Solution structure of an informationally complex high-affinity RNA aptamer to GTP

JAMES M. CAROTHERS,[1] JONATHAN H. DAVIS,[2] JAMES J. CHOU,[3] and JACK W. SZOSTAK

Department of Molecular Biology and Center for Computational and Integrative Biology, Simches Research Center 7215, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

ABSTRACT

Higher-affinity RNA aptamers to GTP are more informationally complex than lower-affinity aptamers. Analog binding studies have shown that the additional information needed to improve affinity does not specify more interactions with the ligand. In light of those observations, we would like to understand the structural characteristics that enable complex aptamers to bind their ligands with higher affinity. Here we present the solution structure of the 41-nt Class I GTP aptamer (K_d = 75 nM) as determined by NMR. The backbone of the aptamer forms a reverse-S that shapes the binding pocket. The ligand nucleobase stacks between purine platforms and makes hydrogen bonds with the edge of another base. Interestingly, the local modes of interaction for the Class I aptamer and an RNA aptamer that binds ATP with a K_d of 6 μM are very much alike. The aptamers exhibit nearly identical levels of binding specificity and fraction of ligand sequestered from the solvent (81%–85%). However, the GTP aptamer is more informationally complex (~45 vs. 35 bits) and has a larger recognition bulge (15 vs. 12 nucleotides) with many more stabilizing base–base interactions. Because the aptamers have similar modes of ligand binding, we conclude that the stabilizing structural elements in the Class I aptamer are responsible for much of the difference in K_d. These results are consistent with the hypothesis that increasing the number of intra-RNA interactions, rather than adding specific contacts to the ligand, is the simplest way to improve binding affinity.

Keywords: aptamer; NMR; affinity; in vitro selection; information; complexity

INTRODUCTION

The amount of information needed to specify an RNA structure in a particular environment (Schneider et al. 1986; Adami and Cerf 2000; for review, see Adami 2004) increases with the level of activity the structure exhibits (Carothers et al. 2004). For a set of 11 different RNA aptamers to guanosine triphosphate (GTP) with dissociation constants (K_d's) that range from 8 μM to 9 nM, each 10-fold improvement in binding affinity requires a structure that is about 1000 times less frequent in a pool of random sequences. We noted the same quantitative relationship between informational complexity and functional activity in two RNAs that catalyze an RNA–RNA ligation reaction. These observations led to the supposition that the similarity of the informational costs, ~10 bits per 10-fold improvement in activity, is due to common underlying physical phenomena.

We hypothesized that the informational cost to improve activity could be the same for different functions if the added information content is used primarily to specify more stable conformations, rather than more or better contacts with a ligand or reaction substrate (Carothers et al. 2004). Others have emphasized that a generalizable information–function relationship may be a natural consequence of the scale of monomer subunits in RNA relative to small molecule ligands or substrates (Lau et al. 2004). With greater understanding of how and why information and function are related in RNA structures it may be possible to devise better methods for creating biosensors and therapeutics (for review, see Sullenger and Gilboa 2002; Rimmele 2003) and gain insight into the origins and evolution of biochemical activity (Carothers and Szostak 2006).

The set of GTP aptamers discussed above was obtained by in vitro selection from a random sequence library where half of the molecules contained a designed stem–tetraloop (Davis and Szostak 2002). This partially engineered sublibrary was clearly a better source of high-affinity aptamers than the completely random sublibrary. Seven of the 11

aptamers came from the partially engineered library, including six of the seven molecules with K_d's for GTP that are >300 nM and all four molecules with K_d's >100 nM (Carothers et al. 2004).

Since in vitro selection produces the simplest solutions to a given problem (Lorsch and Szostak 1996), in vitro-selected aptamers are likely to employ the simplest possible mechanisms for attaining the binding affinities that they possess. Surprisingly (Eaton et al. 1995), analog binding studies (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.) showed that the more informationally complex higher-affinity GTP aptamers do not necessarily make a greater number of specific contacts with the ligand than the lower-affinity aptamers. This finding supports the notion that it is easier to improve aptamer activity by making more contacts within the RNA itself than by making more contacts with the ligand.

Here we report the solution structure of the Class I GTP aptamer (Fig. 1), determined by NMR spectroscopy. The K_d of this aptamer for GTP is 75 nM under the conditions used for its selection. The Class I aptamer is comprised of 41 nt. The asymmetric internal recognition bulge contains a total of 15 residues (9 nt in L1 and 6 nt in L3) that are flanked on one side by a 7-bp external stem (P1). The 4-bp internal stem (P2) and tetraloop (L2) at the other side of the recognition bulge are derived from the partially designed sequence library (Davis and Szostak 2002). The solution structure shows the importance of the stems and designed-tetraloop in anchoring the overall fold of the molecule. We find that a number of the invariant bases within the binding bulge do not contact the GTP ligand at all; instead, much of the information content is needed to specify a network of interactions that stabilize the active
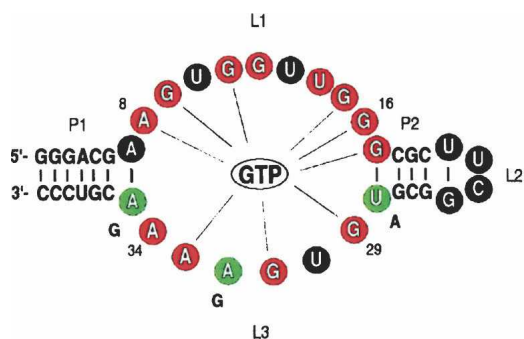

