J. Mol. Biol. (1997) **272**, 645–664

JMB

REVIEW ARTICLE

Structure, Recognition and Adaptive Binding in RNA Aptamer Complexes

Dinshaw J. Patel*, Asif K. Suri, Feng Jiang, Licong Jiang, Pei Fan R. Ajay Kumar and Sylvie Nonin

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

*Corresponding author

Novel features of RNA structure, recognition and discrimination have been recently elucidated through the solution structural characterizion of RNA aptamers that bind cofactors, aminoglycoside antibiotics, amino acids and peptides with high affinity and specificity. This review presents the solution structures of RNA aptamer complexes with adenosine monophosphate, flavin mononucleotide, arginine/citrulline and tobramycin together with an example of hydrogen exchange measurements of the base-pair kinetics for the AMP-RNA aptamer complex. A comparative analysis of the structures of these RNA aptamer complexes yields the principles, patterns and diversity associated with RNA architecture, molecular recognition and adaptive binding associated with complex formation.

Keywords: RNA aptamer complexes; cofactor; amino acid and aminoglycoside antibiotic ligands; sequence specific recognition and discrimination; NMR based solution structures

Introduction

The coding, information transfer and catalytic activities associated with RNA makes it unique amongst biological macromolecules (reviewed by Gesteland & Atkins, 1993). RNA adopts a range of folding topologies reflecting its functional diversity associated with its participation in an array of biological phenomena (reviewed by Nagai & Mattaj, 1994). Structural studies have offered an opportunity to characterize the complex architectural scaffolds adopted by RNA molecules and the details of ligand binding sites associated with catalytic activity. The structural characterization of ligand-RNA aptamer complexes represents a promising approach to broaden this knowledge data base through identification of new patterns of RNA folding and recognition. This approach is based on *in vitro* selection and evolution techniques (Ellington & Szostak, 1990; Tuerk & Gold, 1990; Robertson & Joyce, 1990) that identify RNA aptamers from random sequence RNA libraries that target ligands ranging from cofactors to peptides, proteins and saccharides (reviewed by Joyce, 1994; Gold *et al.*, 1995; Lorsch & Szostak, 1996). RNA aptamers selected in this manner have exhibited binding affinities in the μM to nM range and the ability to discriminate between a targeted ligand and its closely related counterparts.

Many of these RNA aptamers are in the 30 to 50-mer range which puts their ligand complexes well within the capabilities of solution structure determination by heteronuclear multidimensional nuclear magnetic resonance (NMR) methodologies (Varani & Tinoco, 1991; Moore, 1995; Pardi, 1995; Dieckmann & Feigon, 1994) on uniformly ^{13}C,^{15}N-labeled RNA (Batey *et al.*, 1992; Nikonowicz *et al.*,1992). Further, the RNA conformations are frozen in defined compact conformations when bound to high affinity ligands resulting in narrower NMR resonances characterized by readily interpretable NMR spectral parameters. Recently, NMR approaches have been successfully applied to solve the solution structures of a range of ligand-RNA aptamer complexes that include the cofactors ATP (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996) and FMN (Fan *et al.*, 1996), the amino acids L-arginine and L-citrulline (Yang *et al.*, 1996), the aminoglycoside antibiotic tobramycin (Jiang *et al.*, 1997b) and the basic peptide from the HIV-1 Rev protein (Ye *et al.*, 1996).

0022–2836/97/400645–20 $25.00/0/mb971281

RNA Aptamer Complexes

This review focuses on the solution structures of RNA aptamer complexes with cofactors, amino acids, aminoglycoside antibiotics and peptides (see also earlier minireview by Feigon *et al.*, 1996). In addition, it outlines the results and implications of more recent NMR based hydrogen exchange measurements on the AMP-RNA aptamer complex.

AMP-RNA aptamer complex

Sassanfar & Szostak (1993) applied the *in vitro* selection approach to identify an RNA secondary fold (Figure 1(a)) that exhibits μM affinity for ATP attached to an affinity column through its C^8 base position. The consensus segment identified amongst the isolated RNA aptamer sequences was restricted to an asymmetric internal loop containing an 11-residue purine rich segment (G7 to G17, Figure 1(a)) positioned opposite a single guanine (G34, Figure 1(a)) suggesting that this motif constitutes the ATP binding site. This RNA aptamer fold also binds AMP but discriminates against the other three nucleoside triphosphates. This ATP rich binding motif in turn has been used by Lorsch & Szostak (1994b) to engineer ribozymes with polynucleotide kinase activity. Our group (Jiang *et al.*, 1996a) and Feigon's laboratory (Dieckmann *et al.*, 1996) have reported on the NMR based solution structure of the AMP-RNA aptamer complex using somewhat different stem segments which flank the asymmetric internal loop. The contributions from the two laboratories are in agreement on the key conclusions related to the RNA folding topology of the AMP binding site and the intermolecular hydrogen bonding and stacking alignments associated with molecular recognition. In additon, our group (Nonin *et al.*, 1997) has extended these structural studies to hydrogen exchange measurements of the resolved imino protons of the stem and loop segments in the complex.

The imino proton NMR spectra (Jiang *et al.*, 1996b) of the free 40-mer RNA aptamer and its 1:1 AMP complex in 10 mM phosphate containing H_2O solution (pH 6.7) at 0°C are plotted in Figure 1(b) and (c), respectively. The spectrum of the free RNA aptamer exhibits guanine and uracil imino protons from the stems in contrast to the asymmetric internal loop imino protons which exhibit broad overlapped resonances at 10.5 ppm in the 40-mer sequence (Figure 1(b)). By contrast, ten additional narrow exchangeable protons are observed between 9.5 and 14.5 ppm on AMP-RNA aptamer complex formation (Figure 1(c)). Further, the exchangeable protons in the complex are sufficiently narrow and well resolved to undertake a heteronuclear multidimensional NMR characterization of the complex. Exchangeable and non-exchangeable proton assignments have been successfully undertaken on complexes containing either uniformly ^{13}C,^{15}N-labeled RNA aptamer or uniformly ^{13}C,^{15}N-labeled AMP. The assignment of guanine imino protons in the asymmetric internal loop and their correlation to their non-exchangeable proton counterparts in the two studies (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996) was achieved by a combination of specific [^{15}N]-guanine (Jiang *et al.* 1997a), specific inosine (Jiang *et al.*, 1997a) and specific adenine (Dieckmann *et al.*, 1996) substitution for individual guanine residues in the asymmetric internal loop and following utilization of through bond relay correlation methods (Fiala

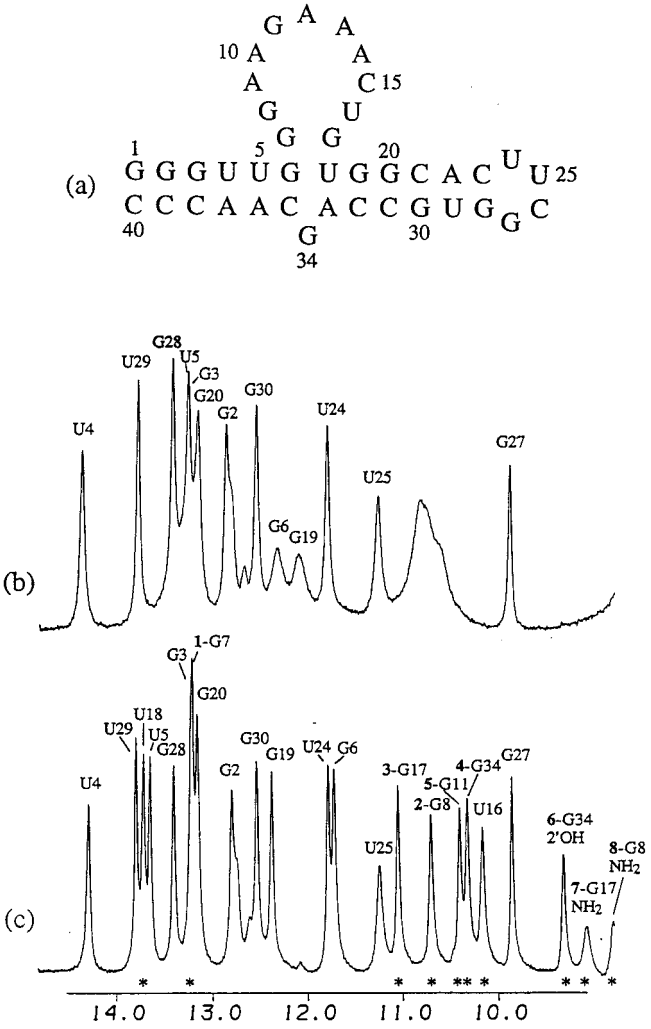

Figure 1. (a) Sequence and numbering system of the AMP-binding 40-mer RNA aptamer. Imino proton NMR spectra (8.8 to 14.8 ppm) of (b) the free RNA aptamer and (c) the AMP-RNA aptamer complex (≈20% excess AMP) in 10 mM sodium phosphate containing H_2O solution at pH 6.7 and 0°C. The asterisks in (c) identify additional exchangeable resonances which appeared upon addition of AMP. The imino proton resonance assignments are indicated above the spectra. The additional guanine imino and amino proton resonances and a 2′-OH proton resonance of the asymmetric internal loop that appear on complex formation are designated by numbers in bold face. (Reproduced from Jiang *et al.* (1996b) with permission of *Biochemistry*.)

et al., 1996; Sklenar *et al*., 1996; Simorre *et al*., 1996). In addition, a narrow exchangeable proton is detected at 9.34 ppm (Figure 1(c)) which has been assigned to the 2′-OH of G34 in the spectrum of the complex. Both groups (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) used a combination of NMR and molecular dynamics to solve the solution structure of the AMP-RNA aptamer complex with the computations guided by ≈45 intermolecular restraints. There was no bias in the starting RNA fold of the asymmetric internal loop in the complex since the computations were initiated using distance geometry protocols in both studies. The pairwise root mean square deviation (r.m.s.d.) values for the well defined core (residues G6-A12, U16-U18, A33-C35 for the asymmetric internal loop sequence in Figure 1(a) and AMP) amongst the distance refined structures of the complex are in the ≈2.0 Å range.

The folding topology of the AMP binding site reported from our laboratory (Jiang *et al*., 1996a) is shown in Figure 2(a). The two helical stems spanning the asymmetric internal loop AMP-binding site are aligned in an approximate orthogonal (≈106°) alignment. The 11-residue purine rich segment forms an S-shaped fold with the 5′-portion (G7 to A12) involved primarily in AMP complex formation while the 3′-portion (A13 to G17) forms a scaffold which closes one face of the binding pocket (Figure 2(b)). The purine ring of AMP is intercalated between the purine rings of A10 and G11 and positioned in an in-plane alignment opposite G8 in the solution structure of the complex (Figure 3(a)). The G7 to G11 segment in the complex adopts a fold (Figure 3(a)) reminiscent of a GNRA (N is any base and R is a purine) hairpin loop closed by a base-pair (Woese *et al*., 1990; Heus & Pardi, 1991; Pley *et al*., 1994). There is extensive stacking amongst the purine rings in this G7-(G8-A9-A10-AMP)-G11 segment except for the G8-A9 step which is involved in chain reversal (Figure 3(a)). In essence, the leftward stem is extended through stacking with the G7·G11 mismatch pair and further into the loop along the G7-G8 and G11-AMP-A10-A9 segments (Figure 3(a)).

This pattern of stacking between adjacent purines is broken in a striking manner for the G11-A12 and A12-A13 steps where the purines are aligned in mutually orthogonal planes in the solution structure of the complex (Figure 3(b)). The A12 residue forms the innermost core of the complex and its orthogonal alignment relative to G11 reflects A12 being part of the stacking extension of the rightward stem while G11 is part of the stacking extension of the leftward stem. Both G11 and A12 residues adopt C2′-*endo* pucker geometries in the complex. The A13 residue forms part of the stacked A13 to G17 segment, forming a scaffold (residues A13 to U16 in yellow and G17 in orange in Figure 2(b)) that prevents entry of the AMP ligand from that face into the binding pocket.

