RNA (1996), 2:628–640. Cambridge University Press. Printed in the USA.

Solution structure of an ATP-binding RNA aptamer reveals a novel fold

THORSTEN DIECKMANN, ELLEN SUZUKI, GERALD K. NAKAMURA, and JULI FEIGON

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90095, USA

ABSTRACT

In vitro selection has been used to isolate several RNA aptamers that bind specifically to biological cofactors. A well-characterized example is the ATP-binding RNA aptamer family, which contains a conserved 11-base loop opposite a bulged G and flanked by regions of double-stranded RNA. The nucleotides in the consensus sequence provide a binding pocket for ATP (or AMP), which binds with a K_d in the micromolar range. Here we present the three-dimensional solution structure of a 36-nucleotide ATP-binding RNA aptamer complexed with AMP, determined from NMR-derived distance and dihedral angle restraints. The conserved loop and bulged G form a novel compact, folded structure around the AMP. The backbone tracing of the loop nucleotides can be described by a Greek zeta (ζ). Consecutive loop nucleotides G, A, A form a U-turn at the bottom of the ζ, and interact with the AMP to form a structure similar to a GNRA tetraloop, with AMP standing in for the final A. Two asymmetric G·G base pairs close the stems flanking the internal loop. Mutated aptamers support the existence of the tertiary interactions within the consensus nucleotides and with the AMP found in the calculated structures.

Keywords: AMP; ATP; G·G base pair; GNRA tetraloop; in vitro selection; NMR; SELEX; 2D NMR; U-turn

INTRODUCTION

RNA aptamers (Ellington & Szostak, 1990) selected for binding to biological cofactors provide a starting point for exploring structural motifs that nucleic acids might form in their potential RNA-world role as catalysts of chemical reactions. One example is the ATP-binding RNA aptamer family (Sassanfar & Szostak, 1993), which contains a conserved 11-base loop opposite a bulged G and flanked by regions of double-stranded RNA (Fig. 1). This consensus sequence has been found in two separate in vitro selection (Ellington & Szostak, 1990; Tuerk & Gold, 1990; Szostak, 1992) experiments: initially from an RNA pool selected for binding to ATP that had been covalently linked to a agarose column via the C8 position (Sassanfar & Szostak, 1993), and subsequently in a selection for aptamers that bind to NAD^+ (Burgstaller & Famulok, 1994). The ATP-binding aptamers bind specifically and selectively to ATP, AMP, adenosine, and NAD^+ with K_d's in the micromolar range; there is no detectable binding to other nucleotides (Sassanfar & Szostak, 1993). Here we present the three-dimensional solution structure of a 36-nucleotide ATP-binding RNA aptamer complexed with AMP. The structure was determined from NMR-derived distance and dihedral angle restraints. The conserved loop and bulged G form a novel compact, folded structure around the AMP, which we describe by a Greek zeta (ζ). Consecutive loop nucleotides G8, A9, and A10 and the AMP form a structure similar to that of a GNRA tetraloop, with the AMP standing in for the final A. Two asymmetric G·G base pairs interact with the ligand by stacking, one with the adenine part of the AMP and one with the ribose. Three nonconserved bases in the aptamer form a linker between conserved residues that form the binding pocket. Aptamers with single and two-base substitutions support the existence of the tertiary interactions within the consensus nucleotides and with the AMP.

Reprint requests to: Juli Feigon, Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90095, USA; e-mail: feigon@ewald.mbi.ucla.edu.

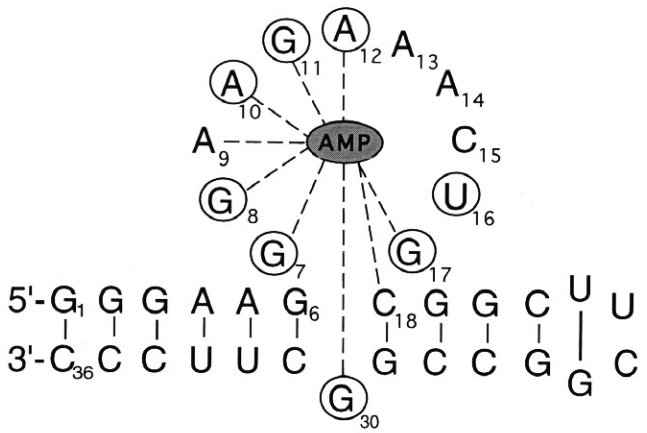

FIGURE 1. Sequence and proposed secondary structure of the ATP-binding aptamer used in this study. The sequence is based on the consensus reported by Sassanfar and Szostak (1993). Nucleotides that were conserved in all but one of the clones are circled. Dashed lines connecting the AMP to bases in the internal loop indicate nucleotides for which AMP–RNA NOEs were observed.

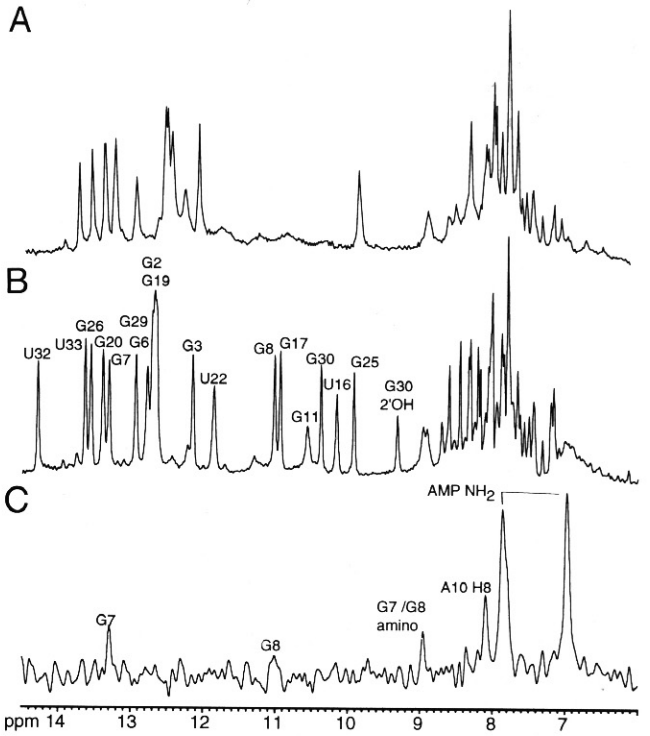

FIGURE 2. One-dimensional exchangeable and aromatic proton spectrum at 500 MHz of the ATP-binding aptamer in 100 mM NaCl, pH 6.0, 1 °C, 1.2 mM RNA (**A**) before and (**B**) after addition of a stoichiometric amount of AMP. Assignments of the imino resonances are indicated on the spectrum. The spectra were acquired at 5 °C using 11 spin-echo water suppression. For each spectrum, 128 scans of 4,096 complex points were acquired. The FIDs were zero-filled to 16K and transformed after multiplication with a 60° shifted squared sinebell. **C:** ^{15}N-filtered 1D NOE spectrum of ^{13}C,^{15}N-labeled AMP complexed to unlabeled RNA at 1 °C. The AMP amino resonances and NOEs are labeled. Suppression of the solvent signal was achieved using a waterflip back WATERGATE sequence. The spectrum was acquired using 28,000 scans of 4,096 complex points. The FID was zero-filled to 16K, line-broadened by 20 Hz, and transformed. The resulting spectrum shows two signals for the AMP amino group (identified from 2D ^{1}H-^{15}N correlation) and several signals originating from NOE connectivities.

RESULTS

Binding of AMP induces a structural change in the internal loop

The ATP-binding aptamer in the absence of AMP folds to form the expected double-helical stem base pairs and UUCG tetraloop (Allain & Varani, 1995) secondary structure (Fig. 1), but both the imino proton NMR spectra that provide evidence for base pairing (Fig. 2A) and the NMR spectra of the nonexchangeable protons (not shown) indicate that the internal loop is largely unstructured. For the free ATP-binding aptamer, only iminos from the stems and UUCG tetraloop are observed. Upon addition of AMP, additional imino resonances from all of the iminos in the internal loop appear; most are in the chemical shift range normally observed for hydrogen-bonded iminos (Fig. 2B). With the exceptions of G30 and G25, the corresponding ^{15}N chemical shifts are also in the range normally observed for nitrogens attached to a hydrogen-bonded imino proton. Assignments of the imino proton resonances of the aptamer were obtained from through bond correlations to the assigned aromatic H8,H6 resonances observed in HCCNH experiments, as previously reported (Sklenář et al., 1996).