FIGURE 1. Secondary structure of the Class I GTP aptamer. Lines connect base-pairing partners in the P1 and P2 stems. Positions in the recognition bulge (L1, residues 8–16; L3, residues 29–34) are colored according to whether they are invariant (red), two-base varying (green), or unconserved (black) (Carothers et al. 2004). The terminal base pair of the outer stem P1 (A7:A35) and inner stem P2 (G17:U28) are noncanonical pairs and are colored in the same way. For sites that are two-base varying, the tolerated substitution is shown *below* the position. The lines connecting aptamer residues to the GTP ligand indicate positions for which we detected intermolecular aptamer-GTP NOEs.

conformation of the aptamer. At the same time, we can now understand how sequence variation is tolerated at other positions within the recognition bulge. Although there are specific intermolecular interactions between the aptamer and parts of the GTP, these occur at few enough places, and within a large enough pocket, that several analogs of GTP are readily accommodated.

We compare the Class I GTP aptamer structure with other in vitro-selected aptamers that bind ligands with planar moieties, particularly the "Sassanfar" ATP aptamer (K_d = 6 μM; Sassanfar and Szostak 1993; Dieckmann et al. 1996; Jiang et al. 1996), and find that the modes of ligand recognition are very similar. We suggest that additional stabilizing intra-aptamer interactions (as opposed to aptamer–ligand interactions) confer most of the 80-fold-improved binding affinity of the Class I aptamer relative to the Sassanfar aptamer. These results advance the hypothesis that the simplest way to improve activity is to increase the stability of the active form of the RNA. We discuss these observations in terms of a larger effort to understand how sequence information is related to functional activity.

RESULTS

Resonance assignments and structure determination by NMR

The bulk of the Class I aptamer (Fig. 1) resonance assignments was made using standard techniques (Lukavsky and Puglisi 2001) and a sample consisting of uniformly labeled ^{13}C/^{15}N Class I RNA and unlabeled GTP ligand. The aptamer becomes well ordered upon the addition of slight molar excesses of Mg^{2+} and GTP ligand. Representative spectra are shown in Figure 2.

Sequence covariation and mutation data (Carothers et al. 2004) suggested that residues 1–6, 18–20, 25–27, and 36–41 are involved in standard Watson-Crick (W-C) base pairs. As expected, NOESY-derived interproton distances for these nucleotides were consistent with W-C base-pairing and A-form geometry, and it was straightforward to make sequential assignments for those regions. Sequential assignments for positions within the recognition bulge were made using 3D HCP data and several segmentally labeled samples (see Materials and Methods). Semiquantitative interproton distances for residues in the recognition bulge were obtained from 2D homonuclear and ^{13}C- or ^{15}N-edited 3D NOESY spectra. COSY and TOCSY data were used to ascribe dihedral angle values for the sugar puckers. NOE data were employed to determine dihedral angles for the glycosidic bonds (X) (see Materials and Methods).

The resonances of the GTP were assigned using a sample of unlabeled RNA and ^{13}C/^{15}N-labeled ligand. 2D HSQC experiments revealed well-resolved peaks corresponding to the sugar, aromatic, and H1 imino protons. We detected 14 well-resolved intermolecular aptamer-GTP NOEs, involv-
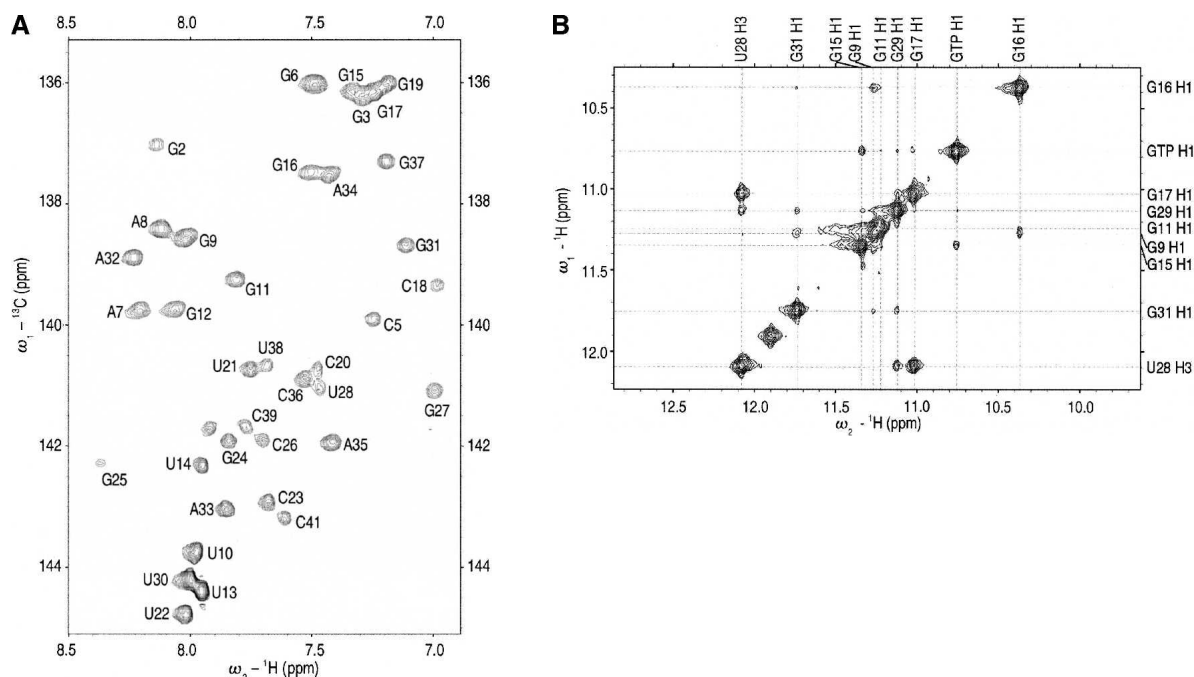
FIGURE 2. NMR spectra of the Class I GTP aptamer. (*A*) The aromatic region of a constant time ^{13}C-HSQC experiment at 15°C with uniformly labeled ^{13}C/^{15}N Class I aptamer and unlabeled GTP ligand. (*B*) Imino–imino region of the recognition bulge in the 100 msec mixing time 2D NOESY in 95% H$_2$O/5% D$_2$O at 15°C. The dashed lines connect NOE cross-peaks among the U (H3) and G (H1) imino protons within the recognition bulge of the Class I aptamer. The resonance at 10.77 ppm marked GTP H1 is the imino proton of the GTP ligand. The H1 imino proton of G6 (11.92 ppm) is not marked in this figure because it is not part of the recognition bulge.

