In the small autocleaving RNAs, such as the hammerhead and lead-dependent ribozymes, the scissile phosphodiester bond is located in a flexible region of the backbone (Hoogstraten *et al.*, 1998; Legault *et al.*, 1998a; Wedekind & McKay, 1999; Scott *et al.*, 1996; Murray *et al.*, 1998). Defined conformational changes of the backbone have been found to be crucial for the cleavage reaction to proceed in these ribozymes (reviewed by Lilley, 1999; McKay & Wedekind, 1999). In the hammerhead ribozyme, a sugar moiety of the nucleotide preceding the scissile phosphodiester bond must flip its ring conformation in order to position the 2′-OH nucleophile in the proper orientation for the substitution at the

adjacent phosphate group (Hermann *et al*., 1997; Murray *et al*., 1998).

In a manner similar to the efforts to study the significance of structural rearrangements for RNA function, the understanding of the roles base and backbone modifications play in RNA architectures is still in its infancy. While it is known that modified nucleotides are required for the precise recognition of tRNAs by proteins (reviewed by Grosjean & Benne, 1998), molecular insight into the influence of modifications on RNA structure and flexibility are yet scarce. In some cases, modified nucleosides contribute to the thermostability of RNAs in thermophilic organisms (Edmonds *et al*., 1991). For a base-methylated GGAA loop in 16 S rRNA it has recently been shown that the methyl groups induce a stable yet completely different conformation compared to that of canoncial GNRA tetraloops (Rife & Moore, 1998). It is likely that the peculiar conformation of the methylated loop is crucial for its function in rRNA.

In the above discussion, the importance of RNA architecture for recognition by small molecules and proteins has occasionally been mentioned. Indeed, the large majority of cellular RNAs function in concert with partner molecules, either as stable permanent complexes, such as the ribosome, or in transient associations such as the mRNA splicing machinery. The potential specificity of interactions between RNAs and small molecules renders RNA a prime target for therapeutic intervention (reviewed by Hermann & Westhof, 1998b; Afshar *et al*., 1999), especially because various RNA molecules are involved in all stages of gene expression. While the rational design of drugs which specifically recognize RNA motifs is still in its infancy, promising strategies for developing small molecules targeting RNAs and RNA-protein complexes have been outlined (reviewed by Hermann, 2000). The growing understanding of RNA architecture and its correlation with function will provide firm grounds for the development of specific RNA-directed effector molecules.

First steps have been undertaken to unravel the principles of molecular recognition in complexes between RNA and proteins (reviewed by De Guzman *et al*., 1998; Cusack, 1999; Draper & Reynaldo, 1999; Kambach *et al*., 1999; Steitz, 1999), peptides (reviewed by Frankel & Smith, 1998; Patel, 1999; Puglisi & Williamson, 1999) and small molecules (reviewed by Feigon *et al*., 1996; Patel *et al*., 1997; Hermann & Westhof, 1998b). Despite the wide diversity of molecular compositions and biological functions of the studied systems, a limited number of common themes associated with intermolecular interactions involving RNA have been uncovered. Satisfyingly, the understanding of the three-dimensional architectures of the participating molecules has allowed the drawing of a preliminary map of some general routes associated with biomolecular structure-function relationships. It is this map that will permit future navigation in the immensely complex and challenging world of biological processes which involve RNA.

## Acknowledgements

Funding was provided by NIH GM-54777.

## References

Aboul-ela, F., Karn, J. & Varani, G. (1996). Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucl. Acids Res.* **24**, 3974-3981.

Abramovitz, D. L. & Pyle, A. M. (1997). Remarkable morphological variability of a common RNA folding motif: the GNRA tetraloop-receptor interaction. *J. Mol. Biol.* **266**, 493-506.

Afshar, M., Prescott, C. D. & Varani, G. (1999). Structure-based and combinatorial search for new RNA-binding drugs. *Curr. Opin. Biotechnol.* **10**, 59-63.

Allain, F. H. & Varani, G. (1995). Structure of the P1 helix from group I self-splicing introns. *J. Mol. Biol.* **250**, 333-353.

Auffinger, P. & Westhof, E. (1998). Hydration of RNA base-pairs. *J. Biomol. Struct. Dynam.* **16**, 693-707.

Baeyens, K. J., DeBondt, H. L. & Holbrook, S. R. (1995). Structure of an RNA double helix including uracil-uracil base-pairs in an internal loop. *Nature Struct. Biol.* **2**, 56-62.