The appearance of additional resonances from all of the imino protons in the internal loop upon addition of AMP in stoichiometric amounts provides initial evidence for formation of a highly structured complex. This is confirmed by analysis of the two- and three-dimensional homo- and heteronuclear NMR spectra used to determine the three-dimensional structure of the ATP-binding aptamer/AMP complex. Results identical to those described above were obtained with samples prepared using ATP rather than AMP. Although Mg^{2+} was used in the selection buffer (Sassanfar & Szostak, 1993), it is not necessary for binding under the NMR conditions. In order to avoid hydrolysis, we therefore did not add Mg^{2+} to the samples.

Identification of a slowly exchanging 2′OH

In addition to the appearance of imino resonances from the internal loop, a surprisingly low-field-shifted, slowly exchanging 2′OH is observed at ~9.3 ppm in spectra in H_2O (Fig. 2B). This exchangeable resonance does not appear in any of the ^{1}H-^{15}N correlation spectra recorded on samples containing ^{15}N-labeled RNA or ^{15}N-labeled AMP, and therefore it can only arise from a 2′OH. The resonance was assigned to G30

based on the characteristic NOE pattern exhibited in NOESY spectra recorded from the sample in H_2O (Allain & Varani, 1995). Although hydrogen bonding of 2′OH is generally believed to be important in a variety of RNA structures, 2′ hydroxyl protons have rarely been observed by NMR due primarily to rapid exchange with water. The best-studied example, the 2′OH from a UUCG tetraloop (which cannot be unambiguously identified in this larger aptamer RNA due to spectral overlap), appears in the expected lower-field region at ~6.8 ppm (Allain & Varani, 1995). The linewidth of the G30 2′OH is 25 Hz, which is comparable to the linewidths of the imino resonances (~20–40 Hz). Hydrogen bonding of G302′OH is consistent with the fact that a single deoxyribose substitution at position 30 abolishes binding of ATP to the aptamer (M. Sassanfar & J.W. Szostak, pers. comm.).

Assignment of the RNA/AMP complex requires a variety of $^{13}C,^{15}N$-labeled RNA and AMP samples

The ATP-binding aptamer/AMP complex gives reasonably well-resolved proton NOESY spectra in the base-H1′ region used for standard sequential assignments of nucleic acid duplexes (Fig. 3). However, unambiguous assignment of the RNA required use of $^{13}C,^{15}N$-labeled samples as well as RNAs containing base substitutions, because (1) the RNA was of relatively large size and (2) more importantly, the unknown structure of the internal loop prevented sequential NOEs from being distinguished a priori from long-range NOEs. The mutated RNAs studied are given in Table 1. Assignments for the complex were determined from NMR spectra obtained on unlabeled, uniformly $^{13}C,^{15}N$-labeled, and uniformly $^{13}C,^{15}N$-A-, G-, or U/C-only la-

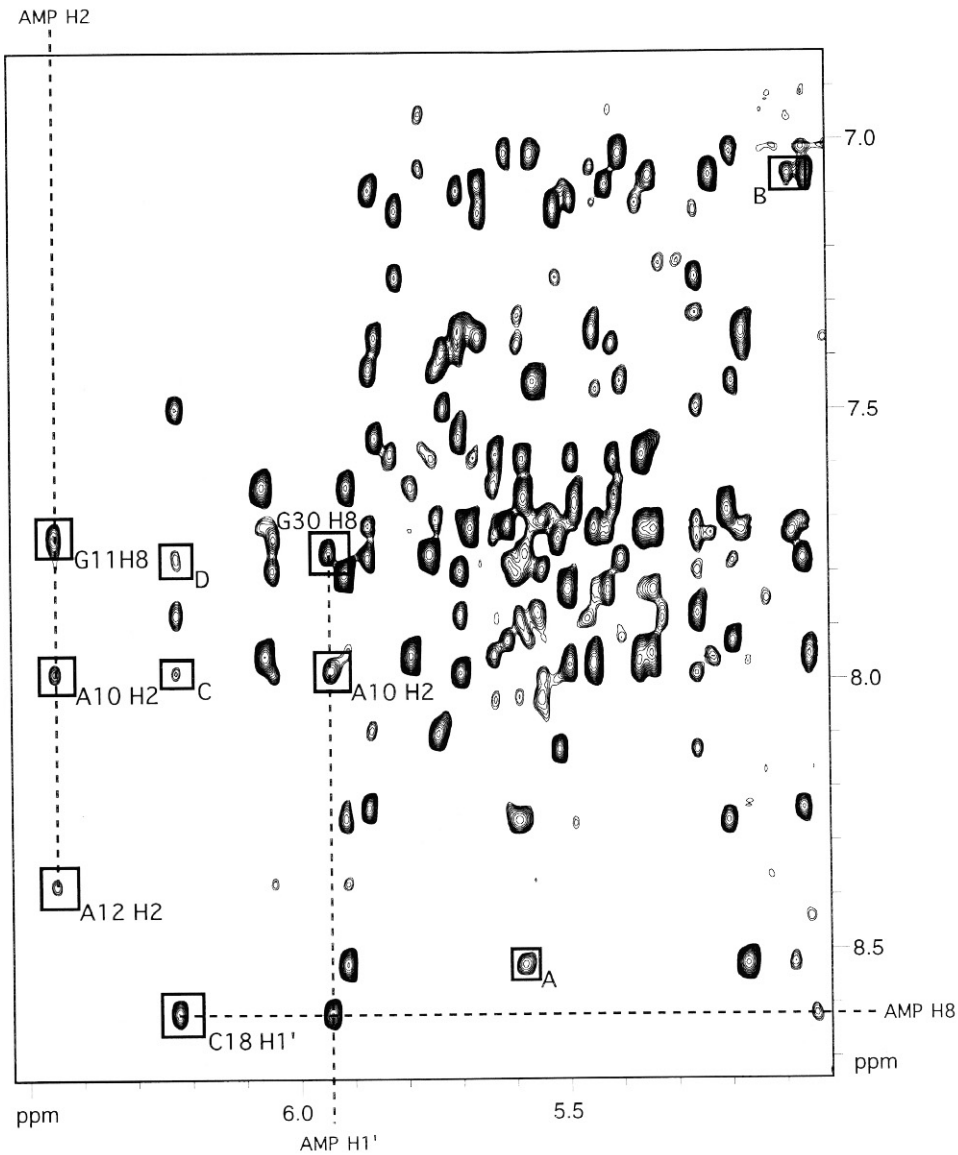

FIGURE 3. Portion of a 750-MHz NOESY spectrum of a 1:1 AMP:ATP-binding RNA aptamer complex at 10 °C showing the base-H1′ crosspeak region. NOEs between the AMP and the nucleotides in the RNA are boxed and labeled. Some nonsequential NOEs that arise from important tertiary interactions are: (**A**) G30H2′–A12H8, (**B**) G11H1′–C31H6, (**C**) C18H1′–A10H2, and (**D**) C18H1′–G30H8. In t_2 and t_1, 1,024 and 300 complex points were acquired, respectively, with 64 scans per t_1 increment and States-TPPI phase cycling. The final data matrix was 2,048 × 2,048 points and was processed with a Gaussian filter function (LB −18, GB 0.08 in f_2 and 0.14 in f_1).

TABLE 1. AMP-binding results for mutants of ATP-binding aptamer.[a]

Mutation	Binding	Conserved base
G7A	+	Yes
G8A	+	Yes
A9U	++	No
G11A	−	Yes
G11C	−	Yes
A12G	+/−	Yes
C15U	++	No
G17A	+	Yes
C18U, G29A	++	No
G30A	+/−	Yes

[a] Binding: −, no bound form detectable; +/−, weak binding, mostly unbound RNA; +, slightly weaker than wild-type binding; ++, comparable to wild-type binding, as assayed by NMR. [RNA] = 0.4 mM in 100 mM NaCl, pH 6.0. AMP was titrated in from equimolar (0.4 mM) to 2-fold excess (1.2 mM), and binding was assessed by NMR spectroscopy at T (°C) = 1, 5, 10, 20 as described in the text.

beled RNA complexed to unlabeled or $^{13}C,^{15}N$-labeled AMP. The use of labeled RNA and unlabeled AMP and vice versa allowed selective observation of intra- and intermolecular RNA and AMP NOEs (Otting & Wüthrich, 1990). Assignments of the nonexchangeable resonances of the RNA were obtained from a large number of homonuclear and heteronuclear two- and three-dimensional NMR experiments (reviewed in Dieckmann & Feigon, 1994; Pardi, 1995). Ambiguities in assignment due to overlap and/or nonsequential NOEs were largely resolved using spectra obtained on the selective base-labeled RNAs and the mutant RNAs. With the exception of many of the H5′,H5″ and aminos, nearly complete ^{1}H, ^{13}C, and ^{15}N assignments of the RNA nucleotides in the loop were obtained. The assignments of the RNA provided the greatest challenge of the structure determination, because numerous nonsequential NOEs are observed throughout the internal loop. These assignments will be discussed in detail elsewhere (ms. in prep.).