The G34 residue, positioned opposite the 11-residue purine rich segment within the asymmetric

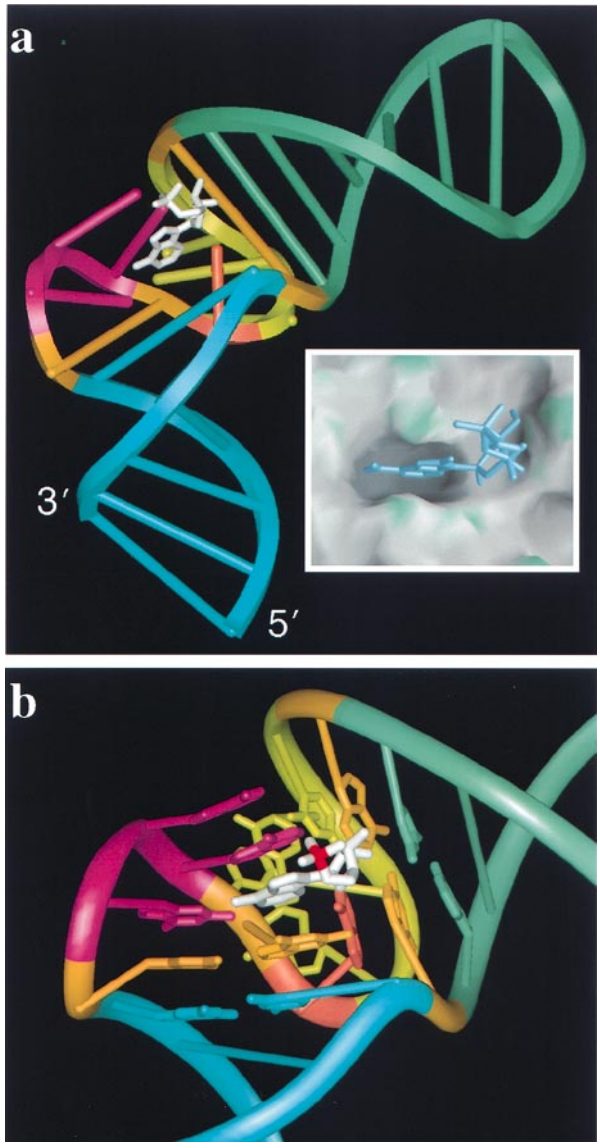

Figure 2. Ribbon representation of the solution structure of the AMP-RNA aptamer complex. The bound AMP is shown in a stick representation in white. The RNA ribbon is colored to distinguish key residues. The left stem (G1 to G6 and C35 to C40) and the right stem-hairpin loop (U18 to A33) are colored cyan and green, respectively. The G7·G11 and G17·G34 mismatch pairs are colored orange. Residues G8 to A10, which together with the bound AMP, constitute the GNRA-like structrual motif are colored magenta. Resisue A12 is colored rusty red. Less well-structured residues A13 to U16 are colored yellow. The bound AMP molecule is shown in white. (a) The entire complex. The two helical stems make an angle of 106(±9)° with the ATP-binding loop folding into two distinct hairpins. The insert shows a GRASP view (Nicholls *et al*., 1991) of the binding of the AMP in the RNA pocket. The RNA accessible surface measured with a probe of radius 1.4 Å around the bound AMP is shaded for curvature with convex and concave surfaces colored green and grey, respectively. The bound AMP in a stick representation is colored blue. (b) A close up view of AMP bound within the folded asymmetric internal loop in the complex. The phosphorus atom of the bound AMP is colored red. (Reproduced from Jiang *et al*. (1996a) with permission from *Nature* for (a) and from Jiang *et al*. (1996b) with permission from *Biochemistry* for (b).)

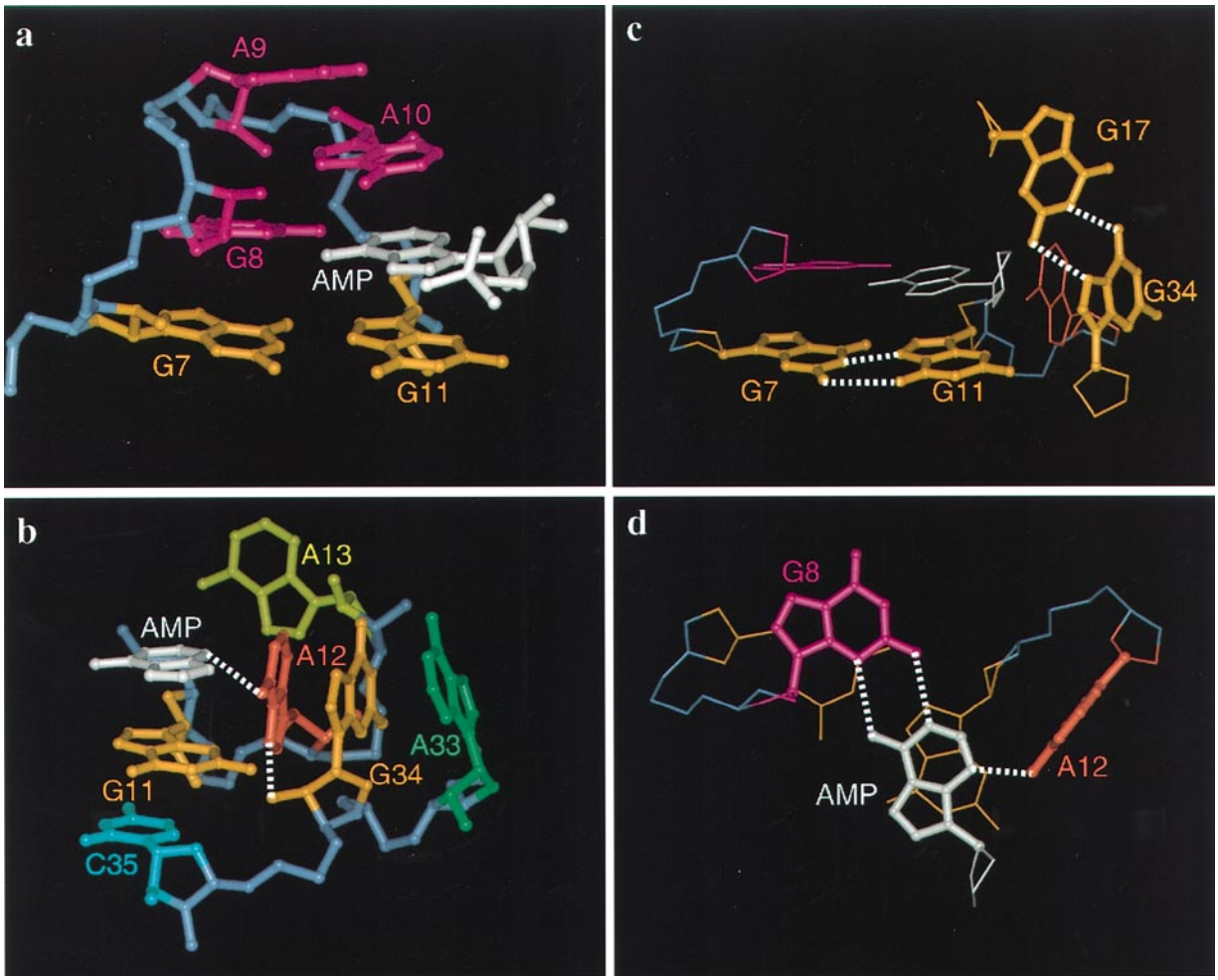

Figure 3. (a) The GNRA-like structural motif with phosphate oxygen atoms removed for clarity. Residues G8-A9-A10 (in magenta) and the base of AMP (in white) constitute the GNRA-like structural motif in the complex. A G7·G11 reversed Hoogsteen mismatch pair (in orange) closes the GNRA-like motif. (b) The purine stacked core of the AMP-binding motif with some atoms omitted for clarity. Residues AMP (in white) and G11 (in orange) stack on the leftward stem (in cyan) while residues A12 (in rusty red) and G34 (in orange) stack on the rightward stem (in green). Broken lines indicate the hydrogen bonds from A12 to the N^3 of AMP and the sugar 2′-OH of G34. Note the orthogonal alignment of the G11 and A12 purine rings and the A12 and A13 purine rings. (c) The G7(*anti*)·G11(*anti*) reversed Hoogsteen and G17(*anti*)·G34(*syn*) Hoogsteen mismatch pairs (in orange) form at the junctions of the stems and the AMP-binding loop. Broken lines indicate hydrogen bonds within the G·G mismatches. (d) The G8·AMP mismatch pairing within the G8-A9-A10-AMP GNRA-like structural motif. Pairing involves the Watson-Crick edge of AMP and the minor groove edge of G8 stabilized by two hydrogen bonds. In addition, a hydrogen bond forms between the exocyclic amino proton of A12 and the N^3 of AMP. Note that views (a) to (d) represent small modifications of the corresponding views containing chirality errors published by Jiang *et al*. (1996a).

internal loop, adopts a *syn* glycosidic bond and C2′-*endo* sugar pucker in the AMP-RNA aptamer complex (Figure 3(b)). The planes of adjacent G34 and C35 residues are orthogonal to each other and are spatially separated by intervening A12 and G11 residues in the complex (Figure 3(b)). This reflects the extension of the rightward stem through formation of the G17·G34 mismatch (Figure 3(c)) with the *syn* G34 residue sandwiched between *anti* A12 and *anti* A33 residues in the complex (Figure 3(b)). The 2′-OH proton of G34 is both narrow and downfield shifted (9.34 ppm) in the AMP-RNA aptamer complex (Figure 1(c)). This reflects its burial in the internal core of the complex, its hydrogen-bonding to the N^7 of A12 and its in-plane positioning relative to the purine rings of G11 and A12 in the complex (Figure 3(b)).

Two G·G mismatches are observed in the AMP-RNA aptamer complex and found to exhibit distinct alignments (Figure 3(c)). The G7·G11 mismatch is of the reverse Hoogsteen type with the Watson-Crick edge of *anti* G7 pairing with the Hoogsteen edge of *anti* G11 through two hydrogen bonds (Figure 3(c)). The G17·G34 mismatch is of the Hoogsteen type with the Watson-Crick edge of *anti* G17 pairing with the Hoogsteen edge of *syn* G34 through two hydrogen bonds (Figure 3(c)).

Proteins are known to contain a hydrophobic core involving packing of hydrophobic side-chains in the interior of the globular fold. Much less is

understood of what constitutes the core of RNA tertiary structures with the focus to date on the potential role of divalent ions in stabilizing the core domain. The AMP-RNA aptamer solution structure provides insights into this question since its core contains stacked purines aligned in three mutually orthogonal planes (Figure 3(b)). This pattern of purine stacking alignments constituting the core of the complex provides a plausible explanation for the invariant concentration of purines in the active site asymmetric internal loop of the AMP-RNA aptamer complex.

Molecular recognition between AMP and the RNA aptamer occurs through a G·A mismatch alignment (Figure 3(d)). The Watson-Crick edge of AMP aligns with the minor groove edge of G8 through formation of two intermolecular hydrogen bonds and to the exocyclic amino group of A12 through formation of one intermolecular hydrogen bond (Figure 3(d)). The donor functionality at the N^6 amino position and the acceptor functionalities at the N^1 and N^3 ring nitrogen positions of AMP are involved in intermolecular hydrogen bonding on complex formation (Figure 3(d)). By contrast, the N^7 position which is adjacent to the C^8 position that was covalently linked to the affinity column during the selection experiments does not participate in intermolecular hydrogen bond formation (Figure 3(d)). This G·A pairing alignment in the AMP-RNA aptamer complex (Figure 3(d)) is distinct from the sheared G·A mismatch alignment proposed previously for GNRA hairpin loops based on modeling studies (Westhof *et al.*, 1989; Jaeger *et al.*, 1994) and observed experimentally in solution (Heus & Pardi, 1991) and crystalline (Pley *et al.*, 1994) states.

The ribose ring of the AMP aligns against the G17·G34 mismatch and U18·A33 Watson-Crick base-pairs positioned along the minor groove edge of the right helical segment in the complex (Figure 2b). The 2′- and 3′-ribose hydroxyls of AMP can form intermolecular hydrogen bonds with the minor groove of the extended right helical segment in the complex. The phosphate group of AMP is directed away from the RNA and exposed to solvent. The RNA aptamer contacts about 50% of the surface area of the bound AMP ligand in the complex.

The solution structure of the AMP-RNA aptamer complex (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996) outlines a range of unique features associated with RNA folding and molecular recognition. Sequence and structure specific targeting of this RNA aptamer involves a combination of mismatch pairing and stacking interactions in order to generate an intercalation cavity that involves an extremely stable GNRA-like hairpin fold with molecular recognition associated with formation of a novel G·A mismatch alignment. A key principle that emerges is the identification of a purine ring stacked core associated with the binding site in the AMP-RNA aptamer complex. A common pattern observed in the complex is associated with the extension of the left and right helical stems through G·G mismatch formation and additional base stacking into the core of the complex. The diversity of interactions is reflected in the distinct pairing alignments adopted by the G7·G11 and G17·G34 mismatches in the complex. The phosphate group of AMP is exposed and available for catalysis permitting the rational design of ribozymes where the availability of a high energy phosphodiester bond is paramount for driving catalysis.

The structural studies on the AMP-RNA aptamer complex have been complemented in our laboratory by monitoring the exchange characteristics (reviewed by Gueron & Leroy, 1995) of the well resolved imino protons in the NMR spectra of the free 40-mer RNA aptamer (Figure 1(b)) and its AMP complex (Figure 1(c)) (Nonin *et al.*, 1997). The imino proton exchange kinetics have been monitored as a function of pH (Figure 4(a)) and added ammonia catalyst (Figure 4(b)) to measure the apparent base-pair dissociation constants (αK_d) of the Watson-Crick and mismatched base-pairs along with the solvent accessibility of the unpaired imino protons in the complex. These αK_d values at individual pairs in the free RNA aptamer and the AMP-RNA aptamer complex are summarized graphically in Figure 4(c). These data establish that complex formation stabilizes the αK_d values for bases within the AMP-binding asymmetric internal loop and also for the two base-pairs of the stem segments on either side of the binding site. Thus, the αK_d value for the G6·C35 base-pair is 5×10^{-3} in the free duplex and decreases to 3.6×10^{-7} in the complex (Figure 4(c)). The αK_d values span the range 10^{-2} to 10^{-7} for the mismatch base-pairs in the asymmetric internal loop of the AMP-RNA aptamer complex. Thus, the αK_d values within the G17·G34 mismatch pair as monitored by the imino protons of G17 (hydrogen-bonded to O^6 of G34 within the mismatch pair) and G34 (directed outwards towards the phosphate oxygen of A14) are 4.3×10^{-7} and 1.6×10^{-2}, respectively in the AMP-RNA aptamer complex (Figure 4(c)). The proton exchange and base-pair kinetics characteristics are consistent with the published solution structures of the AMP-RNA aptamer complex (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996). The kinetic studies also establish that the conformation of the asymmetric internal loop in the complex is most likely stabilized by four additional hydrogen bonds involving the imino protons of G8, G11, U16 and G34 (Nonin *et al.*, 1997) that could be readily accomodated with minor adjustment in the published solution structure of the complex.