Baeyens, K. J., DeBondt, H. L., Pardi, A. & Holbrook, S. R. (1996). A curved RNA helix incorporating an internal loop with G·A and A·A non-Watson-Crick base-pairing. *Proc. Natl Acad. Sci. USA,* **93**, 12851-12855.

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B. & Steitz, T. A. (1990). Placement of protein and RNA structures into a 5 Å-resolution map of the 50 S ribosomal subunit. *Nature,* **400**, 841-847.

Basu, S., Rambo, R. P., Strauss-Soukup, J., Cate, J. H., Ferré-D'Amaré, A. R., Strobel, S. A. & Doudna, J. A. (1998). A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nature Struct. Biol.* **5**, 986-992.

Batey, R. T., Rambo, R. P. & Doudna, J. A. (1999). Tertiary motifs in RNA structure and folding. *Angew. Chem. Int. Ed. Engl.* **38**, 2326-2343.

Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D. & Williamson, J. R. (1996). α helix-RNA major groove recogntion in an HIV-1 rev peptide-RRE RNA complex. *Science,* **273**, 1547-1551.

Been, M. D. & Wickham, G. S. (1997). Self-cleaving ribozymes of hepatitis delta virus RNA. *Eur. J. Biochem.* **247**, 741-753.

Borer, P. N., Lin, Y., Wang, S., Roggenbuck, M. W., Gott, J. M., Uhlenbeck, O. C. & Pelczer, I. (1995). Proton NMR and structural features of a 24-nucleotide RNA hairpin. *Biochemistry*, **34**, 6488-6503.

Brion, P. & Westhof, E. (1997). Hierarchy and dynamics of RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 113-137.

Brodsky, A. S. & Williamson, J. R. (1997). Solution structure of the HIV-2 TAR-argininеamide complex. *J. Mol. Biol.* **267**, 624-639.

Brodsky, A. S., Erlacher, H. A. & Williamson, J. R. (1998). NMR evidence for a base triple in the HIV-2

TAR C-G·C$^+$ mutant-argininamide complex. *Nucl. Acids Res.* **26**, 1991-1995.

Butcher, S. E., Dieckmann, T. & Feigon, J. (1997). Solution structure of a GAAA tetraloop receptor RNA. *EMBO J.* **16**, 7490-7499.

Butcher, S. E., Allain, F. H.-T. & Feigon, J. (1999). Solution structure of the loop B domain from the hairpin ribozyme. *Nature Struct. Biol.* **6**, 212-216.

Cai, Z. & Tinoco, I., Jr (1996). Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.

Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A. & Patel, D. J. (1998). Solution structure of P22 transcriptional antitermination N peptide-box B RNA complex. *Nature Struct. Biol.* **5**, 203-212.

Carola, C. & Eckstein, F. (1999). Nucleic acid enzymes. *Curr. Opin. Struct. Biol.* **3**, 274-283.

Cate, J. H. & Doudna, J. A. (1996). Metal-binding sites in the major groove of a large ribozyme domain. *Structure,* **4**, 1221-1229.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. E. & Doudna, J. A. (1996a). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science,* **273**, 1678-1685.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. E. & Doudna, J. A. (1996b). RNA tertiary structure mediation by adenosine platforms. *Science,* **273**, 1696-1699.

Cate, J. H., Hanna, R. L. & Doudna, J. A. (1997). A magnesium ion core at the heart of a ribozyme domain. *Nature Struct. Biol.* **4**, 553-558.

Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. (1999). X-ray crystal structure of 70 S ribosome functional complexes. *Science,* **285**, 2095-2135.

Chang, K. Y. & Tinoco, I., Jr (1997). The structure of an RNA ''kissing'' hairpin complex of the HIV TAR hairpin loop and its complement. *J. Mol. Biol.* **269**, 52-66.

Chen, X., Kang, H., Shen, L. X., Chamorro, M., Varmus, H. E. & Tinoco, I., Jr (1996). A characteristic bent conformation of RNA pseudoknots promotes −1 frameshifting during translation of retroviral RNA. *J. Mol. Biol.* **260**, 479-483.

Cheong, C. & Moore, P. B. (1992). Solution structure of an unusually stable RNA tetraplex containing G- and U-quartet structures. *Biochemistry,* **31**, 8406-8414.

Cheong, C., Varani, G. & Tinoco, I., Jr (1990). Solution structure of an unusually stable RNA hairpin, 5′GGAC(UUCG)GUCC. *Nature,* **346**, 680-682.

Clemons, W. M., May, J. L. C., Wimberly, B. T., McCutcheon, J. P., Capel, M. S. & Ramakrishnan, V. (1999). Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature,* **400**, 833-840.