AMP has NOE contacts with many of the bases in the internal loop

The AMP resonances for bound and unbound AMP were readily identified from crosspeaks observed in ^{1}H-^{13}C HMQC (Fesik & Zuiderweg, 1988) and ^{1}H-^{15}N long-range HSQC (Sklenář et al., 1994) recorded on a sample containing unlabeled RNA and a slight excess of $^{13}C,^{15}N$-labeled AMP (Fig. 4). Binding of AMP results in a large upfield shift of the AMP H2 resonances, a downfield shift of the AMP H8 resonances, and chemical shifts of N1 and N3 similar to those seen for A in Watson–Crick base pairs. NOE crosspeaks between the AMP and the RNA, as well as intra-AMP crosspeaks, are selectively observed in 2D HMQC-NOESY spectra of 1:1 $^{13}C,^{15}N$-labeled AMP:RNA aptamer sample in D_2O and H_2O. Additional information on the interactions between the AMP amino protons and the RNA was obtained from a ^{15}N-filtered one-dimensional NOE spectrum (Fig. 2C). The AMP amino resonances appear as a resolved pair separated by about 1 ppm, indicating that one amino proton is hydrogen bonded. NOEs to G7 and G8 iminos and A10H8 orient the AMP with respect to these bases. A total of 40 NOEs be-

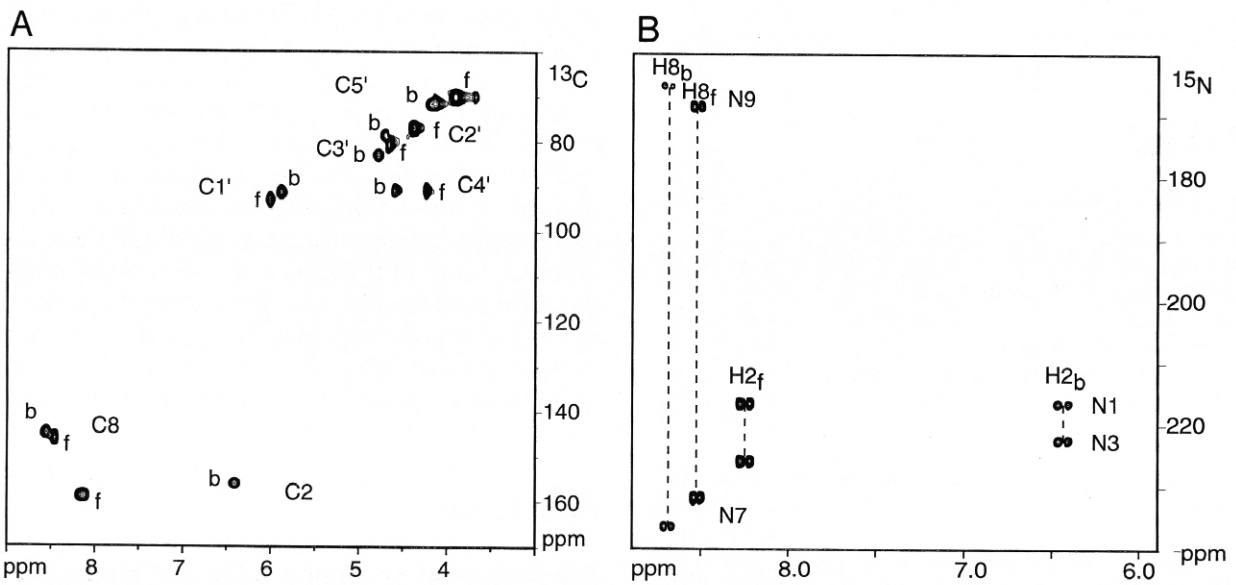

FIGURE 4. (**A**) ^{1}H-^{13}C HMQC and (**B**) ^{1}H-^{15}N long-range HSQC spectra of a 1.5:1 AMP/ATP-binding RNA aptamer complex. Signals originating from bound (b) and free (f) AMP are labeled. Both spectra were acquired at 20 °C with 1,024 and 64 complex points in t_2 and t_1, respectively, using 64 scans per t_1 increment and States–TPPI phase cycling. The final data matrix was 2,048 × 1,024 points and was processed with a 30° and 90° shifted squared sine bell filter function in f_2 and f_1, respectively.

tween AMP and the RNA (Table 2) were observed to seven of the eight conserved bases in the internal loop as well as to A9 and C18 (Fig. 1).

NMR spectra reveal that some conserved bases in the aptamer consensus can be mutated and still retain binding to AMP

Aptamers containing single (or double) base mutations (Table 1) were synthesized and studied by NMR not only to help in obtaining unambiguous sequential assignments for the wild-type aptamer complex but also to see if they provided supporting evidence for base pair interactions prior to completing the aptamer structure calculations. For example, we originally made the G8A and G7A mutants to identify which base interacted directly with the AMP. Binding of AMP to mutated RNAs was assayed first by examination of the changes in the imino proton spectra of the aptamers upon addition of AMP. For those mutants for which evidence for some binding was found, 2D NOESY spectra were used to assay the binding of AMP to the mutant relative to the wild type, based on (1) the appearance of exchange crosspeaks between bound and unbound forms of the AMP and (2) the amount of unbound RNA in the presence of stoichiometric amounts of AMP and RNA. Based on the NMR spectra, all of the aptamers containing base substitutions in nonconserved bases showed wild-type binding to AMP. As expected, several of the aptamers with substitutions in the conserved guanines showed greatly decreased or no detectable binding to AMP. However, surprisingly, three of the five conserved guanines, including G8, which pairs with AMP, can be individually replaced with adenine and still form a fully bound complex with AMP in the presence of <2-fold excess of AMP (Table 1). Analysis of the 2D NOESY spectra clearly shows that in all three of these cases there is no significant change in the tertiary structure or interactions in the complex.

TABLE 2. NMR and refinement statistics for the ATP-binding RNA aptamer.

NMR-derived distance and dihedral angle restraints, internal loop and flanking base pairs and AMP only (residues 6–18, 29–31, and AMP)	
Intranucleotide	64
Sequential	74
Medium-long-range	111
RNA–AMP	40
Dihedral angle	34
Total	263
Overall total for all nucleotides	425
Structure statistics	
NOE violations	None >0.5 Å
Average residual NOE violation	0.09 Å
Angle violations	None >5°
Average residual angle violation	0.11°
Mean deviation from ideal covalent geometry	
Bond lengths	0.012 Å
Bond angles	2.57°
Impropers	1.49°
Pairwise RMSD (Å) for all heavy atoms of the 10 lowest-energy structures	
All residues	2.32 ± 0.42
Internal loop and AMP (residues 6–18, 29–31, AMP)	1.52 ± 0.31
U-turn (residues 6–10)	1.47 ± 0.51
Linker (residues 13–17)	1.29 ± 0.33
Binding pocket and AMP (residues 7, 8, 10–12, 30, AMP)	1.24 ± 0.26

Three-dimensional structure of the ATP-binding RNA aptamer/AMP complex is constrained by a large number of long-range interactions

Starting structures of the ATP-binding RNA aptamer/AMP complex were obtained by distance geometry substructure embedding. Of 100 starting structures, the 22 lowest-energy structures were then refined by molecular dynamics. The large number of NOE interactions identified between the AMP and the RNA and long-range NOEs within the RNA internal loop provide an unusual number of tertiary constraints for a nucleic acid structure (Table 2). The AMP has a total of 45 restraints, and the average number of restraints for the internal loop nucleotides is 15, the majority of which are internucleotide. Although a few intra-ribose NOE distances obtained from well-resolved crosspeaks (which in general occurred for sugar protons with unusual chemical shifts) were included in the structure calculations, we did not include any data from the majority of severely overlapped sugar resonance crosspeaks, since within experimental error these would not affect the final structure. Within the ribose, only the H1′–H4′ distance varies by more than 0.4 Å over the entire range of sugar puckers (Wüthrich, 1986). Because the initial structures clearly indicated that the stems and UUCG tetraloop do not form part of the binding domain, and the NMR data are consistent with standard A-form duplex structure, we did not attempt to optimize them in the structure calculations. Instead, the stem nucleotides, with the exception of the base pairs flanking the internal loop, were restrained to standard A-form geometry, although NOE and dihedral angle restraints for them were also included in the structure calculations.