The structural (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996) and hydrogen exchange (Nonin *et al.*, 1997) measurements establish that complex formation between AMP and its RNA aptamer involves adaptive binding. Thus, the imino protons in the asymmetric internal loop exhibit broad rapidly exchanging resonances centered between 10 and 11 ppm in the free RNA aptamer (Figure 1(b)) in striking contrast to the complex

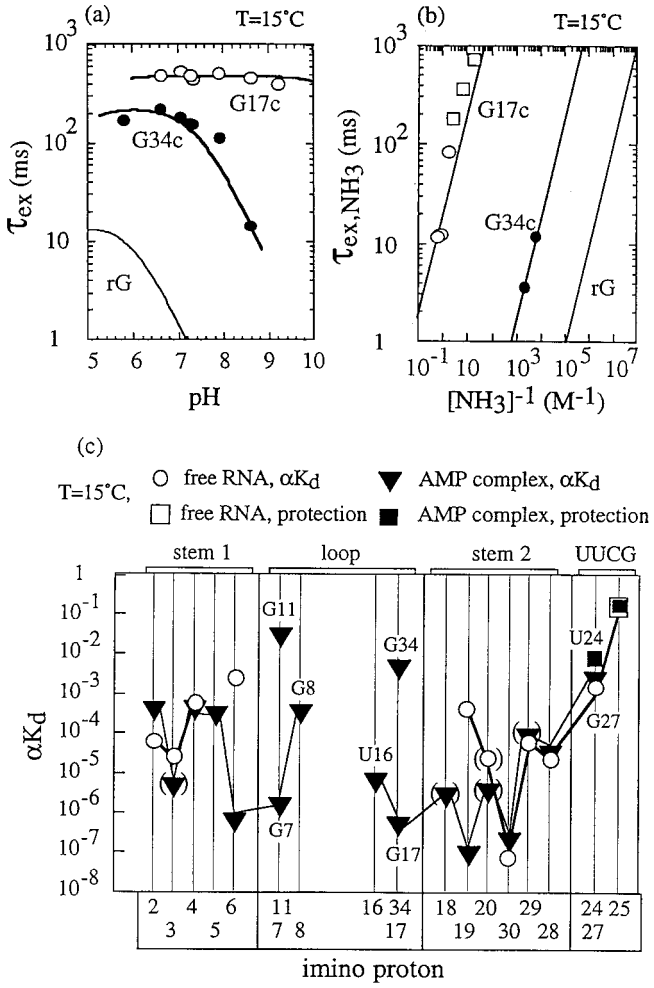

Figure 4. Exchange times of imino protons of G17 and G34 of the G17·G34 mismatch pair in the asymmetric internal loop of the AMP-RNA aptamer complex as a function of (a) pH and (b) the inverse of the ammonia concentration in H_2O at 15°C. Also shown is the corresponding data for rG monomer under the same conditions. Circles and squares relate to data collected with one and four equivalents of AMP bound per RNA aptamer, respectively. (c) Plot of the apparent dissociation constant (αK_d) values (circles and triangles) and the protection factors (squares) as a function of the base or base-pair position in the free RNA aptamer (open symbols) and in the AMP-RNA aptamer complex (filled symbols). αK_d values derived from the measurement of exchange times from poorly resolved imino proton resonances are represented by bracketed symbols. The αK_d values of internal stem Watson-Crick base-pairs exhibit values similar to those reported previously for *B*-DNA. The two base-pairs flanking either side of the AMP-binding asymmetric internal loop are stabilized on complex formation. The αK_d values of imino protons involved in pairing within the mismatch pairs fall in the same range as the corresponding values of stem Watson-Crick pairs with the values of G7 and G17 amongst the most stable in the complex. (Reproduced from Nonin *et al.* (1997) with permission from the *J. Mol. Biol.*)

(Figure 1(c)) where the same imino protons exhibit narrow signals, are dispersed in chemical shift and exhibit low αK_d values, three of which are amongst the lowest in the entire AMP-RNA aptamer complex. Thus, an apparently unstructured asymmetric internal loop in the free RNA aptamer becomes highly structured on complex formation with bound AMP which targets this site.

FMN-RNA aptamer complex

Oxidation-reduction reactions can be mediated by the cofactor flavin mononucleotide (FMN) (Figure 5(a)) when bound as a prosthetic group in flavoproteins. Redox reactions catalyzed by these enzymes are mediated by the isoalloxazine ring of FMN which serves as a transient carrier of a pair of hydrogen atoms abstracted from substrates. The previous research on FMN-protein complexes has now been extended to FMN-RNA complexes through the application of *in vitro* selection methods that have identified RNA aptamers that bind this cofactor with high affinity and specificity (Burgstaller & Famulok, 1994; Lauhon & Szostak, 1995). These efforts have been directed towards the eventual characterization of ribozymes with diverse redox catalytic activities.

The FMN-RNA aptamer complex identified by Burgstaller & Famulok (1994) contains a consensus purine rich asymmetric internal loop with six residues 5′-A-G-G-N-U-A-3′ (where N can be any residue) positioned opposite five residues 5′-G-A-A-G-G-3′ across the loop. Solution structure studies undertaken in our laboratory on the complex of FMN with its high affinity (≈0.5 μM) 35-mer RNA aptamer sequence (Figure 5(b)) have defined the RNA fold at the FMN binding site and the principles associated with FMN-RNA recognition (Fan *et al.*, 1996).

The imino proton NMR spectra (9.5 to 14.5 ppm) of the RNA aptamer and its 1:1 FMN complex in 150 mM NaCl, 4 mM Mg^{2+} containing H_2O solution at (pH 6.5) and 5°C are plotted in Figure 5(c) and (d), respectively. The imino proton spectrum of the complex is exceptionally well resolved with eight additional exchangeable resonances observable on complex formation (Figure 5d). The exchangeable and non-exhangeable resonances in the complex have been assigned from homonuclear and heteronuclear NMR studies on FMN complexes with unlabeled and uniformly ^{13}C,^{15}N-labeled RNA aptamers. The imino and FMN exchangeable proton assignments listed over the spectrum of the complex are based in part on inosine for guanine substitution experiments and establish that all guanine imino protons except G27 (which exhibits rapid exchange) within the asymmetric internal loop give well resolved resonances in the spectrum of the complex (Figure 5(d)). The analysis of the NMR data yielded distance restraints between RNA protons across the asymmetric internal loop with the largest number centered between the G24-A25 and G10-U12-A13 segments in the complex. A total of 18 intermolecular distance restraints were identified between the

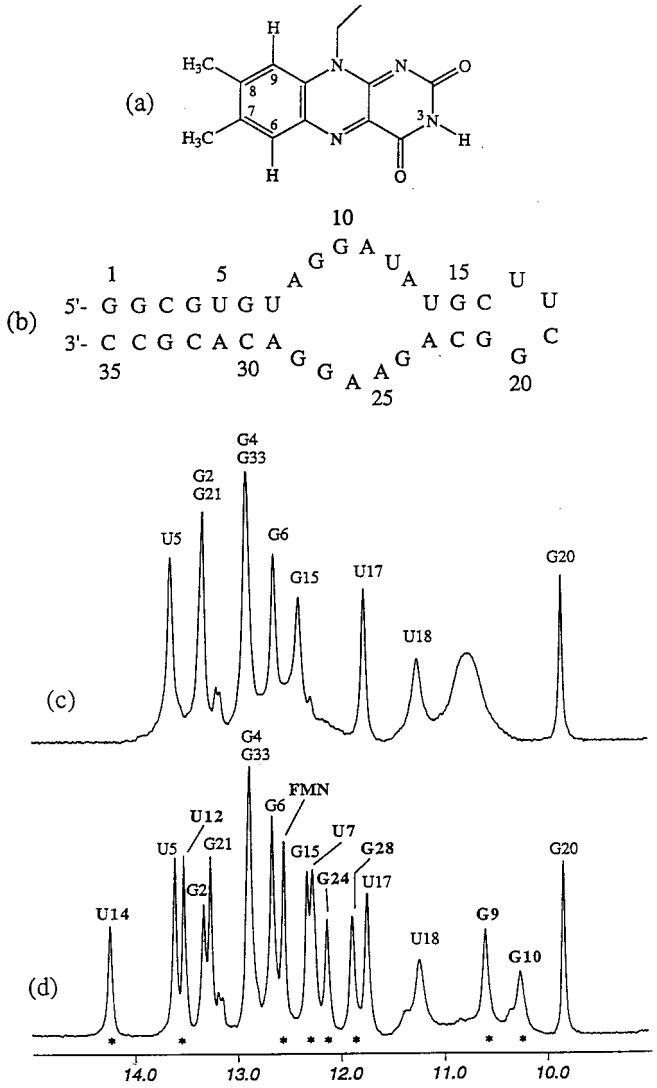

Figure 5. (a) Chemical formula and numbering system of the isoalloxazine ring of flavin mononucleotide (FMN). (b) Sequence and numbering of the FMN-binding 35-mer RNA aptamer. Imino proton NMR spectra (9.0 to 15.0 ppm) of (c) the free RNA aptamer and (d) the 1:1 FMN-RNA aptamer complex in 150 mM NaCl, 10 mM phosphate and 5 mM Mg^{2+} containing H_2O solution at pH 6.5 and 5°C. The imino proton resonance assignments are indicated above the spectra. RNA aptamer imino protons that appear on FMN complex formation are marked by asterisks in (d) and assignments are labeled in bold face. (Reproduced from Fan *et al.* (1996) with permission from the *J. Mol. Biol.*)

isoalloxazine chromophore of FMN and RNA residues G9, G10 and U12 on one side and A25, A26 and G27 on the opposite side of the asymmetric internal loop in the complex. Structure calculations were initiated from different asymmetric internal loop conformations with the FMN molecule placed in randomized orientations at the center of masses of each RNA molecule. The distance refined structures of the FMN-RNA aptamer complex exhibited pairwise r.m.s.d. values of ≈1.25 Å for the asymmetric internal loop RNA binding site (residues A8 to A13 excluding A11 and G24 to G28) and FMN, which represents the core of the complex.

The folding toppology of the FMN binding site in the asymmetic internal loop of the RNA aptamer for one representative refined structure of the complex is plotted in Figure 6(a). The FMN binding site is generated by zippering up the asymmetric internal loop through base mismatch and triple formation with intercalation of the isoalloxazine chromophore of FMN between a G·G mismatch and a G·U·A triple (Figure 6(a)). The flanking stem regions and zippered-up internal loop form a continuous helix except for the non-conserved A11 residue which loops out of the helix to facilitate formation of the G10·U12·A25 base triple in the complex (Figure 6(a)).

The separation between strands is widened at the FMN binding site due to formation of *anti* purine·*anti* purine A8·G28 and G9·G27 mismatch pairs and the G10·U12·A25 base triple. This facilitates insertion of the isoalloxazine ring into the helix with its long axis parallel to the long axis of the flanking G9·G27 mismatch pair resulting in stabilizing stacking interactions (Figure 6(a)). Binding specificity is associated with formation of two intermolecular hydrogen-bonds between the uracil-like edge of the isoalloxazine ring and the Hoogsteen edge of *anti* A26 in the complex (Figure 6(b)). The side-chain of FMN, which is poorly defined in the refined structures, is positioned in the minor groove.

The G10·U12·A25 base triple (Figure 6(c)) was an unexpected feature observed in the solution structure of the FMN-RNA aptamer complex. Residues G10 and A25 on opposite sides of the asymmetric internal loop are too far apart to hydrogen bond in the solution structure of the complex. The U12 residue acts as a hydrogen bonding bridge between G10 and A25 resulting in base triple formation (Figure 6(c)). The participation of G10 and U12 residues in the G10·U12·A25 base triple is facilitated by the looping out of the intervening non-conserved A11 residue into the major groove of the RNA helix (Figure 6(a)). The U12·A25 alignment is of the reversed Hoogsteen type stabilized by two hydrogen bonds while the U12·G10 interaction is stabilized by a single hydrogen bond with the Watson-Crick edge of G10 interacting in addition with the backbone phosphate oxygen atoms of the G24-A25 step (Figure 6(c)). The FMN·A26 recognition pair is positioned over the G10·U12·A25 base triple platform with extensive stacking between A25 and A26, as well as between G10 and U12 with the isoalloxazine ring of FMN in the complex.