Colmenarejo, G. & Tinoco, I., Jr (1999). Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme. *J. Mol. Biol.* **290**, 119-135.

Conn, G. L., Draper, D. E., Lattmann, E. E. & Gittis, A. G. (1999). Crystal structure of a conserved ribosomal protein-RNA complex. *Science,* **284**, 1171-1174.

Convery, M. A., Rowsell, S., Stonehouse, N. J., Ellington, A. D., Hirao, I., Murray, J. B., Peabody, D. S., Phillips, S. E. V. & Stockley, P. G. (1998). Crystal structure of an RNA aptamer-protein complex at 2.8 Å resolution. *Nature Struct. Biol.* **5**, 133-139.

Correll, C. C., Freeborn, B., Moore, P. B. & Steitz, T. A. (1997). Metals, motifs, and recognition in the crystal structure of a 5 S rRNA domain. *Cell,* **91**, 705-712.

Correll, C. C., Munishkin, A., Chan, Y.-L., Ren, Z., Wool, I. G. & Steitz, T. A. (1998). Crystal structure of the ribosomal RNA domain essential for binding elongation factors. *Proc. Natl Acad. Sci. USA,* **95**, 13436-13441.

Costa, M. & Michel, F. (1995). Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14**, 1276-1285.

Crothers, D. M. & Cole, P. E. (1978). Conformational changes in tRNA. In *Transfer RNA* (Altman, S., ed.), pp. 196-247, MIT Press, Cambridge, MA.

Cusack, S. (1999). RNA-protein complexes. *Curr. Opin. Struct. Biol.* **9**, 66-73.

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

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

Dallas, A. & Moore, P. B. (1997). The loop E-loop D region of *Escherichia coli* 5 S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins. *Structure,* **5**, 1639-1653.

De Guzman, R. N., Turner, R. B. & Summers, M. F. (1998). Protein-RNA recognition. *Biopolymers: Nucl. Acids Sci.* **48**, 181-195.

De los Santos, C., Rosen, M. & Patel, D. J. (1989). NMR studies of DNA $(R+)_n\cdot(Y-)_n\cdot(Y+)_n$ triple helices in solution: imino and amino proton markers of T·A·T and C·G·C$^+$ base-triple formation. *Biochemistry,* **28**, 7282-7289.

Deme, E., Nolte, A. & Jacquier, A. (1999). Unexpected metal ion requirements specific for catalysis of the branching reaction in a group II intron. *Biochemistry,* **38**, 3157-3167.

Diener, J. L. & Moore, P. B. (1998). Solution structure of a substrate for the archaeal pre-tRNA splicing endonucleases: the bulge-helix-bulge motif. *Mol. Cell,* **1**, 883-894.

Draper, D. E. & Misra, V. K. (1998). RNA shows its metal. *Nature Struct. Biol.* **5**, 927-930.

Draper, D. E. & Reynaldo, L. P. (1999). RNA binding strategies of ribosomal proteins. *Nucl. Acids Res.* **27**, 381-388.

Earnshaw, D. J., Masquida, B., Müller, S., Sigurdsson, S. T., Eckstein, F., Westhof, E. & Gait, M. J. (1997). Inter-domain cross-linking and molecular modelling of the hairpin ribozyme. *J. Mol. Biol.* **274**, 197-212.

Edmonds, C. G., Crain, P. F., Gupta, R., Hashizume, T., Hocart, C. H., Kowalak, J. A., Pomerantz, S. C., Stetter, K. O. & McCloskey, J. A. (1991). Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria). *J. Bacteriol.* **173**, 3138-3148.

Egli, M., Portmann, S. & Usman, N. (1996). RNA hydration: a detailed look. *Biochemistry,* **35**, 8489-8494.

Eguchi, Y. & Tomizawa, J. (1990). Complex formation by complementary RNA stem-loops and its stabilization by a protein: function of *ColE1* Rom protein. *Cell,* **60**, 199-209.

Famulok, M. (1999). Oligonucleotide aptamers that recognize small molecules. *Curr. Opin. Struct. Biol.* **9**, 324-329.

Fan, P., Suri, A. K., Fiala, R., Live, D. & Patel, D. J. (1996). Molecular recognition in the FMN-RNA aptamer complex. *J. Mol. Biol.* **258**, 480-500.

Feig, A. L. & Uhlenbeck, O. C. (1999). The role of metal ions in RNA biochemistry. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 287-319, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Feigon, J., Dieckmann, T. & Smith, F. W. (1996). Aptamer structures from A to ζ. *Chem. Biol.* **3**, 611-617.