ing eight of the 11 invariant positions within the recognition bulge (Fig. 1). No intermolecular NOEs involving unconserved or two-base varying positions were observed.

We measured a set of aromatic base C-H residual dipolar couplings (RDCs) using Pf1 phage as the alignment medium (Hansen et al. 1998). Systematic grid searches (Tjandra and Bax 1997; see Materials and Methods) and least squares fits (Dosset et al. 2001) yielded positive values for the axial component of the alignment tensor, indicating that the main axis of the alignment tensor is parallel to the magnetic field. This is consistent with an extended conformation for the aptamer with coaxial stems aligned parallel to the phage (Bondensgaard et al. 2002).

We calculated 100 structures using molecular dynamics with NOE and dihedral angle restraints (see Materials and Methods). The 20 lowest energy structures with no NOE violations >0.2 Å and no dihedral angle violations >5° were further refined with the set of RDCs. Table 1 summarizes the restraints used in the calculations and the statistics for the final 10 structures with the lowest energies after refinement with the RDCs. The converged ensemble and its minimized average are shown in Figure 3. The rmsd to the average structure for all heavy atoms in the ensemble including the GTP ligand is 3.4 Å. For the recognition bulge (residues 8–16 and 29–34) and GTP ligand the rmsd to the average structure is 1.97 Å; the rmsd to the average structure for the recognition bulge and GTP is only 1.47 Å if we

exclude the three unconserved U's that function only as pivots and spacers (U10, U13, and U30; see below).

Stems anchor the overall fold

The structure of the Class I aptamer displays a broadly helical form (Fig. 3B) with stretches of base-paired nucleotides (P1 and P2 in Fig. 1) flanking both sides of the recognition bulge. The backbone of the 5′ strand of the recognition bulge (L1, residues 8–16) has the shape of a reverse-S with two sharp turns. The 3′ strand of the recognition bulge (L3, residues 29–34) is shorter and continues the helix with less deviation than the 5′ strand. The GTP ligand is sandwiched into a deep pocket formed by nucleotides from both L1 and L3.

The stems are capped by non-Watson-Crick base pairs with a particular polarity, required so that nucleotides from the recognition bulge lie against the flanking helices. Even though position 7 in P1 is unconserved, its pairing partner at position 35 must be a purine. In the particular sequence variant studied here, the terminal bases of P1 are an unusual A:A pairing between A7 and A35 (Fig. 4A). The X conformation of the A35 base brings the H2 protons of A7 and A35 close enough to one another that we observe a strong NOE between them. Within the experimental error, both the distance (~2.4 Å) and the angle (170°) between the nearer A35 NH6 and A7 N3 are consistent with a

TABLE 1. Statistics of the 10 lowest energy structures

Number of restraints used in structure calculations	
NOE distance	218
Experimentally derived	113
(residues 7–17, 28–35)	99
(intranucleotide)	36
(intermolecular RNA-GTP)	14
Total involving exchangeable protons	47
Model-derived A-form stems	
(residues 1–6, 18–20, 25–27, 36–41)	77
Model-derived UUCG tetraloop	
(residues 21–24)	28
Dihedral angle	149
Dipolar couplings	25
Stem hydrogen bonds	42
Total no. of restraints	434
Average rmsd from experimental restraints	
Distance (no violations >0.2 Å)	0.02 Å
Dihedral (no violations >5°)	0.14°
RDC (no violations >5 Hz)	1.6 Hz
Average rmsd from ideal geometries	
Bonds	0.006 Å
Angles	1.21°
Impropers	0.74°
Heavy atom rmsd from mean structure	
All residues (1–41, GTP)	3.38 ± 0.77 Å
Recognition bulge and ligand (residues 8–16, 29–34, GTP)	1.97 ± 0.34 Å
Recognition bulge and ligand without unconserved U's (see text) (residues 8, 9, 11, 12, 14–16, 29, 31–34, GTP)	1.47 ± 0.38 Å
Recognition bulge alone without unconserved U's (see text) (residues 8, 9, 11, 12, 14–16, 29, 31–34)	1.47 ± 0.36 Å
Ligand alone (GTP)	1.35 ± 0.37 Å