The zippering-up of the asymmetric internal loop about the FMN molecule in the complex is achieved through formation of two types of G·A mismatches at both loop-stem junctions. The A8·G28 mispair involves alignment through the Watson-Crick edges of both purines (Figure 6(d))

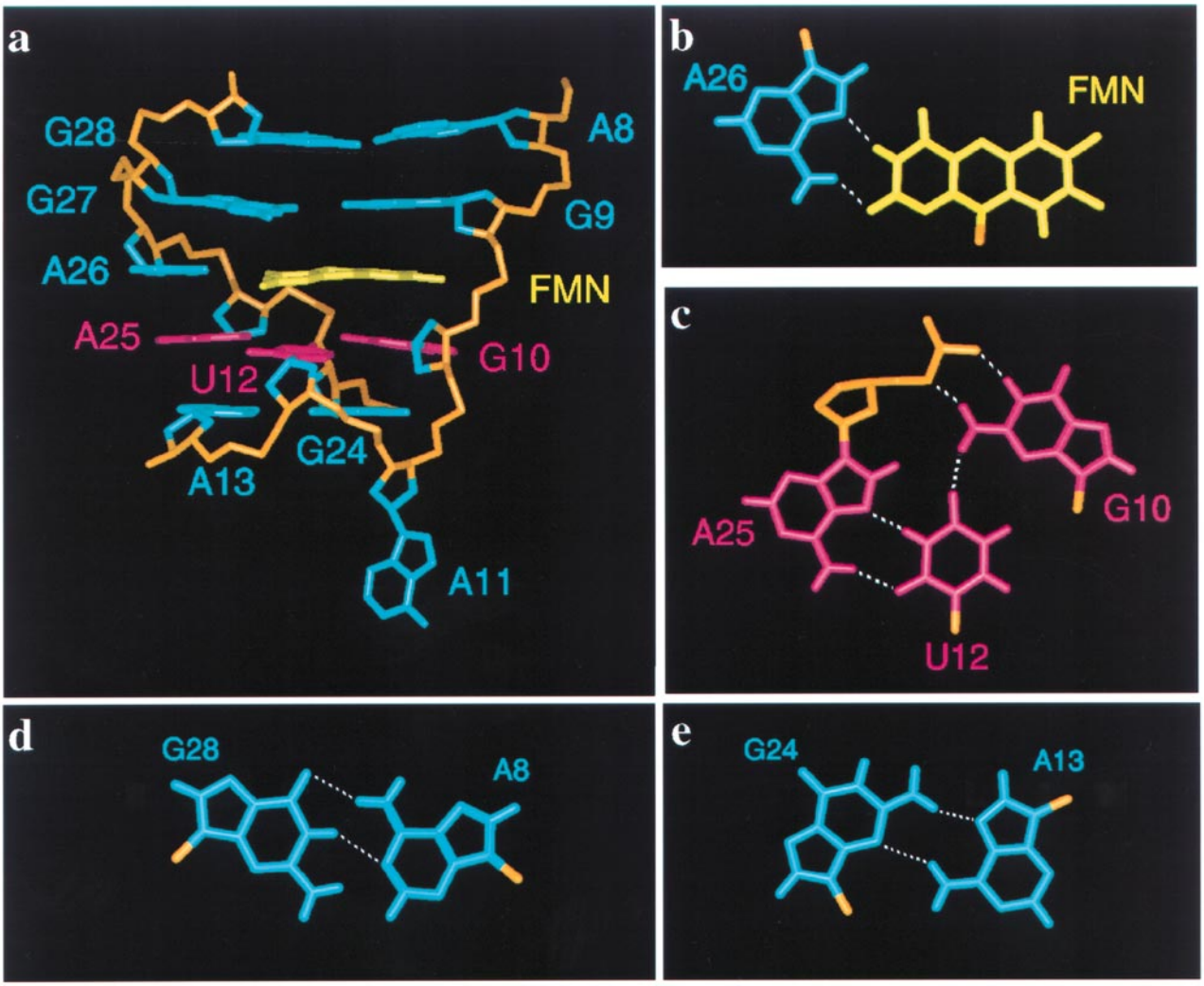

Figure 6. Views of segments of one representative refined structure of the FMN-RNA aptamer complex. The base mismatches and looped out base are colored cyan, the base triple is colored magenta, the backbone is colored orange and the bound isoalloxazine ring of FMN is colored yellow. (a) View looking into the minor groove of the FMN-bound internal loop segment (A8 to A13 and G24 to G28) in the complex. (b) The A26(*anti*)·FMN alignment stabilized by two hydrogen bonds involving the Hoogsteen edge of A26. (c) The G10·U12·A25 triple involving a reversed Hoogsteen U12(*anti*)·A25(*anti*) mispair and G10(*anti*) pairing with U12 and the phosphate oxygen atoms at the G24-A25 step. (d) The A8(*anti*)·G28(*anti*) mismatch pair involving a pair of hydrogen bonds between the Watson-Crick edges of A8 and G28. (e) The A13(*anti*)·G24(*anti*) sheared mismatch pair involving a pair of hydrogen bonds between the major groove edge of A13 and the minor groove edge of G24. (Reproduced from Fan *et al.* (1996) with permission from the *J. Mol. Biol.*)

while the A13·G24 mispair is of the sheared type (Li *et al.*, 1991; Heus & Pardi, 1991) and involves alignment of the minor groove edge of guanine and the major groove edge of adenine (Figure 6(e)).

L-Arginine/L-citrulline-RNA aptamer complexes and ligand discrimination

The molecular basis associated with the ability of RNA aptamers to discriminate between closely related ligands has been probed by structural studies of RNA aptamers that bind L-citrulline (Figure 7(a)) and L-arginine (Figure 7(b)) with high affinity and specificity. Initially, *in vitro* selection studies identified an L-citrulline-binding RNA aptamer (Figure 6(c)) which was subsequently mutagenized and following further selection evolved into an L-arginine-binding RNA aptamer (Figure 6(d)) (Famulok, 1994). The two RNA aptamers contained similar secondary structure folds defined by two base-pair separated asymmetric internal loops flanked at either end by stem segments. The conserved residues span the G9 to G14 (except A11) and C28 to C39 (except U36) segments on opposing strands centered about the larger asymmetric internal loop within each 33-mer RNA aptamer. The L-arginine- and L-citrulline-binding RNA aptamers which differed solely at three nucleotide positions bound their cognate amino acids with 10 μM affinity and without detectable affinity for their non-cognate counterparts. A common ligand binding site centered

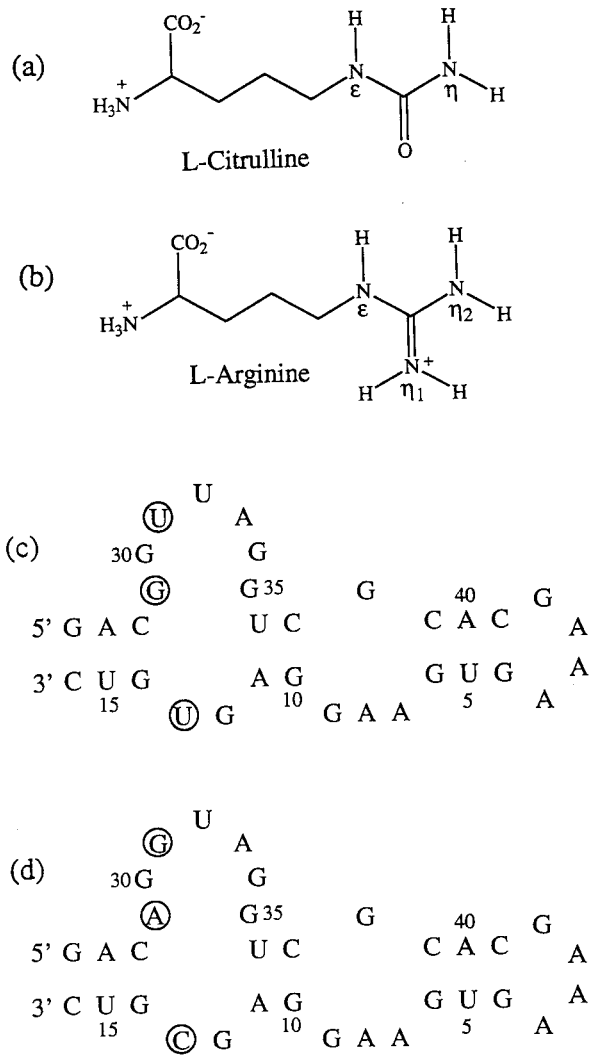

Figure 7. Chemical formula and numbering system of (a) L-citrulline and (b) L-arginine. Sequence and numbering of (c) the L-citrulline-binding RNA aptamer and (d) the L-arginine-binding RNA aptamer.

about the larger asymmetric internal loop was identified from chemical and enzymatic footprinting experiments (Burgstaller *et al.*, 1995).

The solution structures of the L-citrulline-RNA aptamer and L-arginine-RNA aptamer complexes have been determined following a combined NMR-molecular dynamics study (Yang *et al.*, 1996). The unambiguous assignment of exchangeable and nonexchangeable protons in the L-arginine- and L-citrulline-binding RNA aptamer complexes with their pair of asymmetric internal loops poses a considerable challenge (Yang *et al.*, 1996), similar to the one faced in the related studies of the AMP-RNA aptamer complex (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996). This problem was approached by studying complexes containing unlabeled ligand and uniformly ^{15}N-labeled RNA aptamers and through comparitive analysis of the NMR spectral characteristics of the L-arginine- and L-citrulline-RNA aptamer complexes. A set of exchangeable and non-exchangeable proton assignments have been reported for both complexes (Yang *et al.*, 1996). These assignments need confirmation along the lines of single residue isotope labeling and base substitution approaches (Jiang *et al.*, 1997a), the application of relay methods to correlate exchangeable protons with their non-exchangeable counterparts (Fiala *et al.*, 1996; Sklenar *et al.*, 1996; Simorre *et al.*, 1996), as well as the collection and analysis of multinuclear multidimensional NMR data sets (Pardi, 1995) on both complexes. Structure determinations were based on ≈50 long range distance restraints between the RNA aptamer protons and ≈16 intermolecular restraints for both L-arginine- and L-citrulline-RNA aptamer complexes. The computations were also guided by four intermolecular hydrogen-bonding restraints for the L-arginine-RNA aptamer complex and six intermolecular hydrogen-bonding restraints for the L-citrulline-RNA aptamer complex at later stages of the calculations (Yang *et al.*, 1996). The average pairwise r.m.s.d. values amongst distance refined structures were 0.63 Å for base atoms and 1.20 Å for backbone atoms of the RNA binding site segment spanning residues G9 to G14 and C28 to C39 in the L-citrulline-RNA aptamer complex. Similar average pairwise r.m.s.d. values were also observed for the L-arginine-RNA aptamer complex (Yang *et al.*, 1996).

The smaller asymmetric internal loop zippers up with continuous stacking involving adjacent rightward and central stem segments through formation of a bridging A7(*anti*)·G38(*anti*) mismatch involving alignment of Watson-Crick edges and was facilitated through looping out of A8 and G9 residues on complex formation. The central helical stem is extended through formation of a G12(*anti*)·G35(*syn*) mismatch pair involving alignment of the Watson-Crick edge of G12 and the Hoogsteen edge of G35 in both complexes. Additional stacking alignments relate the leftward stem terminating in the G14·C28 base-pair and the extended central/rightward stems terminating in the G12·G35 mismatch pair and involve continuous stacking between the C28, A/G29, G9 and A11 residues and between the G14, C/U13 and G12 residues.

The ligand binding site in both complexes involves participation primarily by RNA aptamer residues G12, C/U13, A/G29, G30 (*syn*), G/U31, A33 and G35 (*syn*) of the larger asymmetric internal loop. The aliphatic side-chains of the arginine and citrulline ligands stack over a platform generated by the G12·G35 mismatch and the G9 base in both complexes. Recognition of L-arginine by its RNA aptamer containing ligand-specific C13, A29 and G31 residues involves intermolecular hydrogen bond formation, specifically involving side-chain $NH_2(\eta)$ and $NH(\varepsilon)$ protons of the ligand and acceptor atoms along the major groove edges of G30 and G31 and the Watson-Crick edge of C13 in the complex (Figure 8(a)). Recognition of

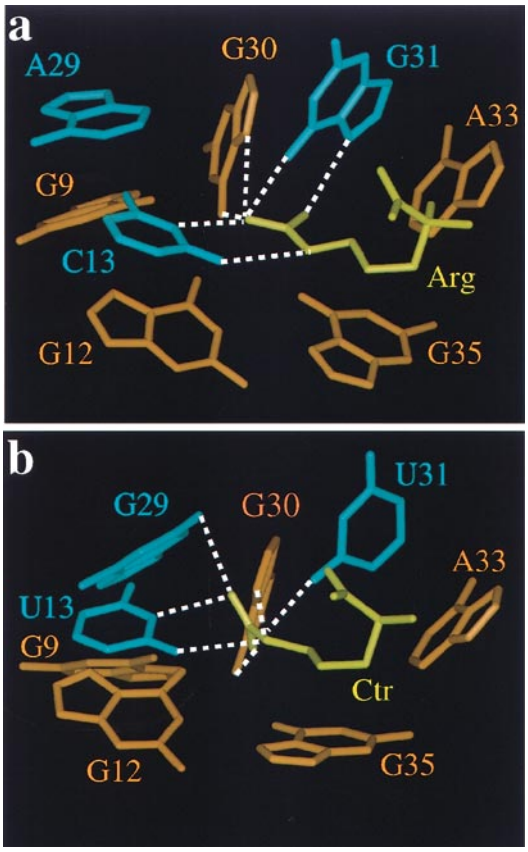

Figure 8. One representative refined structure of the binding pocket in (a) the L-arginine-RNA aptamer complex and (b) the L-citrulline-RNA aptamer complex. The amino acids are shown in yellow, the bases of the binding pocket common to both RNA aptamers (G9, G12, G30, A33 and G35) are shown in orange and the ligand specific bases of the binding pocket characteristic of each RNA aptamer (C/U13, A/G29 and G/U31) are shown in cyan. Intermolecular hydrogen bonds are shown by broken white lines. (Adapted from coordinates in Yang *et al.* (1996).)