Ferré-D'Amaré, A. R. & Doudna, J. A. (1999). RNA folds: insights from recent crystal structures. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 57-73.

Ferré-D'Amaré, A. R., Zhou, K. & Doudna, J. A. (1998). Crystal structure of a hepatitis delta virus ribozyme. *Nature,* **395**, 567-574.

Frankel, A. D. & Smith, C. A. (1998). Induced folding in RNA-protein recognition: more than a simple handshake. *Cell,* **92**, 149-151.

Fresco, J. R., Adams, A., Ascione, R., Henley, D. & Lindahl, T. (1966). Tertiary structure in transfer ribonucleic acids. *Cold Spring Harbor Symp. Quant. Biol.* **31**, 527-537.

Gautheret, D. & Gutell, R. R. (1997). Inferring the conformation of RNA base-pairs and triples from patterns of sequence variation. *Nucl. Acids Res.* **25**, 1559-1564.

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

Gautheret, D., Konings, D. & Gutell, R. R. (1995). G·U base-pairing motifs in ribosomal RNA. *RNA,* **1**, 807-814.

Gehring, K., Leroy, J. L. & Gueron, M. (1993). A tetrameric DNA structure with protonated cytosine·cytosine base-pairs. *Nature,* **363**, 561-565.

Gesteland, R. F. & Atkins, J. F. (1993). Editors of *The RNA World*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Gesteland, R. F., Cech, T. R. & Atkins, J. F. (1999). Editors of *The RNA World*, 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. (1998). A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science,* **282**, 259-264.

Gonzalez, R. L., Jr & Tinoco, I., Jr (1999). Solution structure and thermodynamics of a divalent metal ion binding site in an RNA pseudoknot. *J. Mol. Biol.* **289**, 1267-1282.

Greenbaum, N. L., Radhakrishnan, I., Patel, D. J. & Hirsh, D. (1996). Solution structure of the donor site of a trans-splicing RNA. *Structure,* **4**, 725-733.

Grosjean, H. & Benne, R. (1998). Editors of *Modification and Editing of RNA*, ASM Press, Washington DC.

Hermann, T. (2000). Strategies for the design of drugs targeting RNA and RNA-protein complexes. *Angew. Chem. Int. Edit.* **in the press**.

Hermann, T. & Westhof, E. (1998a). Exploration of metal ion binding sites in RNA folds by Brownian-dynamics simulations. *Structure,* **6**, 1303-1314.

Hermann, T. & Westhof, E. (1998b). RNA as a drug target: chemical, modelling, and evolutionary tools. *Curr. Opin. Biotechnol.* **9**, 66-73.

Hermann, T., Auffinger, P., Scott, W. G. & Westhof, E. (1997). Evidence for a hydroxide ion bridging two magnesium ions at the active site of the hammerhead ribozyme. *Nucl. Acids Res.* **25**, 3421-3427.

Heus, H. A. & Pardi, A. (1991). Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops. *Science,* **253**, 191-194.

Heus, H. A., Wijmenga, S. S., Hoppe, H. & Hilbers, C. W. (1997). The detailed structure of tandem G·A mismatched base-pair motifs in RNA duplexes is context dependent. *J. Mol. Biol.* **271**, 147-158.

Hilbers, C. W., Michiels, P. J. A. & Heus, H. A. (1998). New developments in structure determination of pseudoknots. *Biopolymers,* **48**, 137-153.

Holbrook, S. R. & Kim, S. H. (1997). RNA crystallography. *Biopolymers: Nucl. Acids Sci.* **4**, 3-21.

Holbrook, S. R., Sussman, J. L., Warrant, R. W., Church, G. M. & Kim, S. H. (1977). RNA-ligand interactions. I. Magnesium binding sites in yeast $tRNA^{Phe}$. *Nucl. Acids Res.* **4**, 2811-2820.

Holbrook, S. R., Cheong, C., Tinoco, I., Jr & Kim, S. H. (1991). Crystal structure of an RNA double helix incorporating a track of non-Watson-Crick base-pairs. *Nature,* **353**, 579-581.

Holland, J. A. & Hoffmann, D. W. (1996). Structural features and stability of an RNA triple helix in solution. *Nucl. Acids Res.* **24**, 2841-2848.

Hoogstraten, C. G., Legault, P. & Pardi, A. (1998). NMR solution structure of the lead-dependent ribozyme: evidence for dynamics in RNA catalysis. *J. Mol. Biol.* **284**, 337-350.