DISCUSSION

The consensus sequence of the ATP-binding RNA aptamer forms a highly structured binding pocket for AMP

The calculated three-dimensional structure of the ATP-binding RNA aptamer/AMP complex reveals that the

conserved internal loop forms an independent structural domain that is anchored by the base pairs on the two stems (Fig. 5A). The backbone tracing of the 11-base loop looks like a Greek zeta (ζ), with G7 through G11 forming the bottom right-handed turn of the ζ, A12 through C15 forming the large central left-handed turn, and U16 and G17 making the little turn at the top (Fig. 6). G8–A10 form a loop that is closed by the G7·G11 base pair, which is stacked over the last base pair (G6·C31) in the first stem. A13–C15 form a continuous stack, and G17 reaches across to base pair with G30 at the top of the ζ. A12 is in the interior of the binding pocket and interacts directly with the AMP and likely with the 2′OH of G30. The AMP base is stacked on top of the G7·G11 base pair and is hydrogen bonded to G8 and A12, while the AMP ribose is stacked on the G30·G17 base pair, which lies almost perpendicular to the G7·G11 base pair.

A U-turn forms part of the binding pocket

The G8, A9, A10, and AMP form a structure that is similar in many ways to a GNRA tetraloop (Woese et al., 1990; Heus & Pardi, 1991), with AMP standing in for the A. Nucleotides G8–A10 are positioned almost identically to those found in GNRA tetraloops, such as the GUAA tetraloop in the crystal structure of the hammerhead ribozyme (Pley et al., 1994; Scott et al., 1995) (Fig. 7). This three-base U-turn (Quigley & Rich, 1976) motif (G/UNR) has been characterized by a turn in the backbone between the G/U and N, stacking of the R nucleotide beneath N, and two hydrogen bonds, one from the 2′OH of the G/U to the N7 of R and the other from the amino/imino proton of the G/U to the 3′ phosphate of the R (Jucker & Pardi, 1995). In the GNRA loop, the G is also base paired to the A. In the ATP-binding aptamer structure, the 2′OH of G8 is positioned so that a hydrogen bond to the A10N7 appears likely. Further support for this hydrogen bond comes from ribose substitution experiments. Replacement of the G8 ribose with deoxyribose or 2′O methyl greatly decreases or abolishes binding to ATP, respectively; this is the only other nucleotide beside G30 in which single deoxyribose substitutions strongly affect binding (M. Sassanfar & J.W. Szostak, unpubl. data). A 2′O methyl group would clearly be sterically blocked in the structure. Analysis of the structures indicates that there is also a hydrogen bond between G8N1 and the A10 3′ phosphate, as observed in the U-turn motif, although the stacking of the AMP between A10 and G11 might have been expected to move the position of this phosphate relative to the standard U-turn. The AMP is positioned next to G8, but not in the same orientation as the base pair that is found in GNRA tetraloops. This is not surprising, because the AMP is not constrained by a phosphodiester linkage to the rest of the "tetraloop."

The nonconserved bases in the center of the ζ form a linker between bases in the binding pocket

Nucleotides A12–C15 comprise the central part of the overall ζ-shaped fold of the phosphodiester backbone of the aptamer. There is a left-handed turn between A12 and A13, and then A13 through C15 form a continuous single-strand stack. These stacked bases essentially constitute a linker (Fig. 5D) that connects the conserved bases G30, G17, A12, and U16 that form the back side of the AMP-binding pocket. Nucleotides 13–15 are the only nonconserved nucleotides in the internal loop besides A9, but there is some sequence preference for A. The propensity of the linker nucleotides to be A is consistent with maximizing stacking interactions and with the lack of direct interactions of the residues to the AMP in the binding pocket, since A has fewer functional groups than the other bases.

U16 is partially stacked on C15, but the base is rotated away from the single-strand linker to stack under G17. The G17 continues this rotation to base pair with G30, and this base pair is stacked on the first base pair (G29·C18) of the second stem. Thus, U16 and G17 are represented by the little turn at the top of the ζ, which continues on into the second helix. As discussed above, the appearance of a low-field-shifted, slowly exchanging 2′OH provides evidence for a hydrogen bond of the G30 2′OH. Based on the structures, two possibilities exist for this interaction. One is a hydrogen bond to the N7 of A12. This appears to be the most likely interaction because A12N7 also shows a large upfield shift of about 5 ppm relative to non-hydrogen-bonded N7s, although in most of the structures only the distance but not the geometry of the hydroxyl group is correct for a hydrogen bond. The alternative G30 2′OH hydrogen bond to the O4′ of AMP has the correct geometry, but the distance is too great.

AMP conformation and interactions

The AMP in the aptamer complex is in the *anti* conformation, with a χ angle of $-180° \pm 25°$. The pseudorotation angles for the majority of the AMP ribose vary around $80° \pm 10°$, with the center of the distribution in the O4′-endo range. The phosphate group is not restrained. The pyrimidine ring of the adenine is completely buried in, while its imidazole ring is pointing out of, the binding pocket. The C8 and N7 are accessible from the front of the binding pocket, and the 5′ phosphate group also lies outside (Fig. 5B). This positioning of the AMP within the binding pocket is consistent with how the selection was done, because the ATP was linked to the agarose column via the C8 (Fig. 8).

Binding of the AMP is stabilized by a number of stacking and hydrogen-bonding interactions. As mentioned above, the AMP base is stacked between A10