Unconserved positions are pivots and spacers

There are three unconserved nucleotides, U10, U13, and U30, within the recognition bulge that extend into the solvent and function as spacers and pivots rather than points of interaction. Four nucleotides comprise each turn of the reverse-S in the L1 recognition strand (A8–G9–U10–G11 and G11–G12–U13–U14). In both cases, the first, second, and fourth positions are invariant residues that make interactions with other parts of the aptamer, whereas the third position is the unconserved U. These turns are reminiscent of "hook turns" with a flipped-out base at the apex, except that here the 180° change in the direction of the backbone is achieved over the span of 3 nt rather than 2 nt (Szép et al. 2003). The bases of G29 and G31 are brought close enough together that we detect an NOE between the G29 H1 imino and the G31 H1 imino protons. A 90° kink in the backbone at the unconserved position 30 allows the X *syn* base of G29 to push toward the middle of the aptamer and take part in a base triple with G9 and G16. The kink pulls G31 back relative to G29 so that the GTP ligand sits under the base of G9 and against the W-C face of G31.

There are several lines of experimental evidence consistent with the idea that these three unconserved positions serve only as pivots and spacers. None of these U's have the detectable imino proton resonances at 15°C that would be expected if they made stable interactions on their W-C face. Each has intraresidue NOEs between the sugar protons and the aromatic H5 and H6 protons but no interresidue NOEs involving the aromatic protons. Finally, the U's have overlapping chemical shifts for all the spin systems in the sugar and aromatic base (e.g., Fig. 2A), indicating that they inhabit similar environments. Coincident with their roles as flexible pivots and spacers, U10, U13, and U30 are disordered within the ensemble of structures (Fig. 3A).

hydrogen bond. The pairing may be further stabilized by a CH–N hydrogen bond between A7 CH2 and A35 N1 as the A7 H2–A35 N1 distance is ~2.0 Å and the A7 CH2–A35 N1 angle is ~175°.

The terminal base pair of the internal stem P2 is more constrained at the sequence level than the terminal base pair of the external stem P1. Position 17 must be a G and its pairing partner at position 28 can only be a U or an A. A strong G17 H1 imino-U28 H3 imino NOE indicates that the G17:U28 pairing adopts a G:U wobble conformation with two hydrogen bonds (Chen et al. 2000). A normal G17:A28 imino pairing would be expected to have roughly the same geometry with respect to the position of the G17 base (Nagaswamy et al. 2002). In either case, the pairing broadens the helix relative to a standard W-C interaction. One consequence of the helix broadening is that the base of residue 28 is pushed toward a position where it can partially stack on top of G9. It is also possible that G17:U/A28 is a site for metal binding. The Class I aptamer requires a slight molar excess of Mg^{2+} for maximum binding affinity and both G:U and G:A are known to serve as specific Mg^{2+} binding sites in RNA (Feig and Uhlenbeck 1999).

Base–base interactions stabilize the core

The core of the aptamer consists of platforms of nucleotides making base–base interactions. The base–base contacts hold the strands of the recognition bulge together, shape the binding pocket (Fig. 4B–F), and explain the observed sequence conservation.