Huang, S., Wang, Y.-X. & Draper, D. E. (1996). Structure of a hexanucleotide RNA hairpin loop conserved in ribosomal RNAs. *J. Mol. Biol.* **258**, 308-321.

Hud, N. V., Smith, F. W., Anet, F. & Feigon, J. (1996). The selectivity for $K^+$ *versus* $Na^+$ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR. *Biochemistry,* **35**, 15383-15390.

Ippolito, J. A. & Steitz, T. A. (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.

Jack, A., Ladner, J. E., Rhodes, D., Brown, R. S. & Klug, A. (1977). A crystallographic study of metal-binding to yeast phenylalanine transfer RNA. *J. Mol. Biol.* **111**, 315-328.

Jaeger, L., Michel, F. & Westhof, E. (1994). Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J. Mol. Biol.* **236**, 1271-1276.

James, B. D., Olsen, G. J., Liu, J. & Pace, N. R. (1988). The secondary structure of ribonuclease P RNA, the catalytic element of a ribonucleoprotein enzyme. *Cell,* **52**, 19-26.

Jang, S. B., Hung, L. W., Chi, Y. I., Holbrook, E. L., Carter, R. J. & Holbrook, S. R. (1998). Structure of an RNA internal loop consiting of tandem C-$A^+$ base-pairs. *Biochemistry,* **37**, 11726-11731.

Jiang, L. & Patel, D. J. (1998). Solution structure of the tobramycin-RNA aptamer complex. *Nature Struct. Biol.* **5**, 769-774.

Jiang, L., Suri, A. K., Fiala, R. & Patel, D. J. (1997). Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex. *Chem. Biol.* **4**, 35-50.

Jiang, L., Majumdar, A., Hu, W., Jaishree, T. J., Xu, W. & Patel, D. J. (1999). Saccharide-RNA recognition in a complex formed between neomycin B and an RNA aptamer. *Structure,* **7**, 817-827.

Jossinet, F., Paillart, J.-C., Westhof, E., Hermann, T., Skripkin, E., Lodmell, J. S., Ehresmann, C.,

Ehresmann, B. & Marquet, R. (1999). Dimerization of HIV-1 genomic RNA of subtypes A and B: RNA loop structure and magnesium binding. *RNA,* **5**, 1222-1234.

Jucker, F. M. & Pardi, A. (1995). GNRA loops make a U-turn. *RNA,* **1**, 219-222.

Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. & Pardi, A. (1996). A network of heterogeneous hydrogen bonds in GNRA tetraloops. *J. Mol. Biol.* **264**, 968-980.

Kambach, C., Walke, S. & Nagai, K. (1999). Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Curr. Opin. Struct. Biol.* **9**, 222-230.

Kang, C., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992). Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature,* **356**, 126-131.

Kettani, A., Bouaziz, S., Gorin, A., Zhao, H., Jones, R. A. & Patel, D. J. (1998). Solution structure of a Na cation stabilized DNA quadruplex containing G·G·G·G and G·C·G·C. tetrads formed by G-G-G-C repeats observed in adeno-associated viral DNA. *J. Mol. Biol.* **282**, 619-636.

Kieft, J. S. & Tinoco, I., Jr (1997). Solution structure of a metal binding site in the major groove of RNA complexed with cobalt(III) hexammine. *Structure,* **5**, 713-721.

Kim, S.-H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H., Seeman, N. C. & Rich, A. (1974). Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science,* **185**, 435-440.

Klinck, R., Liquier, J., Taillandier, E., Gouyette, C., Huynh-Dinh, T. & Guittet, E. (1995). Structural characterization of an intramolecular RNA triple helix by NMR spectroscopy. *Eur. J. Biochem.* **233**, 544-553.

Kolk, M., Heus, H. A. & Hilbers, C. W. (1997). The structure of the isolated, central hairpin of the HDV antigenomic ribozyme: novel structural features and similarity of the loop in the ribozyme and free in solution. *EMBO J.* **16**, 3685-3692.

Kolk, M., van der Graaf, M., Wijmenga, S. S., Pleij, C. W. A., Heus, H. A. & Hilbers, C. W. (1998). NMR structure of a classical pseudoknot: interplay of single- and double-stranded RNA. *Science,* **280**, 434-438.

Kuimelis, R. G. & McLaughlin, L. W. (1998). Mechanisms of ribozyme-mediated RNA cleavage. *Chem. Rev.* **98**, 1027-1044.

Laing, L. G., Gluick, T. C. & Draper, D. E. (1994). Stabilization of RNA structure by Mg ions. Specific and non-specific effects. *J. Mol. Biol.* **237**, 577-587.