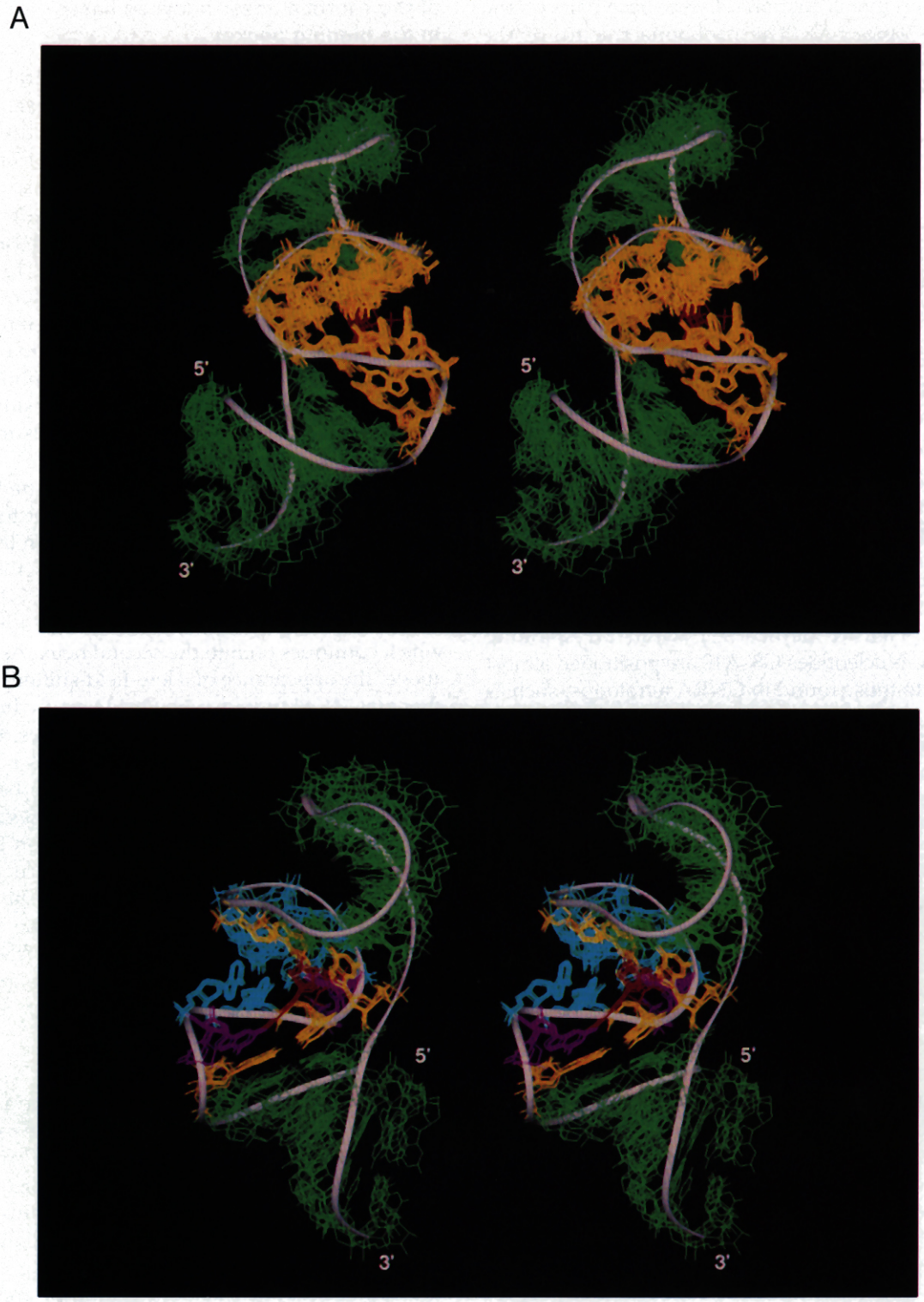

FIGURE 5. Stereo views of the three-dimensional structure of the ATP-binding RNA aptamer. **A:** Superpositions of the 10 lowest-energy structures of the entire molecule with the back of the AMP-binding pocket forward. Superposition is of all heavy atoms of residues 6–18, 29–31, and AMP. Stem base pairs are green, the internal loop is yellow, and the AMP is red. **B:** Superpositions of the five lowest-energy structures, viewed from the front of the binding pocket. **C:** Lowest-energy structure of the internal loop, flanking base pairs, and AMP, viewed from the front of the AMP-binding pocket. **D:** Same as C, but viewed from the side to show linker. For B–D, the stem base pairs are green; the G7·G11 and G17·G30 base pairs are yellow; A8 and A12 are purple; A9, A10, A13–U16 are cyan; and AMP is red. The AMP is displayed as a ball-and-stick model in C and D. (*Continued on facing page.*)

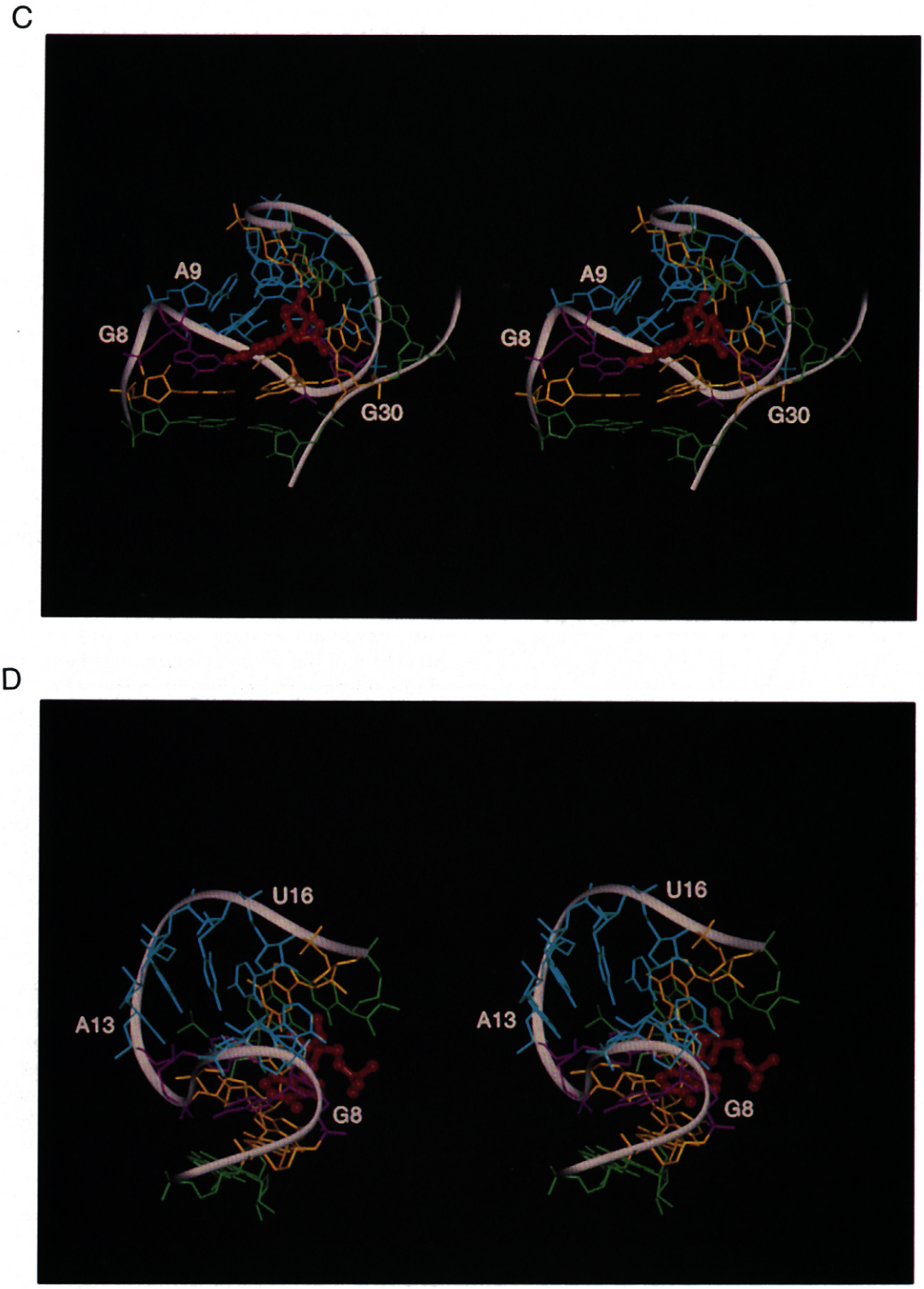

FIGURE 5. *Continued.*

and G11, whereas the AMP ribose is stacked against the G30·G17 base pair. The AMP is positioned to form a base pair with G8, in which the amino group of AMP is hydrogen bonded to G8N3, and N1 of AMP is hydrogen bonded to the amino group of G8. However, an additional or alternative hydrogen bond between the AMP amino and the O4′ of G8 is possible in almost all of the structures. A12 is hydrogen bonded to the N3 of AMP via its amino group. The U16 imino or carbonyl can potentially form a hydrogen bond with the 3′O or

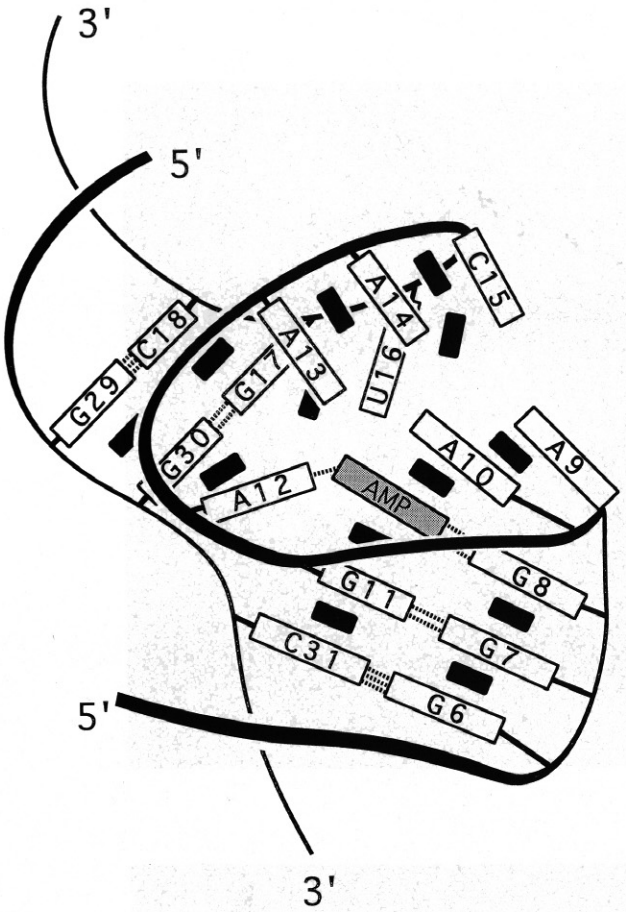

FIGURE 6. Schematic drawing of the ATP-binding aptamer/AMP complex illustrating the overall fold, stacking, and base pair interactions. The view shown is from the back of the binding pocket to illustrate the ζ backbone tracing. The orientation shown is the same as in Figure 5A. The U-turn is at G8-A9, the second turn is at G11-A12-A13, C15-A13 form the linker, and U16 and G17 rotate away from the stacked bases of the linker. All nucleotides but G30 are *anti*. The backbone tracing is shown by the solid curved lines, the nucleotides as numbered rectangles, stacking interactions as filled rectangles between nucleotides, and proposed base pair interactions as dotted lines.