Orienting the aptamer with stem P1 at the bottom (Fig. 3), G12 and A34 (Fig. 4B) comprise the lowermost level of the recognition bulge and are the only two invariant purines far enough from the GTP that they show no NOEs with the ligand. The base of G12 partially stacks on top of A35, while A34 lies across the A7:A35 base pair that caps the outer stem. NOEs between G12 H1–A34 H2 and G12 H1–A34 H1′, and a very weak resonance for G12 H1 itself, are indicative of the imino–N3 single hydrogen bond pairing (Nagaswamy et al. 2002) in the structure. The sequence conservation data (Davis and Szostak 2002; Carothers et al. 2004) demonstrate that A34G variants are not functional, even though the N3 hydrogen bond acceptor
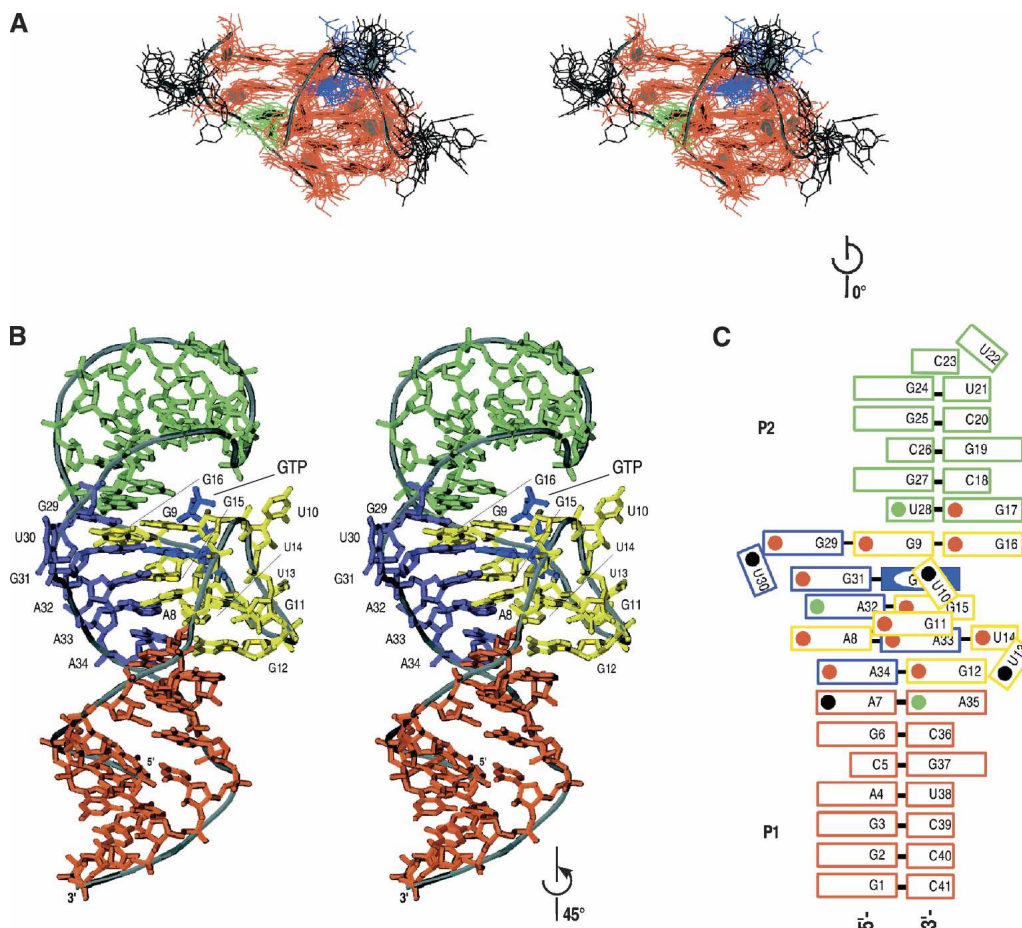
FIGURE 3. Class I GTP aptamer structure. (*A*) Ensemble of 10 lowest energy structures. Stereo view of the superimposed recognition bulge structures (residues 8–17, 29–34) overlaid on the average structure (gray). The residues are colored according to sequence conservation as in Figure 1: red = invariant, green = two-base varying, black = unconserved. The GTP ligand is colored blue, oriented with the disordered triphosphate region pointing upward. (*B*) Minimized average structure. Stereo view of the average structure of the converged ensemble after energy minimization. Nucleotides from the outer stem (P1) are colored orange. Nucleotides from the inner stem and UUCG tetraloop (P2) are colored green. Nucleotides from the 5′ strand of the recognition bulge (L1) are yellow. Nucleotides from the 3′ strand of the recognition bulge (L3) are purple. The GTP ligand is blue. The 5′ and 3′ ends of the aptamer are shown. The complex is rotated ∼45° counterclockwise relative to the image in *A*. (*C*) Cartoon model of the Class I GTP aptamer. In this representation, the *bottom* stem is P1. The outlining corresponds to the colors in *B* (e.g., orange = P1 nucleotides, green = P2, etc.). The nucleotide positions are colored with circles as in Figure 1 according to sequence conservation: red = invariant positions, green = two-base varying positions, black = unconserved positions. The 5′ and 3′ ends are marked.

would be preserved. Presumably, the need to tie down the second turn of the reverse-S restricts the angle between the W-C face of G12 and the minor-groove face of A34 such that an A34G substitution cannot be sterically accommodated.

The reverse W-C pairing of U14:A33 (Fig. 4C) is the only instance in the core of the aptamer where the interaction takes place exclusively among W-C faces. In reverse W-C pairings the backbone of the strand containing the U is inverted with respect to a standard A:U such that U O2, rather than U O4, hydrogen bonds with an A H6 amino proton. A8 lies in the same plane as U14:A33 with an A33 amino proton close enough to hydrogen bond with A8 N1. In the averaged structure, the nearest A8 H6 is ∼4 Å from A33 N7, though we do detect two distinct but extremely exchange-broadened A8 H6 resonances, which could signify a A8 H6–A33 N7 hydrogen bond.

G15:A32 (Fig. 4D) sits above A8:A33:U14 and forms the floor of the binding pocket. G11 lies partly between these two levels, beneath the GTP. The G11 W-C face is positioned to interact with the minor groove edge of A8 (Fig. 4C). Although another base triple could potentially replace A8:A33:U14, it seems unlikely that any substitution exists that would simultaneously maintain the G11:A8 interaction.

Aside from the three unconserved pivots and spacers, A32 occupies the only other position within the recognition bulge where sequence variation is tolerated. The G31 nucleobase, which directly contacts the GTP ligand, stacks on top of A32. Both the G15 H1–A32 N3 hydrogen bond implicated by the NMR structure (Fig. 4D) and stacking with G31 could persist in an A32G substitution.