Laughlan, G., Murchie, A. I., Norman, D. G., Moore, M. H., Moody, P. C., Lilley, D. M. & Luisi, B. (1994). The high-resolution crystal structure of a parallel-stranded guanine tetraplex. *Science,* **265**, 520-524.

Lee, A. J. & Crothers, D. M. (1998). The solution structure of an RNA loop-loop complex: the *ColE1* inverted loop sequence. *Structure,* **6**, 993-1005.

Legault, P., Hoogstraten, C. G., Metlitzky, E. & Pardi, A. (1998a). Order, dynamics and metal-binding in the lead-dependent ribozyme. *J. Mol. Biol.* **284**, 325-335.

Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. (1998b). NMR structure of the bacteriophage λ N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell,* **93**, 289-299.

Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. (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. Biol.* **3**, 993-1009.

Leonard, G. A., McAuley-Hecht, K. E., Ebel, S., Lough, D. M., Brown, T. & Hunter, W. N. (1994). Crystal and molecular structure of r(CGCGAAUUAGCG): an RNA duplex containing two G(anti)·A(anti) base-pairs. *Structure,* **2**, 483-494.

Leontis, N. B. & Westhof, E. (1998a). The 5 S rRNA loop E - chemical probing and phylogenetic data versus crystal structure. *RNA,* **4**, 1134-1153.

Leontis, N. B. & Westhof, E. (1998b). A common motif organizes the structure of multi-helix loops in 16 S and 23 S ribosomal RNAs. *J. Mol. Biol.* **283**, 571-583.

Lietzke, S. E., Barnes, C. L., Berglund, J. A. & Kundrot, C. E. (1996). The structure of an RNA dodecamer show how tandem U-U base-pairs increase the range of stable RNA structures and the diversity of recognition sites. *Structure,* **4**, 917-930.

Lilley, D. M. J. (1999). Structure, folding and catalysis of the small nucleolytic ribozymes. *Curr. Opin. Struct. Biol.* **9**, 330-338.

Malhotra, A. & Harvey, S. A. (1994). A quantitative model of the *Escherichai coli* 16 S rRNA in the 30 S ribosomal subunit. *J. Mol. Biol.* **240**, 308-340.

Marino, J. P., Gregorian, R. S., Jr, Csankovszki, G. & Crothers, D. M. (1995). Bent helix formation between RNA hairpins with complementary loops. *Science,* **268**, 1448-1454.

Masquida, B. & Westhof, E. (1999). Crystallographic structures of RNA oligoribonucleotides and ribozymes. In *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., ed.), pp. 533-565, Oxford University Press, Oxford.

Masquida, B., Felden, B. & Westhof, E. (1997). Context dependent RNA-RNA recognition in a three-dimensional model of the 16 S rRNA core. *Bioorg. Med. Chem.* **5**, 1021-1035.

Massire, C., Jaeger, L. & Westhof, E. (1998). Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. *J. Mol. Biol.* **279**, 773-793.

McKay, D. B. & Wedekind, J. E. (1999). Small ribozymes The role of metal ions in RNA biochemistry. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 265-286, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Michel, F. & Westhof, E. (1990). Modelling of the three-dimensional architecture of group I introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585-610.

Michel, F. & Westhof, E. (1996). Visualizing the logic behind RNA self-assembly. *Science,* **273**, 1676-1677.

Moore, P. B. (1998). The three-dimensional structure of the ribosome and its components. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 35-58.

Moore, P. B. (1999). Structural motifs in RNA. *Annu. Rev. Biochem.* **67**, 287-300.

Mueller, F. & Brimacombe, R. (1997). A new model for the three-dimensional folding of *Escherichia coli* 16 S ribosomal RNA. *J. Mol. Biol.* **271**, 524-544.

Mueller, U., Schubel, H., Sprinzl, M. & Heinemann, U. (1999). Crystal structure of acceptor stem of tRNA(Ala) from *Escherichia coli* shows unique G·U wobble base-pair at 1. 16 Å resolution. *RNA,* **5**, 670-677.

Murray, J. B., Terwey, D. P., Maloney, L., Karpeisky, A., Usman, N., Beigelman, L. & Scott, W. G. (1998). The structural basis of hammerhead ribozyme self-cleavage. *Cell,* **92**, 665-673.

Nowakowski, J. & Tinoco, I., Jr (1999). RNA structure in solution. In *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., ed.), pp. 567-602, Oxford University Press, Oxford.

Pan, T. & Uhlenbeck, O. C. (1992). A small metalloribozyme with a two-step mechanism. *Nature,* **358**, 560-563.