3'OH, respectively, of the AMP. This is consistent with the fact that U16 shows a sharp imino resonance and would explain why U16 is a highly conserved base. There are several possible hydrogen-bonding interactions between the AMP 2'OH and conserved bases in the loop; the most frequent ones are G30N7, A12N1, and G17N3.

Evidence for base pair interactions within the binding pocket

The results of the structure calculations combined with specific proton and nitrogen chemical shifts and NOEs provide strong evidence for three base pairs in the tertiary structure of the aptamer complex (Fig. 9). In all of the DG/RMD structures, the G30·G17 base pair comes out oriented to pair as shown. Weak G17 imino–G30 imino and G17 imino–G30H8 NOEs observed in the NOESY spectra in H_2O are only consistent with this type of pairing, and the ^{15}N chemical shifts indicate that G17N1 is hydrogen bonded as well. The G7·G11 base pair also comes out as shown in the majority of the initial DG/RMD structures and in all of the refined structures; this arrangement is also supported by the NOE between G7 imino and G11H8 as well as the G7N1 chemical shift that indicates that the G7 imino is hydrogen bonded. Supporting evidence for a G8·AMP base pair is provided by the pattern of NOEs seen and the nitrogen chemical shifts. The G8 amino protons appear as two well-resolved resonances, which is quite unusual. NOEs are observed from the G8 amino to both the AMP amino and the H2. The amino nitrogens of AMP and G8 are shifted in the downfield direction consistent with hydrogen bonding, which is consistent with the proton chemical shifts as well.

Binding of AMP to mutant aptamers supports the proposed tertiary interactions

The binding results with the mutant aptamers (Table 1) can be explained on the basis of the three-dimensional structure of the wild-type aptamer. Examination of the deduced base pairs (Fig. 9) in the wild-type complex shows that for each aptamer in which substitution of a conserved base still retains binding (G7A, G8A, and G17A), the mutant A can bind in essentially the same position as the wild-type G, albeit with the loss of some hydrogen-bonding interactions. However, replacement of the other G in each of the proposed base pairs with an A would be sterically hindered, and consistent with this, those mutants (G11A, G30A) do not show appreciable binding. In the case of the G8A mutant, the U-turn hydrogen bond from the G amino to the A10 3' phosphate would be abolished as well as one potential hydrogen bond to the AMP, and consistent with this, the G8A mutant binds AMP less tightly than do G7A and G17A. Although only partial assignments for the mutated aptamer complexes have been obtained so far, the identified NOEs are also consistent with the structural predictions.

Analysis of data from chemical modification, ATP analogues, and other selections

The three-dimensional structure of the ATP-binding RNA aptamer largely explains the reported results of binding assays with ATP analogues (Sassanfar & Szostak, 1993). All of the ATP analogues that did not show binding have substitutions at positions that would delete or sterically block hydrogen-bonding interactions. The reported kethoxal modification data, which indicate that G8, G30, and to a lesser extent G11 are protected from modification by this reagent when

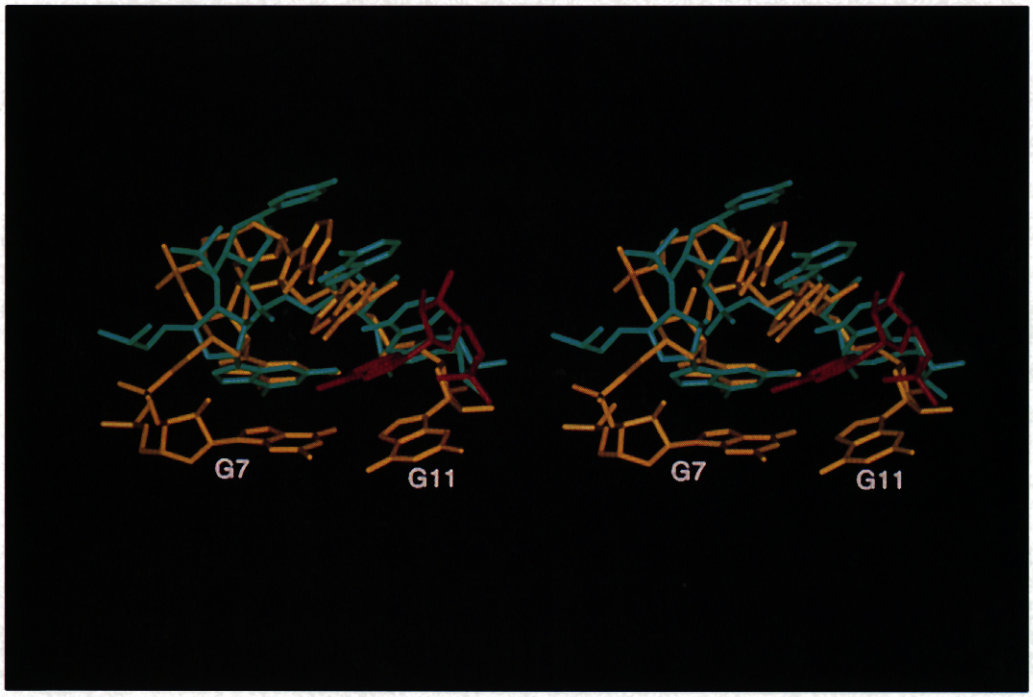

FIGURE 7. Stereo view of a superposition of the backbone G7-G8-A9-A10-G11 and AMP (yellow and red, respectively) and the GUAA tetraloop (green) from the crystal structure of the hammerhead ribozyme (Scott et al., 1995). The three-base U-turns (underlined bases above) have a backbone RMSD of 1.7 Å.

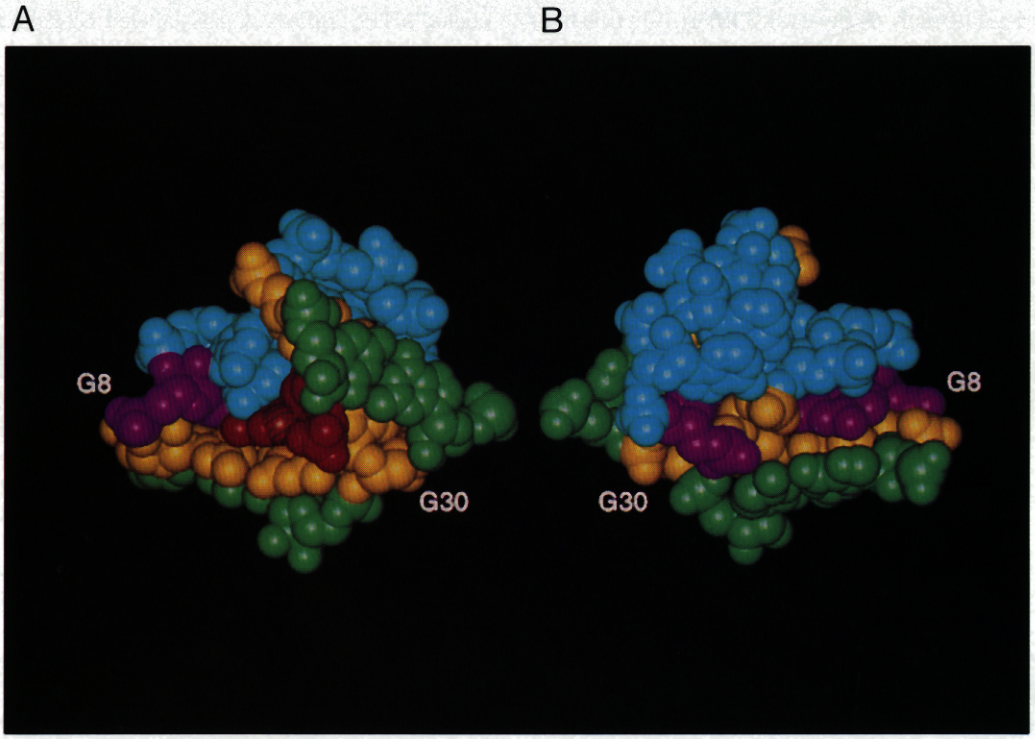

FIGURE 8. Space-filling representations of the ATP-binding domain, residues 6–18, 29–31, and AMP. **A:** View from the front of the binding pocket ("view from the column"). **B:** View from the back of the binding pocket. Colors are the same as in Figure 5B–D. The C8 atom on the AMP is striped.