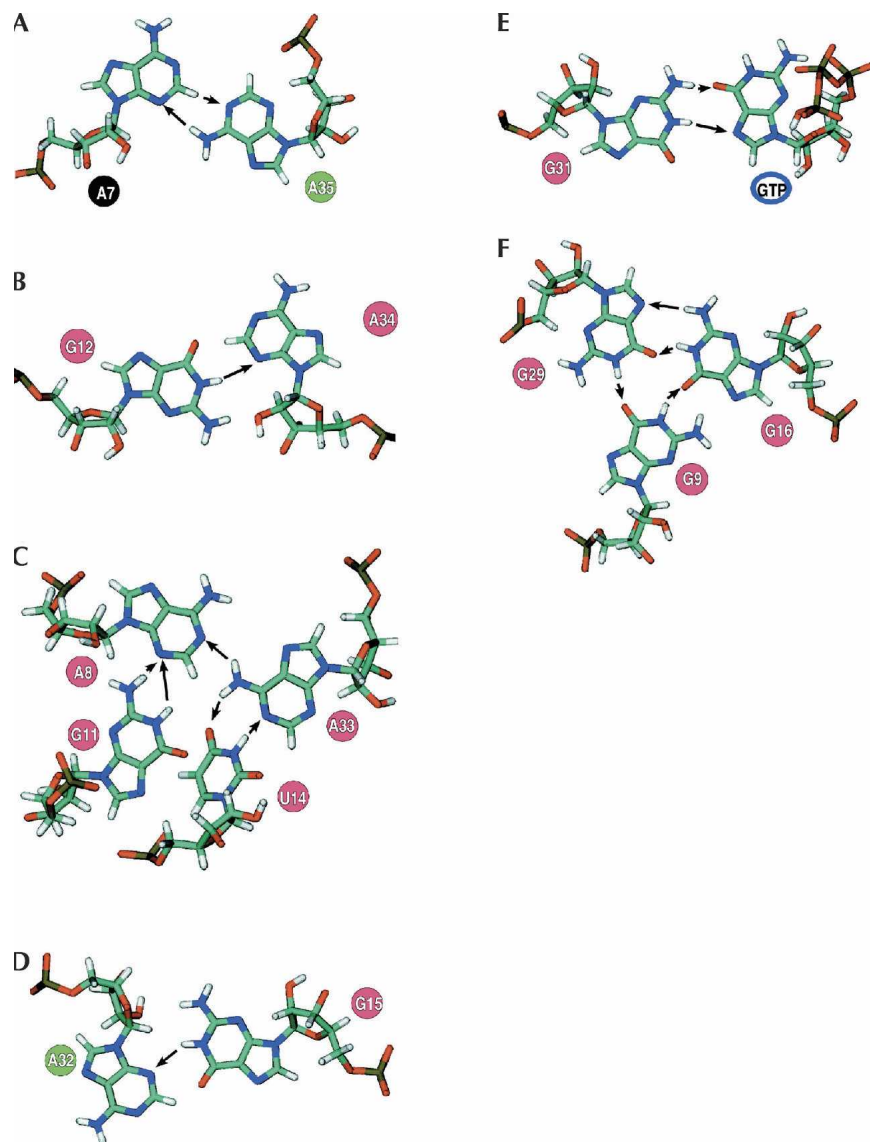Each G in the G9:G16:G29 base triple (Fig. 4F) is involved in N1H-carbonyl hydrogen bonds with the other

FIGURE 4. (*A–F*) Predicted base–base interactions. (*A*) A:A pairing that caps the P1 stem. (*B–F*) Base–base interactions that stabilize the core. Putative hydrogen bonds consistent with the structural and spectroscopic data are marked by arrows from donor to acceptor. The identity of each position is colored for sequence conservation in the same manner as in Figure 1. Note that the predicted base–base interactions are present throughout the family of 10 low energy structures shown in Figure 3A.

two. The presence of well-resolved G16 H2 amino resonances is consistent with a G16 H2–G29 N7 interaction. G29 adopts a X *syn* conformation in order to take part in the G9:G16:G29 platform. This conformation of G29, combined with the need to position the G9 base so it can stack on top of the GTP nucleobase dictates absolute sequence conservation of the residues in this triple.

GTP occupies the center

The GTP ligand sits within a pocket at the very center of the aptamer (Fig. 3). Drawing upon previously collected data

(Carothers et al., in prep.), we know that the Class I aptamer binds both GTP and a number of GTP analogs with high affinity (Fig. 5). The structural data illustrate that the recognition bulge forms specific surfaces that can nonetheless accommodate chemically modified versions of GTP.

Like the 10 other characterized GTP aptamers (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.), the Class I aptamer interacts more specifically with the nucleobase than the ribose or triphosphate portions of the ligand (Fig. 5). In the structure, the GTP ribose sugar nestles underneath the first turn of L1 while the triphosphate protrudes into the solvent (Fig. 6A,B,C). Although we cannot rule out water- or metal ion-mediated contacts, the NMR structure is consistent with the analog binding data in that we cannot discern any specific aptamer-GTP ribose or phosphate interactions. Instead, a wall comprised of the phosphodiester backbone and riboses of A8, G9, U10, and G11 makes intermolecular van der Waals contacts and topologically constrains those parts of ligand.

The GTP nucleobase stacks below G9 and on top of G15 (Fig. 6C,D). The GTP H1 imino proton lies over the G15 base, resulting in a chemical shift (10.77 ppm) that is several ppm upfield of the W-C base-paired iminos (12–13 ppm). NOEs involving GTP H1 are detected to all of the immediately surrounding bases, G9, G15, G16, and G31. The fact that we observe the GTP H1 as a well-resolved resonance at 15°C suggests involvement in a hydrogen bond. The G16 N7–GTP H1 distance is only ~1.8 Å, short enough for a hydrogen bond even though the bases are not coplanar (Hoffmann et al. 2003). Steric clashes (Fig. 6D) and loss of the putative GTP H1–G16 N7 interaction explain why binding to 1-methyl GTP is reduced >100-fold relative to GTP.