L-citrulline by its RNA aptamer containing ligand-specific U13, G29 and U31 residues involves intermolecular hydrogen bond formation involving side-chain $NH_2(\eta)$ and NH(ε) protons and carbonyl (ε) oxygen of the ligand and polar functionalities on the Watson-Crick edges of U13, G29 and U31 in the complex (Figure 8(b)). The discrimination between L-arginine and L-citrulline can be readily rationalized since donor $NH_2(\eta)$ protons on L-arginine which hydrogen bond to the acceptor atoms at N^3 of C13, O^6 and N^7 of G30 and O^6 of G31 on its RNA aptamer (Figure 8(a)) are replaced by the acceptor carbonyl (ε) oxygen functionality on L-citrulline which hydrogen bonds to the donor atoms at N^1H of U13 and N^2H_2 of G29 on its RNA aptamer (Figure 8(b)).

The amino acid side-chains of L-arginine and L-citrulline constitute the core of their RNA aptamer complexes with the folded RNA surrounding the bound ligands. Molecular recognition and discrimination are associated with defined intermolecular hydrogen bonding alignments in the L-arginine and L-citrulline-RNA aptamer complexes. The L-arginine and L-citrulline ligands were covalently linked to the column through their peptide amine groups (Famulok, 1994) and it is therefore not surprising that the peptide functionalites are accessible to solvent in the solution structures of both RNA aptamer complexes (Yang *et al.*, 1996). Both complexes represent examples of adaptive binding since chemical footprinting (Burgstaller *et al.*, 1995) and the NMR parameters (Yang *et al.*, 1996) establish substantial conformational changes in the RNA aptamers on complex formation.

Tobramycin-RNA aptamer complex

Aminoglycoside antibiotics are amino-modified saccharide containing antimicrobial agents which came into clinical prominence because of their ability to block protein synthesis. They can bind to a variety of RNA targets with sequence specificity ranging from functional sites on 16 S ribosomal RNA (Moazed & Noller, 1987; Recht *et al.*, 1996), to catalytic group I intron RNAs (von Ahsen *et al.*, 1991), the hammerhead ribozyme (Clouet-d'Orval *et al.*, 1995) and the Rev protein binding site on the Rev response element of HIV-1 RNA (Zapp *et al.*, 1993). The aminoglycoside antibiotics are polycationic at neutral pH and it appears likely that ionic interactions make potentially important contributions to the binding affinity associated with complex formation with the RNA polyanion (Clouet-d'Orval *et al.*, 1995).

The binding affinities of aminoglycoside antibiotics to their biological target RNAs are in the μM range. This has stimulated efforts to identify RNA aptamers under highly stringent selection conditions in order to increase the binding affinity into the nM range while retaining binding selectivity (Wang & Rando, 1995; Lato *et al.*, 1995; Wallis *et al.*, 1995). One such system involves the aminoglycoside antibiotic tobramycin (Figure 9(a)) which binds to a bulge containing stem-hairpin loop RNA aptamer (Figure 9(b)) with nM affinity (Wang & Rando, 1995).

Our understanding of the structural basis for aminoglycoside antibiotic-RNA recognition has taken a significant step forward following publication of the solution structure of paromomycin bound to a 27-nucleotide RNA containing the A-site of *Escherichia coli* 16 S ribosomal RNA from the Puglisi laboratory (Fourmy *et al.*, 1996) and the solution structure of the tobramycin-RNA aptamer complex from our laboratory (Jiang *et al.*, 1997b). The high resolution structure of the paramomycin-27-mer ribosomal RNA complex provides a wealth of information on the intermolecular contacts that account for the specificity of complex formation. Specifically, several OH and charged NH_2 groups on the paramomycin form intermolecular hydrogen bonds to base edges, sugar ring oxygen atoms and backbone phosphates to anchor the aminogly-

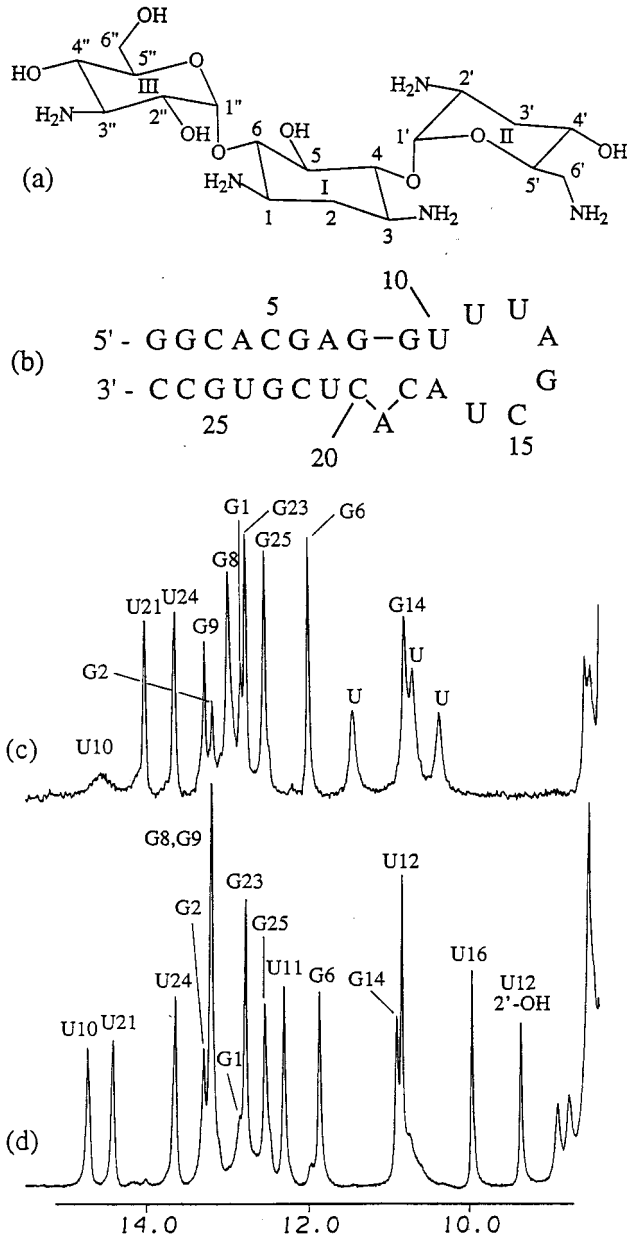

Figure 9. (a) Chemical formula and numbering system of tobramycin. The amino protons are positively charged through protonation at neutral pH. (b) Sequence and numbering system of the tobramycin-binding 27-mer RNA aptamer. Imino proton NMR spectra (8.5 to 15.5 ppm) of (c) the free RNA aptamer and (d) the 1:1 tobramycin-RNA aptamer complex in 10 mM phosphate containing H_2O buffer at pH 6.8 and 5°C. The imino proton assignments are indicated over the spectra. (Reproduced from Jiang *et al.* (1997b) with permission from *Chem. Biol.)*

coside antibiotic within its RNA major groove binding pocket created through A·A mismatch and A bulge formation (Fourmy *et al.*, 1996). This structure explains in molecular terms the wealth of biochemical data that has identified key positions on the antibiotic and conserved residues on the RNA that are critical for complex formation.

We outline below the structural features of the tobramycin-RNA aptamer complex in some detail, which despite being solved to lower resolution, better fits the thematic focus of this review on RNA aptamer complexes. Formation of the tobramycin-RNA aptamer complex can be readily monitored by recording the imino proton spectra (9.0 to 15.0 ppm) of the free RNA aptamer (Figure 9(c)) and the 1:1 complex (Figure 9d) in 10 mM phosphate containing H_2O solution (pH 6.8) at 5°C. The RNA aptamer (Figure 9(b)) contains a six residue U-U-A-G-C-U hairpin loop and an A bulge separated by two base-pairs. The U11, U12 and U16 imino protons narrow significantly and shift on complex formation as does the U10 imino proton of the loop closing U10·A17 base-pair (Figure 9(c) and (d)). In addition, a narrow exchangeable proton is detected at 9.36 ppm (Figure 9(d)) which has been assigned to the 2′-OH proton of U12 in the spectrum of the complex. The exceptional quality of the NMR spectra of the tobramycin-RNA aptamer complex (Figure 9(d)) has permitted a detailed heteronuclear multidimensional NMR characterization of the complex containing unlabeled tobramycin and uniformly ^{13}C,^{15}N-labeled RNA aptamer (Jiang *et al.*, 1997b). It was possible to extensively assign the exchangeable and non-exchangeable RNA protons in the key G8 to C20 segment of the RNA aptamer, as well as all the non-exchangeable and a few exchangeable OH and NH_2 tobramycin protons in the complex. An unfortunate limitation was the degeneracy of the chemical shifts of several tobramycin protons in the complex which could not be overcome in the absence of an isotopically labeled sample of the aminoglycoside antibiotic. This limited the number of intermolecular distance restraints to 34 to guide the molecular dynamics calculations on the tobramycin-RNA aptamer complex. The tobramycin-RNA aptamer complex was solved (Jiang *et al.*, 1997b) using a computational procedure similar to the one used to solve the FMN-RNA aptamer complex (Fan *et al.*, 1996). The resulting distance refined structures exhibited pairwise r.m.s.d. values of ≈2.0 Å for the A7 to U21 (excluding C15) RNA aptamer segment and tobramycin which reduced to ≈1.25 Å for the G9 to C18 (excluding C15) RNA aptamer segment and tobramycin. The somewhat higher r.m.s.d. values for this complex reflect a limited spread in the alignment of the bound tobramycin on its RNA aptamer binding site and was the result of our inability to use the maximum number of intermolecular restraints due to overlap amongst tobramycin protons in the complex.

A representative refined structure of the tobramycin-RNA aptamer complex, spanning the A7 to U21 segment of the RNA aptamer, is shown in Figure 10(a). Tobramycin is positioned within the major groove spanning the hairpin loop-stem junction segment of the RNA aptamer. The amino sugar rings II and III are directed towards the hairpin loop and the bulged adenine residue, respect-

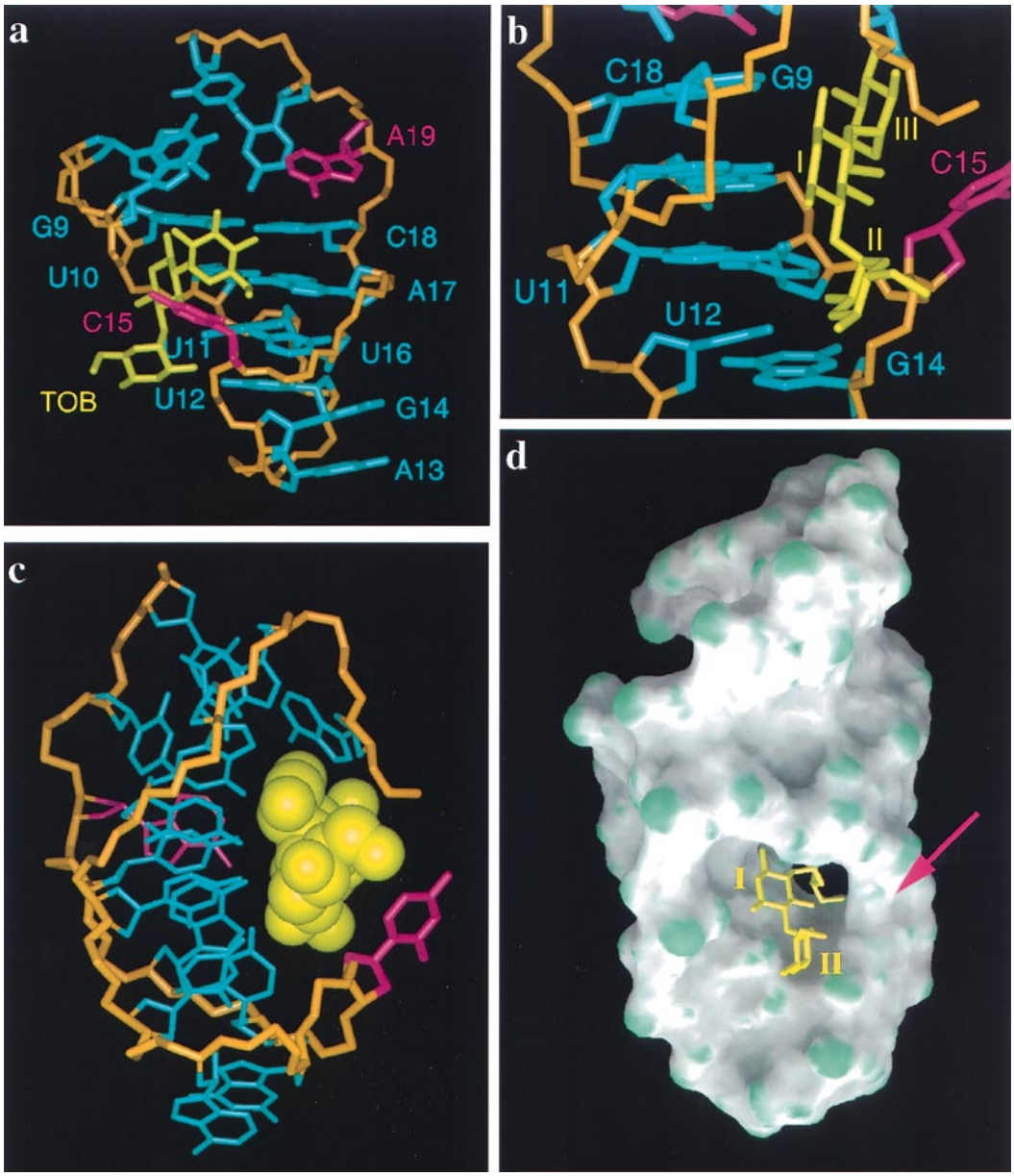

Figure 10. Views of segments of one representative refined structure of the tobramycin-RNA aptamer complex. The stacked bases are colored cyan, the C15 and A19 bases are colored magenta, the backbone is colored orange and the tobramycin is colored yellow. (a) A view looking into the major groove of the core of the complex (A7 to U21 and tobramycin). (b) A more detailed view of the molecular recognition interface between tobramycin and the RNA aptamer. (c) View looking down the major groove with tobramycin in yellow represented in a space filling surface. Note the position of C15 in magenta which acts as a flap over the bound tobramycin. (d) Surface view looking into the binding channel from the major groove and its occupancy in the complex. The RNA aptamer surface representation is displayed using GRASP (Nicholls *et al*., 1991) with concave and convex surfaces shown in green and grey, respectively. The bound tobramycin in yellow is shown in a stick representation. The position of the C15 base which acts as a flap to close the binding site is indicated by a magenta colored arrow. (Reproduced (Jiang *et al*., 1997b) with permission from *Chem. Biol*.)

ively while the aliphatic non-sugar ring I is centered about the U10·A17 base-pair in the complex (Figure 10(b)). The RNA binding site spans an unpaired base (U12), a mismatch pair (U11·U16), two Watson-Crick pairs (G9·C18 and U10·A17) and the looped out C15 base which orients itself as a flap over ring III of tobramycin in the complex (Figure 10(a) and (b)). The potential stacking between G9·C18 and C8·G20 base-pairs that flank the bulge site is perturbed as a consequence of the bulged A19 base being displaced partially into the major groove in the complex (Figure 10(a)).