Pan, B., Mitra, S. N. & Sundaralingam, M. (1998). Structure of a 16-mer RNA duplex r(GCAGACUUAAAUCUGC)$_2$ with wobble C·A$^+$ mismatches. *J. Mol. Biol.* **283**, 977-984.

Pan, B., Mitra, S. N. & Sundaralingam, M. (1999). Crystal structure of an RNA 16-mer duplex r(GCAGAGUUAAAUCUGC)2 with nonadjacent G(syn)·A$^+$(anti) mispairs. *Biochemistry,* **38**, 2826-2831.

Patel, D. J. (1997). Structural analysis of nucleic acid aptamers. *Curr. Opin. Chem. Biol.* **1**, 32-46.

Patel, D. J. (1999). Adaptive recognition in RNA complexes with peptides and protein modules. *Curr. Opin. Struct. Biol.* **9**, 74-87.

Patel, D. J., Suri, A. K., Jiang, F., Jiang, L., Fan, P., Kumar, R. A. & Nonin, S. (1997). Structure, recognition and adaptive binding in RNA aptamer complexes. *J. Mol. Biol.* **272**, 645-664.

Pleij, C. W., Rietveld, K. & Bosch, L. (1985). A new principle of RNA folding based on pseudoknotting. *Nucl. Acids Res.* **13**, 1717-1731.

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

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994b). 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.

Portmann, S., Grimm, S., Workman, C., Usman, N. & Egli, M. (1996). Crystal structure of an *A*-form duplex with single-adenosine bulges and a conformational basis for site-specific RNA self-cleavage. *Chem. Biol.* **3**, 173-184.

Puglisi, J. D. & Williamson, J. R. (1999). RNA interaction with small ligands and peptides. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 403-425, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Puglisi, J. D., Wyatt, J. R. & Tinoco, I., Jr (1988). A pseudoknotted RNA oligonucleotide. *Nature,* **331**, 283-286.

Puglisi, J. D., Chen, L., Blanchard, S. & Frankel, A. D. (1995). Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science,* **270**, 1200-1203.

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

Quigley, G. J. & Rich, A. (1976). Structural domains of transfer RNA molecules. The ribose 2′ hydroxyl which distinguishes RNA from DNA plays a key role in stabilizing tRNA structure. *Science,* **194**, 796-806.

Quigley, G. J., Teeter, M. M. & Rich, A. (1978). Structural analysis of spermine and magnesium ion binding to yeast phenylalanine transfer RNA. *Proc. Natl Acad. Sci. USA,* **75**, 64-68.

Rajagopal, P. & Feigon, J. (1989). Triple-strand formation in the homopurine:homopyrimidine DNA oligonucleotides d(G-A)$_4$ and d(T-C)$_4$. *Nature,* **339**, 637-640.

Rife, J. P. & Moore, P. B. (1998). The structure of a methylated tetraloop in 16 S ribosomal RNA. *Structure,* **6**, 747-756.

Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C. & Klug, A. (1974). Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature,* **250**, 546-551.

Saenger, W. (1984). *Principles of Nucleic Acid Structure*, Springer, New York.

Schindelin, H., Zhang, M., Bald, R., Fürste, J. P., Erdmann, V. A. & Heinemann, U. (1995). Crystal structure of an RNA dodecamer containing the *Escherichia coli* Shine-Dalgarno sequence. *J. Mol. Biol.* **249**, 595-603.

Schmitz, U., James, T. L., Lukavsky, P. & Walter, P. (1999). Structure of the most conserved internal loop in SRP RNA. *Nature Struct. Biol.* **6**, 634-638.

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

Scott, W. G., Finch, J. T. & Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. *Cell,* **81**, 991-1002.

Scott, W. G., Murray, J. B., Arnold, J. R. P., Stoddard, B. L. & Klug, A. (1996). Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. *Science,* **274**, 2065-2069.

Smith, D. & Pace, N. R. (1993). Multiple magnesium ions in the ribonuclease P reaction mechanism. *Biochemistry,* **32**, 5273-5281.

Sontheimer, E. J., Gordon, P. M. & Piccirilli, J. A. (1999). Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev.* **13**, 1729-1741.

Steitz, T. A. (1999). RNA recognition by proteins. In *The RNA World* (Gesteland, R. F., Cech, T. R. & Atkins, J. F., eds), 2nd edit., pp. 427-450, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Steitz, T. A. & Steitz, J. A. (1993). A general two-ion-mechanism for catalytic RNA. *Proc. Natl Acad. Sci. USA,* **90**, 6498-6502.