The Hoogsteen face of GTP abuts the W-C face of G31 (Fig. 4E). Two potential hydrogen bonds stabilize this G:G N7 imino-type interaction (Nagaswamy et al. 2002); one is between G31 H1 and GTP N7, and the other, between G31 H2 and GTP O6. The analog binding results demonstrate the importance of these GTP Hoogsteen face contacts. The affinity of 7-methyl GTP is reduced more than 1000-fold compared with GTP. No binding was observed when the
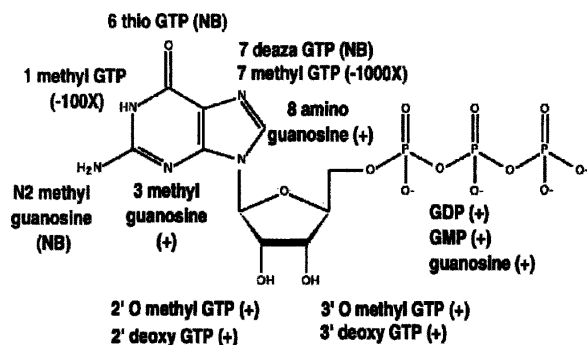
FIGURE 5. Class I GTP analog binding. The analog binding results from J.M. Carothers, S.C. Oestreich, and J.W. Szostak (in prep.) are summarized in this diagram. The numbers in parentheses are the fold reduction in binding affinity. NB signifies no binding detected. Plus sign (+) denotes that the aptamer K_d for the analog was not more than 100-fold worse than the aptamer K_d (76 nM) for GTP.

GTP O6 was changed to sulfur (6 thio GTP) or when GTP N7 was changed to CH (7 deaza GTP).

The NMR data and structure do not indicate crucial roles for the GTP H2 amino protons, N3 (Fig. 6C,D) or 8 position (Fig. 6B). The GTP H2 amino protons are in fast enough exchange with the solvent that they resonate as a single peak (6.35 ppm) at 15°C and do not show any intramolecular GTP H1–H2 NOEs. However, no aptamer binding was observed to a GTP analog methylated at the N2 position. This is probably due to a steric clash with the ribose of G15. No portion of the aptamer is close enough to GTP N3 to interfere with modification at that position. As expected, binding to 3-methyl GTP was not greatly diminished. The 8 position points toward the outside of the aptamer with little obstruction. Not surprisingly, the aptamer has the same affinity for 8 amino-guanosine as for GTP.

DISCUSSION

We have determined the solution NMR structure of the Class I aptamer bound to GTP. The Class I recognition bulge forms a reverse S that shapes a pocket that engulfs the ligand (Figs. 3, 7). We find that three unconserved positions within the recognition bulge do not make specific inter- or intramolecular contacts but are necessary because of their role as backbone pivots and spacers. However, most positions within the core of the aptamer are invariant. Platforms involving base–base interactions among invariant nucleotides stabilize the core and explain the observed sequence conservation and thus the informational complexity of the aptamer. Now that we have generated initial models of the Class I aptamer structure, future NMR work can be directed toward obtaining direct evidence for the predicted base–base hydrogen bonds, examining the roles of the obligatory metal ions and investigating the dynamics of the structure (Latham et al. 2005).

The Class I aptamer inner stem (P2 in Fig. 1) and tetra-loop (L2) correspond to the secondary structure element that was engineered into half of the original random-sequence library (Davis and Szostak 2002). As expected, the engineered stem–tetraloop anchors one end of the recognition bulge (Fig. 3), a condition that is likely repeated in the six other GTP aptamers containing the designed element (Davis and Szostak 2002; Carothers et al. 2004). Stem P1 anchors the other side of the recognition bulge. Both empirical findings (Davis and Szostak 2002; Carothers et al. 2004) and theoretical considerations (Gevertz et al. 2005) argue that designing additional secondary structure elements into sequence libraries may further increase the likelihood of obtaining highly-active molecules.

The G9:G16:G29 base triple forms the ceiling of the ligand binding pocket and G15:A32 constitutes the floor (Figs. 3, 5, 7). Specific interactions between the W-C face of G31 and the Hoogsteen face of GTP contribute to the binding affinity and specificity that we have observed (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.; Fig. 5). The structure shows that many of the positions on the GTP can be altered without introducing substantial steric clashes or destroying potential hydrogen-bonding contacts. Although some RNA aptamers can specifically recognize phosphates (Sazani et al. 2004), neither the analog binding data (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.; Fig. 5) nor the NMR structure (Fig. 6) indicates that the Class I aptamer makes specific contacts with the triphosphate region of the GTP ligand.

The mode of ligand binding for the Class I GTP aptamer ($K_d \approx 75$ nM in 5 mM Mg^{2+}) resembles the way in which other RNA aptamers interact with ligands containing planar heterocyclic groups, such as ATP ($K_d \approx 6000$ nM in 5 mM Mg^{2+}; Sassanfar and Szostak 1993; Dieckmann et al. 1996; Jiang et al. 1996; Fig. 7), theophylline ($K_d \approx 200$ nM; Jenison et al. 1994; Zimmermann et al. 1997), and FMN ($K_d \approx 500$ nM; Burgstaller and Famulok 1994; Fan et al. 1996). In each of these aptamer structures, the planar portion of the ligand stacks between platforms of purine bases. Hydrogen bonds between atoms on the planar moiety of the ligand and a base at the "docking site" dictate discrimination against several chemically related analogs (Hermann and Patel 2000).