All three tobramycin rings adopt a chair conformation in the complex with an average of 52% of the surface area of tobramycin being buried on complex formation. Tobramycin rings III and I interact in an approximately face down orientation with the floor of the major groove. By contrast,

tobramycin ring II lies partially in the major groove with its 2′-NH_2 group directed into this groove. There appears to be shape complementarity between the interacting tobramycin surface and the floor of the major groove in its RNA binding pocket as visualized in a view looking down the major groove of the complex (Figure 10(c)).

The U11-U12-A13-G14-U16 segment of the hairpin loop is the best defined domain of the RNA aptamer in the complex. The U11 and U16 bases form a defined U·U mismatch pair involving one of two possible Wobble alignments. The U12 base is not paired but its imino proton can hydrogen bond to the backbone phosphate oxygen at the G14-C15 step in the complex. Chain reversal occurs at A13 and the conformation is stabilized through extensive stacking between the purine rings of A13 and G14 supplemented by a hydrogen bond between the 2′-OH of U12 and the N^7 of G14 in the complex. This folding pattern for the U12-A13-G14 segment of the hairpin loop in the tobramycin-RNA aptamer complex is strikingly similar to the corresponding three residues in the anticodon and TψC hairpin loops which form U-turns in yeast $tRNA^{Phe}$ (Kim *et al*., 1974; Robertus *et al*., 1974). Even more striking, is the similarity between the global features of the six nucleotide U-U-A-G-C-U segment of the hairpin loop of the tobramycin-RNA aptamer complex and a conserved hexanucleotide hairpin loop in ribosomal RNA (Huang *et al*., 1996; Fountain *et al*., 1996).

The looped out C15 residue plays a key role in generating the tobramycin binding site since sugar ring III and to a lesser extent non-sugar ring I are sandwiched between the C15 base and the floor of the major groove in the complex. This is best visualized in a GRASP (Nicholls *et al*., 1991) view of the entire complex (Figure 10(d)) where the tobramycin is binding in a channel generated between the plane of the C15 base (magenta arrow, Figure 10(d)) and the floor of the major groove. The encapsulation of amino sugar ring III is reflected in a large number of intermolecular NOEs between protons on this ring and the base protons of C15 in the complex.

There is a spread in the relative alignment of the A19 bulge residue which adopts a conformation intermediate between stacked and looped out alignments amongst the refined structures of the complex. This in turn effects the relative orientation of the stem segments that flank the bulged adenine amongst the refined structures in the complex. A larger propeller twist is observed for the flanking G8·C20 base-pair in one direction relative to the flanking G9·C18 base-pair in the other direction in the complex.

Tobramycin was also shown to bind with high specificity and affinity to RNA aptamer analogs where the G8·C20 base-pair was replaced by the I8·C20 or A8·U20 base-pairs, the A19 bulge residue was replaced by G19, the G14 loop residue was replaced by A14 and the C15 flap residue was replaced by A15 (Jiang *et al*., 1997b). Our solution structure of the tobramycin-RNA aptamer complex provides novel global insights into the molecular recognition features associated with high affinity sequence specific binding of aminoglycoside antibiotics to their RNA targets. The RNA major groove width at the binding site is defined, in part, by the conformation of the bulged A19 residue and, in part, by the pyrimidine U11·U16 mismatch and the sharp turn at the U12-A13-G14 segment in the complex. The sequence and orientation of residues encompassing the binding site (G9·C18, U10·A17, U11·U16 pairs and U12·phosphate alignment) define the heteroatom functionalities along the base edges that line the floor of the major groove and which are targeted by the bound tobramycin. An unexpected feature of the solution structure of the complex was the establishment of a looped out C15 base which forms a flap over the binding site in the complex. The current resolution does not provide answers to whether specific intermolecular contacts exist between hydroxyl and protonated amine protons on the tobramycin and backbone phosphates, 2′-OH groups and major groove functional edges on the RNA aptamer in the tobramycin-RNA aptamer complex (Jiang *et al*., 1997b). Current efforts underway to define the solution structure of a related tobramycin-RNA aptamer complex with a different stem-loop junctional sequence (Wang & Rando, 1995) could provide the missing details associated with intermolecular contacts provided that the structure can be solved to higher resolution.

Peptide/protein-RNA aptamer complexes

There are now several RNA aptamers that have been identified based on their ability to target proteins with high affinity and specificity (reviewed by Gold *et al*., 1995). The structural characterization of these protein-RNA aptamer complexes represents a future challenge both to the crystallographic and NMR communities. A recently reported solution structural characterization of a 17-mer HIV-1 Rev peptide bound to a 35-mer Rev response element (RRE) RNA aptamer target (nM binding affinity) from our laboratory (Ye *et al*., 1996) represents an important step towards the goal of protein-RNA aptamer structure determination. The solution structure of the Rev peptide-RRE RNA aptamer complex (Ye *et al*., 1996) and a companion contribution on the Rev peptide-RRE RNA stem loop IIB complex (Battiste *et al*., 1996) establish that the Rev peptide adopts an α-helical fold along its entire length and binds in the widened major groove of the RRE RNA target site. The RNA binding pocket in these complexes adopts a unique fold defined by base mismatches, looped out bases, and in the case of the RNA aptamer complex, a U·(A·U) base triple. Both peptide and RNA undergo adaptive conformational transitions on complex formation. These structures are striking in that an isolated α-helix binds deep within the major groove of the RNA without the

need for additional structural domains to stabilize the complex. The details of the Rev peptide-RRE RNA aptamer complex (Ye *et al*., 1996) are not included in this review since this subject matter is part of a recently written companion review submitted to this journal (Patel *et al*., 1998) that compares the structures of regulatory immunodeficiency viral Tat (binds as a β-hairpin) and Rev (binds as an α-helix) peptides bound to their RNA targets with high affinity and specificity.

Principles, Patterns and Diversity

The growing data base of solution structures of RNA aptamer complexes provides new insights into the principles, patterns and diversity associated with nucleic acid architecture, as well as the intermolecular interactions that contribute to ligand recognition and discrimination by RNA motifs in solution.

Intercalative recognition of planar chromophores

The AMP and FMN ligands contain planar chromophores and their respective RNA aptamer binding site architectures involve intercalative stacking of the chromophores between either purine bases as observed for the AMP-RNA aptamer complex (Figure 3(a)) (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) or between base-pairs/triples as observed for the FMN-RNA aptamer complex (Figure 6(a)) (Fan *et al*., 1996). Specificity is associated with a pair of in-plane intermolecular hydrogen bonds between the chromophore and one of the edges of a base residue positioned opposite it such as observed with the minor groove edge of guanine in the AMP-RNA aptamer complex (Figure 3(d)) (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) and the Hoogsteen edge of adenine in the FMN-RNA aptamer complex (Figure 6(b)) (Fan *et al*., 1996).

A novel feature of the FMN-RNA aptamer complex is the extensive stacking between the intercalated isoalloxazine chromophore of FMN and a G·(U·A) base triple platform at the binding site (Fan *et al*., 1996). It remains to be established whether this observation will turn out to be general in that base triples provide the extended surface area to maximally stack with large aromatic chromophores in RNA aptamer complexes.

Encapsulative recognition of amino acid ligands

The amino acids L-arginine and L-citrulline lack aromatic chromophores and hence their binding sites involve pockets where the non-polar methylene segments of the side-chains are positioned over platforms generated by alignments of purine bases (Figure 8). Molecular recognition involves intermolecular hydrogen bonds contributed by polar edges of multiple bases that surround the RNA binding pocket (Yang *et al*., 1996).

Encapsulative recognition of amino acids by nucleic acid folds has been observed recently for a DNA aptamer complex in our laboratory (Lin & Patel, 1996). In this case, L-argininamide is encapsulated within the DNA hairpin loop with the binding cavity generated by the tip of the loop folding back onto the stem in the complex. In addition, the amino acid is sandwiched between mismatch and regular base-pairs that form within the loop on complex formation, with binding specificity attributed to the formation of intermolecular hydrogen bonds between the guanidinium group on the L-argininamide and the Watson-Crick edge of a loop cytidine residue (Lin & Patel, 1996). It is noteworthy that salt bridges were not detected between the guanidinium group of the charged amino acid side-chain and the backbone phosphates in either the L-arginine-RNA aptamer (Yang *et al*., 1996) or the L-argininamide-DNA aptamer (Lin & Patel, 1996) complexes.

Bases that close the binding cavity play an important role in the encapsulation of ligands within nucleic acid folds. Thus, rings II and III of tobramycin are sandwiched between the floor of the RNA major groove and a looped out cytidine residue that acts as a flap over the bound antibiotic in the tobramycin-RNA aptamer complex (Figure 10(d)) (Jiang *et al*., 1997b). An adenine base forms a similar flap in the L-argininamide-DNA aptamer complex (Lin & Patel, 1996) and contributes to the encapsulation of the amino acid within the DNA fold. One face of a binding site can also be closed by an array of stacked bases as observed for the encompassing residues A13 to U16 that blocks cofactor entry from one face in the AMP-RNA aptamer complex (Jiang *et al*., 1996a; Dieckmann *et al*., 1996).

Discrimination between related amino acid ligands

RNA aptamers discriminate between closely related ligands through formation of defined donor-acceptor intermolecular hydrogen bonding alignments. This feature was best characterized by the arrangement of acceptor atoms on the RNA scaffold that target the donor $NH_2(\eta)$ protons of the L-arginine side-chain in the L-arginine-RNA aptamer complex (Figure 8(a)) which contrasts with the arrangement of donor atoms on the RNA scaffold that target the acceptor carbonyl (ε) oxygen on the L-citrulline side-chain in the L-citrulline-RNA aptamer complex (Figure 8(b)) (Yang *et al*., 1996).

Theophylline (1,3-dimethylxanthine) and caffeine (1,3,7-trimethylxanthine) are closely related xanthine analogs that differ by a single methyl group. An RNA aptamer has been identified that binds theophylline with 0.1 μM affinity and discriminates by a factor of 10^4 against caffeine (Jenison *et al*., 1994). Structural studies currently underway in the Pardi laboratory should provide a molecular explanation for the theophylline-binding

RNA aptamer's ability to discriminate between these two closely related xanthine analogs.

Stacking alignments within the core of an RNA fold

There is a preponderence of purine bases at the binding sites of cofactor-RNA aptamer complexes (Figures 1(a) and 5(b)) and these purines tend to participate in base stacking interactions when not involved in mismatch or triple formation. Indeed, such stacking interactions between purine planes as observed in the AMP-RNA aptamer complex (Figure 3(b)) (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) can constitute the core of folded RNA architectures. It is striking that three mutually orthogonal arrays of stacked bases/base-pairs converge into the core of the AMP-RNA aptamer complex (Figure 3(b)).

The stacked core can also be centered about the bound ligand as in the case of the intercalated isoalloxazine ring in the FMN-RNA aptamer complex (Fan *et al*., 1996). Here, the asymmetric internal bubble of the free RNA aptamer (Figure 5(b)) zippers up through mismatch and base triple formation to form a continuous stacked helix (Figure 6(a)) that links up with adjacent stem regions in the FMN-RNA aptamer complex.

Diversity of RNA architecture

The majority of nucleotides adopt C3′-*endo* sugar puckers and *anti* glycosidic bonds in the RNA aptamer complexes studied to date. However, there are exceptions with *syn* guanines observed in both the AMP-RNA aptamer (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) and L-arginine/L-citrulline-RNA aptamer (Yang *et al*., 1996) complexes. Generally, the *syn* guanines participate in G(*syn*)·G(*anti*) mismatch pair formation and extend existing helical stems in the RNA aptamers associated with complex formation. It is too early to predict whether examples will emerge of guanine, adenine and cytosine residues adopting *syn* alignments, and with what frequency, within folds of RNA aptamer complexes. The C2′-*endo* sugar pucker conformation has also been occassionally observed in RNA aptamer complexes at sites of abrupt chain reversal and where bases are looped out of the helix. Examples of C2′-*endo* sugar puckers can be found in the AMP-RNA aptamer (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) and the HIV-1 Rev peptide-RRE RNA (Battiste *et al*., 1996; Ye *et al*., 1996) complexes.