Streicher, B., Westhof, E. & Schroeder, R. (1996). The environment of two metal ions surrounding the splice site of a group I intron. *EMBO J.* **15**, 2556-2564.

Strobel, S. A. (1999). A chemogenetic approach to RNA function/structure analysis. *Curr. Opin. Struct. Biol.* **9**, 346-352.

Strobel, S. A. & Doudna, J. A. (1997). RNA seeing double: close-packing of helices in RNA tertiary structure. *Trends Biochem. Sci.* **22**, 262-266.

Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. (1999). Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nature Struct. Biol.* **6**, 285-292.

Szewczak, A. A. & Moore, P. B. (1995). The sarcin/ricin loop, a modular RNA. *J. Mol. Biol.* **247**, 81-98.

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

Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, M., Stormo, G., D'Aubenton-Carafa, Y., Uhlenbeck, O. C. & Tinoco, I., Jr (1988). CUUCGG hairpins: extraordinarily stable RNA secondary

structures associated with various biochemical processes. *Proc. Natl Acad. Sci. USA,* **85**, 1364-1368.
Valegard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature,* **371**, 623-626.
Varani, L., Hasegawa, M., Spillantini, M. G., Smith, M. J., Murrell, J. R., Ghetti, B., Klug, A., Goedert, M. & Varani, G. (1999). Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and parkinsonism linked to chromosome 17. *Proc. Natl Acad. Sci. USA,* **96**, 8229-8234.
Viani-Puglisi, E., Green, R., Noller, H. F. & Puglisi, J. D. (1997). Structure of a conserved RNA component of the peptidyl transferase center. *Nature Struct. Biol.* **4**, 775-778.
Wang, Y. X., Lu, M. & Draper, D. E. (1993). Specific ammonium ion requirement for functional ribosomal RNA tertiary structure. *Biochemistry,* **32**, 12279-12282.
Wedekind, J. E. & McKay, D. B. (1999). Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nature Struct. Biol.* **6**, 261-268.
Weeks, K. M. & Crothers, D. M. (1993). Major groove accessibility of RNA. *Science,* **261**, 1574-1577.
Westhof, E. (1988). Water: an integral part of nucleic acid structure. *Annu. Rev. Biophys. Biophys. Chem.* **17**, 125-144.
Westhof, E. & Michel, F. (1998). Ribozyme architectural diversity made visible. *Science,* **282**, 251-252.
Westhof, E., Dumas, P. & Moras, D. (1985). Crystallographic refinement of yeast aspartic acid transfer RNA. *J. Mol. Biol.* **184**, 119-145.
Westhof, E., Dumas, P. & Moras, D. (1988). Hydration of transfer RNA molecules: a crystallographic study. *Biochimie,* **70**, 145-165.
Wimberly, B., Varani, G. & Tinoco, I., Jr (1993). The conformation of loop E of eucaryotic 5 S ribosomal RNA. *Biochemistry,* **32**, 1078-1087.
Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W. & Ramakrishnan, V. (1999). A detailed view of a ribosomal active site: the structure of the L11-RNA complex. *Cell,* **97**, 491-502.
Woese, C. R., Winker, S. & Gutell, R. R. (1990). Architecture of ribosomal RNA: constraints on the sequence of ''tetra-loops''. *Proc. Natl Acad. Sci. USA,* **87**, 8467-8471.
Wyatt, J. R. & Tinoco, I., Jr (1993). RNA structural elements and RNA function. In *The RNA World* (Gesteland, R. F. & Atkins, J. F., eds), pp. 465-496, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Wyatt, J. R., Puglisi, J. D. & Tinoco, I., Jr (1990). RNA pseudoknots: stability and loop size requirements. *J. Mol. Biol.* **214**, 455-470.
Ye, X., Kumar, R. A. & Patel, D. J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem. Biol.* **2**, 827-840.
Ye, X., Gorin, A., Ellington, A. D. & Patel, D. J. (1996). Deep penetration of an α-helix into a widened RNA major groove in the HIV-1 rev peptide-RNA aptamer complex. *Nature Struct. Biol.* **3**, 1026-1033.
Yonath, A. & Franceschi, F. (1998). Functional universality and evolutionary diversity: insights from the structure of the ribosome. *Structure,* **6**, 679-684.
Zimmermann, G. R., Jenison, R. D., Wick, C. L., Simorre, J.-P. & Pardi, A. (1997). Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nature Struct. Biol.* **6**, 644-649.

***Edited by P. E. Wright***

*(Received 18 August 1999; received in revised form 14 October 1999; accepted 14 October 1999)*