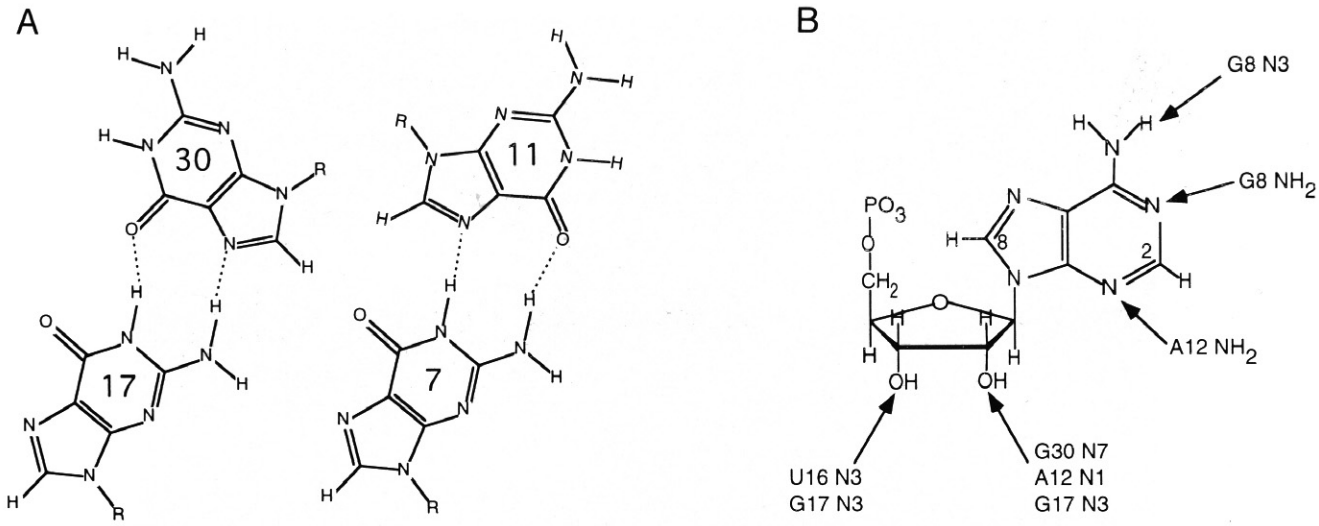

FIGURE 9. **A:** G17·G30 and G7·G11 base pairs proposed to form in the complex. **B:** Proposed hydrogen-bonding interactions between the AMP and the RNA nucleosides.

ATP is bound to the RNA aptamer, are also consistent with the compact and extensively stacked loop structure. The higher reactivities of G6, G7, and G17 may be attributed to their locations at the helix–loop interfaces, although the N1H and/or amino protons of the bases are hydrogen bonded.

The consensus sequence for the ATP-binding RNA aptamer was incorporated into the RNA library constructed for selection of ribozymes with polynucleotide kinase activity (Lorsch & Szostak, 1994). Of the four classes of kinases that retain ATP-binding aptamer-like consensus sequences, we would predict that only the Class IV kinase, which has a single A10G mutation, should still bind ATP with near wild-type affinity. The other three classes of kinases all have mutations that should strongly decrease binding. This is consistent with what is reported for binding of ATP-γS, which was used in selection for the kinase activity.

Recently, the solution structure of an FMN/aptamer complex was reported (Fan et al., 1996). In contrast to the complex interactions for the AMP found in the ATP-binding aptamer/AMP complex, the FMN binds via a Hoogsteen-like interaction between an A of the aptamer and the U-like moiety of the FMN that is stabilized by intercalation into the RNA duplex.

Summary

The structure of the ATP-binding RNA aptamer/AMP complex provides a rational basis for thinking about how ribozymes could incorporate biological cofactors to catalyze reactions. Although AMP is normally found in RNA, the selected binding site does not incorporate a standard Watson–Crick base pair and simple stacking interactions. Rather, a complicated binding pocket is formed that uses all of the conserved bases in the consensus sequence.

MATERIALS AND METHODS

Sample preparation

RNA was synthesized enzymatically on a DNA template using T7 RNA polymerase (Milligan et al., 1987) and unlabeled NTPs and/or ^{13}C,^{15}N-labeled NTPs to prepared unlabeled, base-specifically G-, A-, or U/C-labeled, or fully labeled RNA samples. The labeled NMPs were isolated from methanolotrophic bacteria that had been grown on ^{13}C-methanol and ^{15}N-ammonia, purified, and converted to NTPs as described (Batey et al., 1992; Nikonowicz et al., 1992; Peterson et al., 1994). For the base-specifically labeled samples, the labeled NMPs used in the enzymatic synthesis were first purified by HPLC using a semi-preparative scale C18 column (Waters PrepPak Cartridge, 25 × 100 mm) with a water/methanol gradient. Samples of the wild-type RNA aptamer used in obtaining assignments included: (1) unlabeled RNA with unlabeled AMP, (2) unlabeled RNA with labeled AMP, (3) unlabeled RNA with labeled ATP, (4) labeled RNA with unlabeled AMP, (5) G-labeled RNA with unlabeled AMP, (6) A-labeled RNA with unlabeled AMP, and (7) C,U-labeled RNA with unlabeled AMP in both H_2O and D_2O. Samples of the mutant RNAs that bound AMP well enough to give interpretable spectra (G7A, G8A, A9U, C15U, G17A, C18UG29A) were also used to confirm assignments. NMR samples were typically 0.5–1.6 mM in RNA, pH 6.0, 100 mM NaCl.

NMR spectroscopy

All heteronuclear NMR spectra were acquired at 500 MHz on Bruker AMX and DRX spectrometers. Homonuclear spectra were acquired at 500 and 750 MHz on Bruker DRX and DMX spectrometers. Assignments of the nonexchangeable reso-

nances were obtained by combining information from a variety of heteronuclear and homonuclear NMR experiments. The heteronuclear experiments used were 2D HCN (Sklenář et al., 1993a), ^{13}C-CT-HSQC (Santoro & King, 1992), ^{13}C- or ^{15}N-filtered NOESY (Otting & Wüthrich, 1990), ^{13}C-selected NOESY (Otting & Wüthrich, 1990), ^{1}H-^{15}N long-range HSQC (Sklenář et al., 1994), 3D ^{1}H-^{13}C-NOESY-HMQC (Marion et al., 1989b), 3D ^{1}H-^{15}N-NOESY-HMQC (Marion et al., 1989a), and 3D HCCH TOCSY (Bax et al., 1990; Clore et al., 1990). Heteronuclear experiments on the base-specifically labeled samples were particularly useful for assignment and identification of NOE contacts. The homonuclear experiments used were 2D and 3D NOESY (Kumar et al., 1980), DQF-COSY (Piantini et al., 1982), ROESY (Bothner-By et al., 1984), and TOCSY (Bax & Davis, 1985). Assignments of the exchangeable resonances were obtained from 2D HCCNH-TOCSY (Sklenář et al., 1996), ^{15}N-^{1}H HSQC (Sklenář et al., 1993b), and ^{15}N-filtered NOESY (Otting & Wüthrich, 1990). Water suppression was achieved by using the 11 spin-echo (Otting & Wüthrich, 1990) or WATERGATE (Piotto et al., 1992; Sklenář et al., 1993b) sequences. All data were processed using the Bruker XWINNMR software and analyzed with the assistance of AURELIA.