RNA differs from DNA in having a OH group at the the 2′ position of the sugar ring. Generally, sugar hydroxyls exchange rapidly with solvent water and, if observable, resonate at ≈6.8 ppm (Allain & Varani, 1995). The observation of a narrow slowly exchanging 2′-OH proton at 9.34 ppm in the AMP-RNA aptamer complex (Figure 1(c)) (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) and at 9.36 ppm in the tobramycin-RNA aptamer complex (Figure 9(d)) (Jiang *et al*., 1997b) were received with considerable surprise in our laboratory. These narrow downfield-shifted 2′-OH protons are buried from solvent, participate in hydrogen bond formation and are positioned in the planes of one or more purine rings in the complex.

Diversity of mismatch and triple pairing alignments

The RNA aptamer complexes utilize a diverse array of mismatches and triples to generate the RNA binding pockets in the complex. Thus, a range of G·G and G·A mismatches have been identified in the various RNA aptamer complexes whose structures have been solved to date. The G·G mismatch alignments include pairing of the Watson-Crick edges of *anti* guanines as observed in the FMN-RNA aptamer complex (Fan *et al*., 1996), pairing of the Watson-Crick and Hoogsteen edges of *anti* guanines in a reverse pair (Figure 3(c)) as observed in the AMP-RNA aptamer (Jiang *et al*., 1996a; Dieckmann *et al*., 1996), and the pairing of Watson-Crick and Hoogsteen edges of *anti* and *syn* guanines, respectively (Figure 3(c)) as observed in the AMP-RNA aptamer (Jiang *et al*., 1996a; Dieckmann *et al*., 1996) and L-arginine/L-citrulline-RNA aptamer (Yang *et al*., 1996) complexes. The corresponding G·A mismatch alignments include pairing of the Watson-Crick edges of *anti* guanines and adenines (Figure 6(d)) as observed in the FMN-RNA aptamer (Fan *et al*., 1996), L-arginine/L-citrulline-RNA aptamer (Yang *et al*., 1996) and HIV-1 Rev peptide-RRE RNA aptamer (Battiste *et al*., 1996; Ye *et al*., 1996) complexes and pairing of the minor groove edge of *anti* guanine and the major groove edge of *anti* adenine to form a sheared mispair (Figure 6(e)) as observed in the FMN-RNA aptamer complex (Fan *et al*., 1996). In addition, G·A mispair formation involving the Watson-Crick edge of *anti* AMP ligand and the minor groove edge of *anti* guanine (Figure 3(d)) are associated with specific recognition of the bound ligand in the AMP-RNA aptamer complex (Jiang *et al*., 1996a; Dieckmann *et al*., 1996).

The U·(A·U) base triple that involves pairing of the Watson-Crick edge of the uracil and the major groove edge of the Watson-Crick A·U base-pair has been observed in both the HIV-1 Rev-RRE RNA aptamer (Ye *et al*., 1996) and BIV Tat-TAR (Ye *et al*., 1995) complexes. The uracil base of the U·(A·U) base triple is positioned in the major groove along with the bound peptide in these peptide-RNA complexes. The triple-forming uracil's role is to buttress the helix, locally align the phosphodiester backbone and act as a scaffold for positioning amino acid side-chains of the peptide in its vicinity. The G·(U·A) triple serves a different role in the FMN-RNA aptamer complex (Fan *et al*., 1996). Here, the uracil in the G·(U·A) triple acts as a bridge to span the guanine and adenine residues on partner strands across the asymmetric internal

loop. Further, the G·(U·A) triple serves as a platform within the intercalative binding site resulting in extensive stacking between this triple and the bound isoalloxazine ring in the complex. Clearly, there are most likely a range of base triple alignments in folded RNA and novel base triples should emerge following solution of additional RNA aptamer complexes.

Groove dimensions and recognition

The major groove appears to be the target of both aminoglycoside antibiotics (Fourmy *et al.*, 1996; Jiang *et al.*, 1997b) and peptides (Battiste *et al.*, 1996; Ye *et al.*, 1995, 1996; Puglisi *et al.*, 1995) that bind RNA. The major groove is narrow in *A*-form RNA and needs to be widened in order to accomodate the bound ligand on complex formation. Helix interruptions such as mismatches and bulges have been shown to widen the major groove of RNA and provide accessibility of bound ligands to its very deep interior (Weeks & Crothers, 1993). Thus, it is not surprising that mismatches and looped out bases, together with base triples are a common feature of the major groove binding sites of RNA aptamer complexes. It is less clear whether the RNA minor groove will also accomodate bound ligands and additional structures of ligand-RNA complexes will need to be solved to address this issue.

Adaptive binding

The bulges and internal loops of free RNA aptamers generally adopt partially structured and often interconverting conformations in solution. This is reflected in the observation of broad exchangeable proton resonances centered about ≈10.7 ppm for guanine and uracil imino protons located at bulge and loop sites (Figures 1(b) and 5(c)). By contrast, the same segments adopt defined secondary and tertiary RNA folds on formation of high affinity and specificity ligand-RNA aptamer complexes. The bulge and loop imino protons are narrower at the complex level and are dispersed over a much wider spectral range that extends several ppm downfield of ≈10 ppm (Figures 1(c) and 5(d)). Such transitions associated with complex formation are accompanied by the formation of base-pair mismatches and triples that extend existing Watson-Crick stem segments. The adaptive transitions reflect the need to generate an RNA binding pocket on complex formation.

Adaptive binding is not necessarily restricted to the RNA aptamer alone in its complexes with high affinity and specificity ligands. Thus, both the peptide and the RNA undergo adaptive conformational changes on formation of the HIV-1 Rev peptide-RRE RNA complex in solution (Battiste *et al.*, 1996; Ye *et al.*, 1996). A related adaptive transition of both components has also been observed on formation of the BIV Tat peptide-TAR RNA complex in solution (Puglisi *et al.*, 1995; Ye *et al.*, 1995).

Folding motifs common to RNA aptamers and natural RNA sequences

Several of the RNA folds identified in ligand-RNA aptamer complexes exhibit structural features that are common to natural RNA folds reported in the literature. Thus, it is quite remarkable that the non-covalently bound AMP becomes part of a GNRA-like hairpin loop closed by a G·G mismatch pair in the solution structure of the AMP-RNA aptamer complex (Figure 3(a)) (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996). The backbone torsion angles for the G8-A9-A10 segment in the AMP-RNA aptamer complex (Jiang *et al.*, 1996a; Dieckmann *et al.*, 1996) are very similar to what has been published for GNRA hairpin loops both in solution (Heus & Pardi, 1991) and in the crystalline state (Pley *et al.*, 1994).

It is of interest that internal loops closed by a sheared G·A pair (Figure 6(e)) stacked adjacent to a reversed Hoogsteen U·A pair (Figure 6(c)) as observed in the solution structure of the FMN-RNA aptamer complex (Fan *et al.*, 1996) have been reported previously for NMR based solution structures of the internal loop E in eukaryotic 5 S RNA (Wimberly *et al.*, 1993) and the sarcin-ricin hairpin loop of 28 S ribosomal RNA (Szewczak *et al.*, 1993). In addition, it has been proposed that a bulged out guanine may form a base triple with the reversed-Hoogsteen U·A pair in the solution structure of the internal loop E in eukaryotic 5 S RNA (Wimberly *et al.*, 1993). Thus, the structural motif formed through pairing alignments within the internal loop closing (G10-U12-A13)·(G24-A25) segment in the FMN-RNA aptamer complex exhibits striking similarities to internal and hairpin loops evolved through evolutionary selection in natural RNA systems.

This point is reiterated in the remarkable similarity in the global fold of the hairpin loop in the tobramycin-RNA aptamer complex (Jiang *et al.*, 1997b) and a conserved hexanucleotide loop in ribosomal RNA (Huang *et al.*, 1996; Fountain *et al.*, 1996).

Future Prospects

The solution structures of ligand-RNA aptamer complex outlined above represent the first contributions of structural biology to a field with significant therapeutic potential. Ample opportunities exist for additional structure-function correlations in the RNA aptamer field and a few potential challenges are outlined below.

Binding affinity, specificity and *in vivo* stability of RNA aptamers

The second generation of RNA aptamers have been designed to contain base and sugar ring modifications that have the potential to increase binding affinities/specificites and *in vivo* stability. These modifications range from 1-pentenyl substi-

tution at the pyrimidine base 5 position (Latham *et al.*, 1994) to 2′-NH_2 (Green *et al.*, 1995) and 2′-F (Pagratis *et al.*, 1997) substitution of the sugar ring of RNA aptamers. The 2′-F substituted RNA aptamers are of special interest because of their unusual thermostabilities, unusually high binding affinities (in the pM range) and nuclease resistance.

An alternate route around the problem of weak affinities would be to cross-link the ligand to the RNA aptamer. This has indeed been accomplished in the case of HIV-1 Rev protein which has been covalently linked to its selected RNA aptamer target (Jensen *et al.*, 1995). There are significant advantages to the NMR based structural characterization of such covalent ligand-RNA aptamer complexes since such an approach circumverts issues related to the potential broadening of spectra due to unfavorable complex dissociation kinetics.

RNA aptamers targeted to transition state analogs

Solution structures of ligand-RNA aptamer complexes characterized to date are limited to ligands in their ground state. However, several examples have been reported of RNA aptamers that target transition state analogs with high affinity and specificity (Morris *et al.*, 1994; Prudent *et al.*, 1994; Conn *et al.*, 1996). The most recent example is of a 35-mer stem-loop RNA aptamer that catalyzes the insertion of Cu(II) into mesoporphyrin IX and serves as a transition state analog of the corresponding mesoporphyrin metalation with Fe(II) catalyzed by human ferrochelatase (Conn, *et al.*, 1996). Structural studies of RNA aptamers bound to transition state analogs could provide critical insights into the molecular basis of RNA catalysis.

Mirror image RNA aptamers

More recently, the issue of chiral specificity has been addressed by *in vitro* selection studies using RNA libraries containing either the natural D-nucleotides or their mirror image L-nucleotide counterparts (Klußmann *et al.*, 1996; Nolte *et al.*, 1996). Thus, a D-RNA aptamer was identified that targeted L-adenosine along with its L-RNA counterpart that bound D-adenosine. Reciprocal chiral specificity was evident in the observed $\approx 10^4$ ligand discrimination in this series. Comparative structural studies of such mirror image related ligand-RNA aptamer complexes are of pharmacological importance since the L-RNA aptamers displayed extraordinary stability in human serum under conditions where their natural D-RNA aptamer counterparts were rapidly degraded by nucleases.

Pseudoknot and G-quartet RNA aptamer folding motifs

The solution structures of ligand-RNA aptamer complexes reported to date are limited to RNA sequences whose secondary structures contain mismatches, bulges and asymmetric internal loops. However, ligand-RNA aptamer complexes have been reported where the RNA secondary structure has the potential to form either pseudoknots (Tuerk *et al.*, 1992; Lorsch & Szostak, 1994a; Binkley *et al.*, 1995; Ringquist *et al.*, 1995) or G-quartets (Lin *et al.*, 1994; Lauhon & Szostak, 1995). It is conceivable that the pseudoknot and G-quartet RNA folds in these aptamer complexes will be stabilized by the bound ligands and hence be amenable to detailed structural characterization. Such studies should provide an opportunity to expand on the existing knowledge base of the RNA structure of pseudoknots (Puglisi *et al.*, 1991) and G-quartets (Cheong & Moore, 1992) and, in addition, provide insights into the recognition of these higher order RNA architectures.

Structure determination of larger RNA aptamer complexes

In principle, both NMR spectroscopy and X-ray crystallography should be the methods of choice for structure determination of RNA aptamer complexes to high resolution. Surprisingly, NMR has provided all the structures of complexes that involve adaptive transitions in the RNA aptamers associated with complex formation. These structures have involved complexes with small ligands ranging from cofactors to antibiotics to peptides targeting RNA sequences up to 40-mers in length. There are uncertainties as to the molecular weight limits attainable by NMR for studies of RNA aptamer complexes. There has been one example of an NMR based structure of a protein-RNA complex reported to date (Allain *et al.*, 1996) and more examples are likely to emerge in the future. The ability to uniformly ^{13}C,^{15}N-label both the protein and/or RNA components should greatly help to overcome resonance assignment ambiguities in larger RNA aptamer complexes through simplification of experimental data sets based on spectral editing approaches of data dispersed in multiple dimensions. It is conceivable that RNA aptamers containing more than 50 residues will require selective ^{13}C,^{15}N-labeling of specific residues and/or segments to overcome issues of spectral overlap, as well as uniform ^{2}H labeling (60 to 80%) to narrow residual proton line widths in larger RNA aptamer complexes. Clearly, X-ray crystallography will also contribute to the structural characterization of higher molecular weight systems such as protein-RNA aptamer complexes.

Coordinates deposition

The deposited coordinates of the RNA aptamer complexes discussed in this review have the following accession numbers and can be retrieved from the Brookhaven protein data bank: AMP-RNA aptamer complex (Jiang *et al.*, 1996a), 1am0; AMP-RNA aptamer complex (Dieckmann *et al.*, 1996), 1raw; FMN-RNA aptamer complex

(Fan *et al*., 1996), 1fmn; L-arginine-RNA aptamer (Yang *et al*., 1996), 1kod; L-citrulline-RNA aptamer (Yang *et al*., 1996), 1koc; Tobramycin-RNA aptamer (Jiang *et al*., 1997b), 1tob; HIV-1 Rev peptide-RRE RNA aptamer (Ye *et al*., 1996), 1ull.