Structure calculations

Interproton distance restraints were obtained from NOESY spectra of the sample in D_2O at 10 °C with mixing times of 150 and 300 ms at 750 MHz, as well as 40, 80, 100, 150, 200, 250, and 300 ms at 500 MHz and in H_2O at 5 °C with mixing times of 100 and 150 ms. For the D_2O NOEs, crosspeaks were picked from the 750-MHz 300-ms NOESY, and the other mixing times were used for confirmation and calibration of the NOE intensity. NOE crosspeak intensities were classified as strong (2.5 + 0.6 Å), medium strong (3.0 + 0.8 Å), medium (3.5 + 0.9 Å), medium weak (4.0 + 1.0 Å), weak (4.5 + 1.5 Å), and very weak (5.0 + 1.8 Å). An additional upper bound of 1.0 Å was used for the exchangeable NOEs and in the case of NOEs involving pseudoatoms. All lower bounds were set to van der Waals distance. Dihedral angle restraints ν_1 and ν_2 for the riboses were obtained from analysis of DQF COSY spectra; residues with a strong H1′-H2′ crosspeak were restrained to S-type, those with no detectable H1′-H2′ crosspeak were restrained to N-type, and all others were left unrestrained. (None of the residues show significant differences in linewidths of the H1′ and H2′ resonances, so absence of a crosspeak is not due to linewidth.) All structure calculations were done with X-PLOR version 3.1 (Brünger, 1992). For the internal loop, base pairs flanking the loop, and AMP, input data for the structure calculations included only the experimentally derived restraints listed above. Although NOE and dihedral angle restraints were included for the stem and UUCG tetraloop, the stems (except the base pairs flanking the internal loop) were constrained to standard A-form geometry. Initially, 100 starting structures were generated using distance geometry substructure embedding (Havel & Wüthrich, 1984). These structures were then subjected to simulated annealing (25 ps at 2,000 K, followed by 15 ps of cooling to 100 K, with time steps of 1 fs). At this point the 22 lowest-energy structures were analyzed for potential hydrogen-bonding interactions, and weak hydrogen-bond restraints were introduced for those internal loop base pairs (G7·G11, G17·G30) for which independent evidence of hydrogen bonding (slowly exchanging protons, ^{1}H and ^{15}N chemical shifts) had been obtained from the NMR data. No hydrogen-bond restraints were introduced for the AMP interactions. These structures were then subjected to an additional round of simulated annealing (cooling from 2,000 K to 100 K in 35 ps with 1-fs time step). In a final molecular dynamics step, the full van der Waals term was introduced (40 ps at 300 K with 2-fs time step). The 10 lowest-energy structures were used in the structural analysis.

Coordinates for the structures will be deposited in the Protein Data Bank, Brookhaven National Laboratory.

ACKNOWLEDGMENTS

We thank Professor J.W. Szostak for advice and encouragement and for the deoxyribose substitution data, Dr. Samuel Butcher for many helpful discussions, and Dr. Flint Smith for help with the illustrations. T.D. was supported by an HFSPO postdoctoral fellowship, and J.F. acknowledges grant support from NIH and NSF. We also acknowledge the Stable Isotope Facility at Los Alamos for labeled cells and the National NMR Facility at Madison for 750-MHz NMR spectrometer time.

Manuscript accepted without revision June 20, 1996

REFERENCES

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

Batey RT, Inada M, Kujawinski E, Puglisi JD, Williamson JR. 1992. Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA. *Nucleic Acids Res 20*:4515–4523.

Bax A, Clore GM, Gronenborn AM. 1990. ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for assigning ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J Magn Reson 88*:425–431.

Bax A, Davis DG. 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn Res 65*:355–360.

Bothner-By AA, Stephens RL, Lee Jm, Warren CD, Jeanloz RW. 1984. Structure determination of a tetrasaccharide: Transient nuclear Overhauser effect in the rotating frame. *J Am Chem Soc 106*:811–813.

Brünger AT. 1992. *X-PLOR (version 3.1) manual*. New Haven, Connecticut/London: Yale University Press.

Burgstaller P, Famulok M. 1994. Isolation of RNA aptamers for biological cofactors by in vitro selection. *Angew Chem Int Ed Engl 33*:1084–1087.

Clore GM, Bax A, Driscoll PC, Wingfield PT, Gronenborn AM. 1990. Assignment of the side-chain ^{1}H and ^{13}C resonances of interleukin-1β using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy. *Biochemistry 29*:8172–8184.

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

Ellington AD, Szostak JW. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature 346*:818–822.

Fan P, Suri AK, Fiala R, Live D, Patel DJ. 1996. Molecular recognition in the FMN–RNA aptamer complex. *J Mol Biol 258*:480–500.

Fesik SW, Zuiderweg ERP. 1988. Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. *J Magn Reson 78*:588–593.

Havel T, Wüthrich K. 1984. A distance geometry program for determining the structures of small proteins and other macromolecules from nuclear magnetic resonance measurements of intermolecular ^{1}H-^{1}H proximities. *Bull Math Biol 46*:673–698.

Heus HA, Pardi A. 1991. Structural features that give rise to the un-

usual stability of RNA hairpins containing GNRA loops. *Science 253*:191–194.

Jucker FM, Pardi A. 1995. GNRA tetraloops make a U-turn. *RNA 1*:219–222.

Kumar A, Ernst RR, Wüthrich K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Commun 95*:1–6.

Lorsch JR, Szostak JW. 1994. In vitro evolution of new ribozymes with polynucleotide kinase activity. *Nature 371*:31–36.

Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. 1989a. Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H-^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: Application to naturenterleukin 1β. *Biochemistry 28*:6150–6156.

Marion D, Kay LE, Sparks SW, Torchia DA, Bax A. 1989b. Three-dimensional heteronuclear NMR of ^{15}N-labeled proteins. *J Am Chem Soc 111*:1515–1517.

Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. 1987. Oligoribonucleotide synthesis using T7-RNA polymerase and synthetic DNA templates. *Nucleic Acids Res 15*:8783–8798.

Nikonowicz EP, Sirr A, Legault P, Jucker FM, Baer LM, Pardi A. 1992. Preparation of C-13 and N-15 labeled RNAs for heteronuclear multi-dimensional NMR studies. *Nucleic Acids Res 20*:4507–4513.

Otting G, Wüthrich K. 1990. Heteronuclear filters in two-dimensional [^{1}H,^{1}H]-NMR spectroscopy: Combined use with isotope labeling for studies of macromolecular conformation and intermolecular interactions. *Q Rev Biophys 23*:39–96.

Pardi A. 1995. Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. *Methods Enzymol 261*:350–380.

Peterson RD, Bartel DP, Szostak JW, Horvath SJ, Feigon J. 1994. 1H NMR studies of the high-affinity Rev binding site of the Rev responsive element of HIV-1 mRNA: Base pairing in the core binding element. *Biochemistry 33*:5357–5366.

Piantini U, Sørensen OW, Ernst RR. 1982. Multiple quantum filters for elucidating NMR coupling networks. *J Am Chem Soc 104*:6800–6801.

Piotto M, Saudek V, Sklenář V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J Biomol NMR 2*:661–665.

Pley HW, Flaherty KM, McKay DB. 1994. Three-dimensional structure of a hammerhead ribozyme. *Nature 372*:68–74.

Quigley GJ, 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.

Santoro J, King G. 1992. A constant-time 2D Overbodengausen experiment for inverse correlation of isotopically enriched species. *J Magn Reson 97*:202–207.

Sassanfar M, Szostak JW. 1993. An RNA motif that binds ATP. *Nature 364*:550–553.

Scott WG, Finch JT, Klug A. 1995. The crystal structure of an all-RNA hammerhead ribozyme—A proposed mechanism for RNA catalytic cleavage. *Cell 81*:991–1002.

Sklenář V, Dieckmann T, Butcher SE, 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.

Sklenář V, Peterson RD, Rejante MR, Feigon J. 1993a. Two and three-dimensional HCN experiments for correlating base and sugar resonances in ^{15}N, ^{13}C labeled RNA oligonucleotides. *J Biomol NMR 3*:721–727.

Sklenář V, Peterson RD, Rejante MR, Feigon J. 1994. Correlation of nucleotide base and sugar protons in ^{15}N labeled HIV RNA oligonucleotide by ^{1}H-^{15}N HSQC experiments. *J Biomol NMR 4*:117–122.

Sklenář V, Piotto M, Leppik R, Saudek V. 1993b. Gradient-tailored water suppression for ^{1}H-^{15}N HSQC experiments optimized to retain full sensitivity. *J Magn Reson Ser A 102*:241–245.

Szostak JW. 1992. In vitro genetics. *Trends Biochem Sci 17*:89–93.

Tuerk C, Gold L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science 249*:505–510.

Woese CR, Winker S, Gutell RR. 1990. Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops." *Proc Natl Acad Sci USA 87*:8467–8471.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: John Wiley & Sons.