Acknowledgments

This research was funded by NIH grant GM54777 and start up funds from the Memorial Sloan-Kettering Cancer Center. I thank Radovan Fiala and David Live of our NMR facility and John Hubbard of our computational facility for their technical assistance. Wei-jun Xu provided biochemical technical assistance in our laboratory. The Roger Jones laboratory provided our group with RNA oligomers containing specifically ^{15}N-labeled guanines that were critical for deducing unambiguous assignments of guanines in the asymmetric internal loop in the AMP-RNA aptamer complex.

References

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

Allain, F. H. T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D. & Varani, G. (1996). Specificity of ribonucleoprotein interaction determined by RNA folding during complex formation. *Nature,* **380**, 646–650.

Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. & Williamson, J. R. (1992). Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucl. Acids Res.* **20**, 4515–4523.

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 recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science,* **273**, 1547–1551.

Binkley, J., Allen, P., Brown, D. M., Green, L., Tuerk, C. & Gold, L. (1995). RNA ligands to human nerve growth factor. *Nucl. Acids Res.* **23**, 3198–3205.

Burgstaller, A. T. & Famulok, M. (1994). Isolation of RNA aptamers for biological cofactors by *in vitro* selection. *Angew. Chem. Intl. Edn. Eng.* **33**, 1084–1087.

Burgstaller, P., Kochoyan, M. & Famulok, M. (1995). Structural probing and damage selection of citrulline- and arginine-specific RNA aptamers identify base positions required for binding. *Nucl. Acids Res.* **23**, 4769–4776.

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

Clouet-d'Orval, B., Stage, T. K. & Uhlenbeck, O. C. (1995). Neomycin inhibition of the hammerhead ribozyme involves ionic interactions. *Biochemistry,* **34**, 11186–11190.

Conn, M. M., Prudent, J. R. & Schultz, P. G. (1996). Porphyrin metalation catalyzed by a small RNA molecule. *J. Am. Chem. Soc.* **118**, 7012–7013.

Dieckmann, T. & Feigon, J. (1994). Heteronuclear techniques in NMR studies of RNA and DNA. *Curr. Opin. Struct. Biol.* **4**, 745–749.

Dieckmann, T., Suzuki, E., Nakamura, G. K. & Feigon, J. (1996). Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA,* **2**, 628–640.

Ellington, A. & Szostak, J. W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature,* **346**, 812–822.

Famulok, M. (1994). Molecular recognition of amino acids by RNA aptamers: an L-citrulline binding RNA motif and its evolution into an L-arginine binder. *J. Am. Chem. Soc.* **116**, 1698–1706.

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.

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

Fiala, R., Jiang, F. & Patel, D. J. (1996). Direct correlation of exchangeable and nonexchangeable protons in purine bases in ^{13}C,^{15}N-labeled RNA using HCCNH-TOCSY experiment. *J. Am. Chem. Soc.* **118**, 689–690.

Fountain, M. A., Serra, M. J., Krugh, T. R. & Turner, D. (1996). Structural features of a six-nucleotide RNA hairpin loop found in ribosomal RNA. *Biochemistry,* **35**, 6539–6548.

Fourmy, D., Recht, M. I., Blanchard, S. C. & Puglisi, J. D. (1996). Structure of the A site of *Escherichia coli* 16 S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science,* **274**, 1367–1371.

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

Gold, L., Polisky, B., Uhlenbeck, O. C. & Yarus, M. (1995). Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64**, 763–797.

Green, L. S., Jellinek, D., Bell, C., Beeble, L. A., Feistner, B. D., Gill, S. C., Jucker, F. M. & Janjic, N. (1995). Nuclease-resistant nucleic acid ligands to vascular permeability factor/vascular endothelial growth factor. *Chem. Biol.* **2**, 683–695.

Gueron, M. & Leroy, J. L. (1995). Studies of base-pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.* **261**, 383–413.

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

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.

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

Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. (1994). High resolution molecular discrimination by RNA. *Science,* **263**, 1425–1429.

Jensen, K. B., Atkinson, B. L., Willis, M. C., Koch, T. H. & Gold, L. (1995). Using *in vitro* selection to direct the covalent attachment of human immunodeficiency virus type 1 Rev protein to high-affinity RNA ligands. *Proc. Natl Acad. Sci. USA,* **92**, 12220–12224.

Jiang, F., Kumar, R. A., Jones, R. A. & Patel, D. J. (1996a). Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature,* **382**, 183–186.

Jiang, F., Fiala, R., Live, D., Kumar, R. A. & Patel, D. J. (1996b). RNA folding topology and intermolecular

contacts in the AMP-RNA aptamer complex. *Biochemistry,* **35**, 13250–13256.
Jiang, F., Patel, D. J., Zhang, X., Zhao, H. & Jones, R. A. (1997a). Specific labeling approaches to guanine and adenine imino and amino proton assignments in the AMP-RNA aptamer complex. *J. Biomol. NMR.* **9**, 55–62.
Jiang, L., Suri, A. K., Fiala, R. & Patel, D. J. (1997b). Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex. *Chem. Biol.* **4**, 35–50.
Joyce, G. F. (1994). *In vitro* evolution of nucleic acids. *Curr. Opin. Struct. Biol.* **4**, 331–336.
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 phenyalanine transfer RNA. *Science,* **185**, 435–439.
Klußmann, S., Nolte, A., Bald, R., Erdmann, V. A. & Furste, J. P. (1996). Mirror-image RNA that binds D-adenosine. *Nature Biotechnol.* **14**, 1112–1115.
Latham, J. A., Johnson, R. & Toole, J. J. (1994). The application of a modified nucleotide in aptamer selection: novel thrombin aptamers containing 5-(1-pentenyl)-2′-deoxyuridine. *Nucl. Acids Res.* **22**, 2817–2822.
Lato, S. M., Boles, A. R. & Ellington, A. D. (1995). *In vitro* selection of RNA lechtins: using combinatorial chemistry to interpret ribozyme evolution. *Chem. Biol.* **2**, 291–303.
Lauhon, C. T. & Szostak, J. W. (1995). RNA aptamers that bind flavin and nicotinamide factors. *J. Am. Chem. Soc.* **117**, 1246–1257.
Li, Y., Zon, G. & Wilson, W. D. (1991). NMR and molecular evidence for a G·A mismatch base-pair in a purine-rich DNA duplex. *Proc. Natl Acad. Sci. USA,* **88**, 26–30.
Lin, C. H. & Patel, D. J. (1996). Encapsulating an amino acid in a DNA fold. *Nature Struct. Biol.* **3**, 1046–1050.
Lin, Y., Qiu, Q., Gill, S. C. & Jayasena, S. D. (1994). Modified RNA sequence pools for *in vitro* selection. *Nucl. Acids Res.* **22**, 5229–5234.
Lorsch, J. R. & Szostak, J. W. (1994a). *In vitro* selection of RNA aptamers specific for cyanocobalamin. *Biochemistry,* **33**, 973–982.
Lorsch, J. R. & Szostak, J. W. (1994b). *In vitro* evolution of new ribozymes with polynucleotide kinase activity. *Nature,* **371**, 31–36.
Lorsch, J. R. & Szostak, J. W. (1996). Chance and necessity in the selection of nucleic acid catalysts. *Acc. Chem. Res.* **29**, 103–110.
Moazed, D. & Noller, H. F. (1987). Interaction of antibiotics with functional sites in 16 S ribosomal RNA. *Nature,* **327**, 389–394.
Moore, P. B. (1995). Determination of RNA conformation by nuclear magnetic resonance. *Acc. Chem. Res.* **28**, 251–256.
Morris, K. N., Tarasow, T. M., Julin, C. M., Simons, S. L., Hilvert, D. & Gold, L. (1994). Enrichment for RNA molecules that bind a Diels-Alder transition state analog. *Proc. Natl Acad. Sci. USA,* **91**, 13028–13032.
Nagai, K. & Mattaj, I. W. (eds) (1994). *RNA-Protein Interactions,* IRL Press, New York.
Nicholls, A., Sharp, K. A. & Honig, B. H. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
Nikonowicz, E. P., Sirr, A., Legault, P., Jucker, F. M., Baer, L. M. & Pardi, A. (1992). Preparation of ^{13}C and ^{15}N labeled RNAs for heteronuclear multidimensional NMR studies. *Nucl. Acids Res.* **20**, 4508–4513.
Nolte, A., Klußmann, S., Bald, R., Erdmann, V. A. & Furste, J. P. (1996). Mirror design of L-oligonucleotide ligands binding to L-arginine. *Nature Biotechnol.* **14**, 1116–1119.
Nonin, S., Jiang, F. & Patel, D. J. (1997). Imino proton exchange and base-pair kinetics in the AMP-RNA aptamer complex. *J. Mol. Biol.* **268**, 359–379.
Pagratis, N. C., Bell, C., Chang, Y-F., Jennings, S., Fitzwater, T., Jellinek, D. & Dang, C. (1997). Potent 2′-amino- and 2′-fluoro-2′-deoxyribonucleotide RNA inhibitors of keratinocyte growth factor. *Nature Biotechnol.* **15**, 68–73.
Pardi, A. (1995). Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. *Methods Enzymol.* **261**, 350–380.
Patel, D. J., Ye, X., Kumar, R. A. & Gorin, A. (1998). Structure, recognition and adaptive binding in regulatory immunodeficiency viral peptide-RNA complexes. *J. Mol. Biol..*
Pley, H. W., Flahorty, K. M. & McKay, D. B. (1994). Three dimensional structure of a hammerhead ribozyme. *Nature,* **372**, 111–113.
Prudent, J. R., Uno, T. & Schultz, P. G. (1994). Expanding the scope of RNA catalysis. *Science,* **264**, 1924–1927.
Puglisi, J. D., Wyatt, J. R. & Tinoco, I., Jr (1991). RNA pseudoknots. *Acc. Chem. Res.* **24**, 152–158.
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.
Recht, M. I., Fourmy, D., Blanchard, S. C., Dahlquist, K. D. & Puglisi, J. D. (1996). RNA sequence determinants for aminoglycoside binding to an A-site rRNA model oligonucleotide. *J. Mol. Biol.* **262**, 421–436.
Ringquist, S., Jones, T., Snyder, E. E., Gibson, T., Boni, I. & Gold, L. (1995). High affinity RNA ligands to *E. coli* ribosomes and ribosomal protein S1: comparision of natural and unnatural binding sites. *Biochemistry,* **34**, 3640–3648.
Robertson, D. L. & Joyce, G. F. (1990). Selection *in vitro* of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature,* **344**, 467–468.
Robertus, J. D., Ladner, J. E., Finch, J. R., Rhodes, D., Brown, R. S., Clark, B. F. & Klug, A. (1974). Structure of yeast phenyalanine tRNA at 3 Å resolution. *Nature,* **250**, 546–551.
Sassanfar, M. & Szostak, J. W. (1993). An RNA motif that binds ATP. *Nature,* **364**, 550–553.
Simorre, J. P., Zimmermann, G. R., Mueller, L. & Pardi, A. (1996). Correlation of the guanosine exchangeable and nonexchangeable base protons in $^{13}C,^{15}N$-labeled RNA with an HNC-TOCSY-CH experiment. *J. Biomol. NMR,* **7**, 153–156.
Sklenar, V., Dieckmann, T., Butcher, S. E. & Feigon, J. (1996). Through bond correlation of imino and aromatic resonances in $^{13}C,^{15}N$-labeled RNA via heteronuclear TOCSY. *J. Biomol. NMR,* **7**, 83–87.
Szewczak, A. A., Moore, P. B., Chan, 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. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to

bacteriophage T4 DNA polymerase. *Science,* **249**, 505–510.
Tuerk, C., MacDougal, S. & Gold, L. (1992). RNA pseudoknots that inhibit human immunodeficiency virus type I reverse transcriptase. *Proc. Natl Acad. Sci. USA,* **89**, 6988–6992.
Varani, G. & Tinoco, I., Jr (1991). RNA structure and NMR spectroscopy. *Quart. Rev. Biophys.* **24**, 479–532.
von Ahsen, U., Davies, J. & Schroeder, R. (1991). Antibiotic inhibition of group I ribozyme function. *Nature,* **353**, 368–370.
Wallis, M. G., von Ahsen, U., Schroeder, R. & Famulok, M. (1995). A novel RNA structure for neomycin recognition. *Chem. Biol.* **2**, 543–552.
Wang, Y. & Rando, R. R. (1995). Specific binding of aminoglycoside antibiotics to RNA. *Chem. Biol.* **2**, 281–290.
Weeks, K. M. & Crothers, D. M. (1993). Major groove accessibility of RNA. *Science,* **261**, 1574–1577.
Westhof, E., Romby, P., Romaniuk, P. J., Ebel, J. P., Ehresmann, C. & Ehresmann, B. (1989). Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5 S RNAs. *J. Mol. Biol.* **207**, 417–431.
Wimberly, B., Varani, G. & Tinoco, I., Jr (1993). The conformation of loop E of eukaryotic 5 S ribosomal RNA. *Biochemistry,* **32**, 1078–1087.
Woese, C. R., Winker, S. & Guttell, R. R. (1990). Architecture of ribosomal RNA: constraints on the sequences of ''tetra-loops''. *Proc. Natl Acad. Sci. USA,* **87**, 8467–8471.
Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. & Famulok, M. (1996). Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science,* **272**, 1343–1347.
Ye, X., Kumar, R. A. & Patel, D. J. (1995). Molecular recognition of 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.
Zapp, M. L., Stern, S. & Green, M. R. (1993). Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell,* **74**, 969–978.

Edited by P. E. Wright

(Received 26 March, 1997; received in revised form 14 July, 1997; accepted 14 July, 1997)