For the two aptamers that bind ribonucleotides, there are further similarities in ligand recognition. Both the Sassanfar ATP and Class I GTP aptamers are relatively insensitive to changes in the ligand sugar and exhibit little preference for the triphosphate compared with the nucleoside form of the ligand (Sassanfar and Szostak 1993; J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.). The fraction of the surface area of a ligand buried upon formation of a complex is known to make an important contribution to binding affinity (Brooijmans et al. 2002; Fraternali and Cavallo 2002). In this regard, it is significant that the Class I and Sassanfar aptamers bury nearly identical fractions of calculated ligand solvent accessible surface area. The Class I GTP aptamer buries 94% of the hydrophobic surface area and
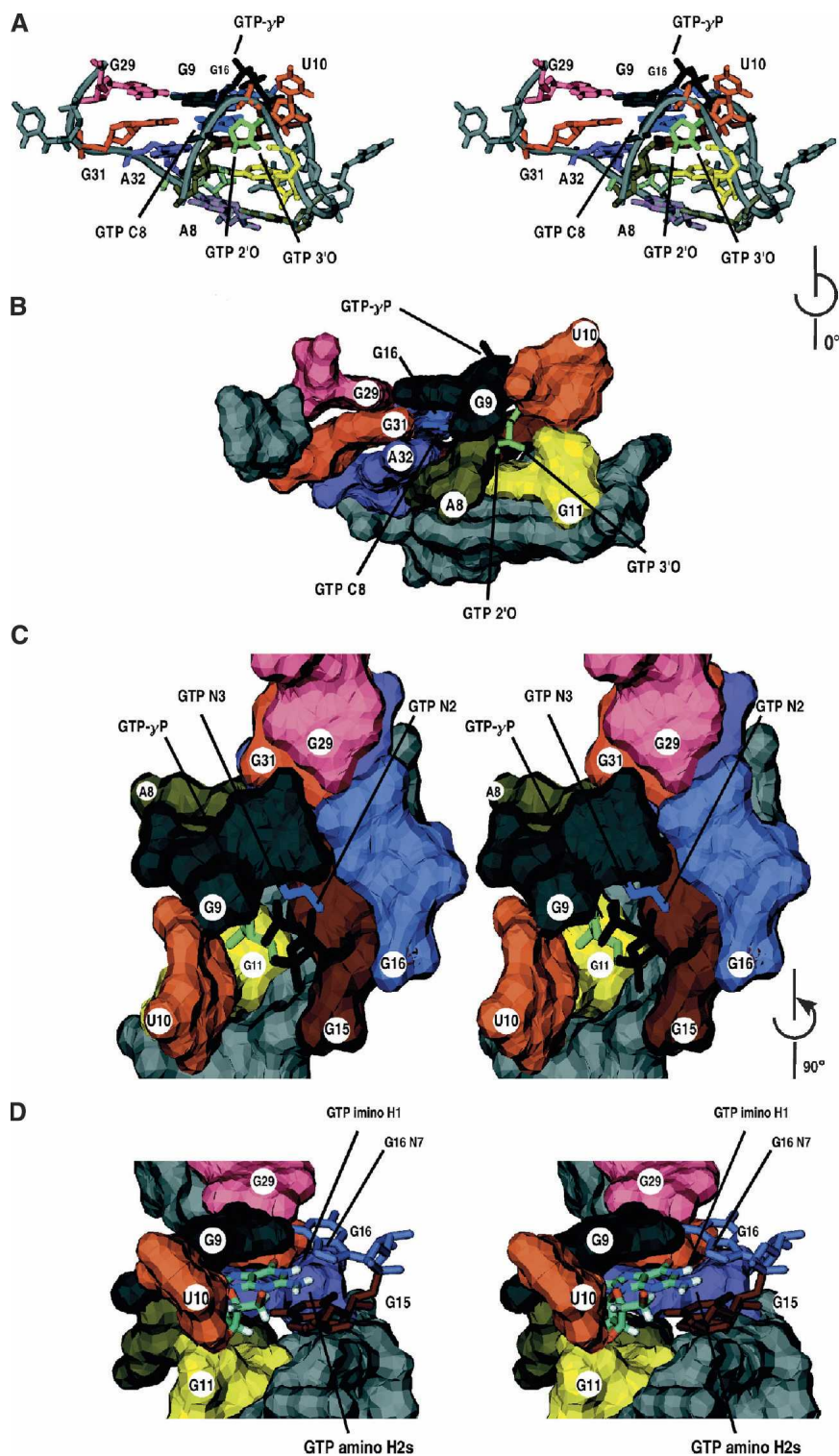
FIGURE 6. Structure of the recognition bulge. (*A*) Stereo view of the recognition bulge in the same orientation as in Figure 3A. (*B*) Molecular surface depiction of the image in *A*. (*C*) Stereo view of the molecular surface from the top of the aptamer, rotated 90° counterclockwise relative to *B*. (*D*) Stereo view with G15 and G16 in stick form. The GTP triphosphate is removed for clarity. The residues are numbered and colored consistently in *A–D*. The GTP ribose is green. The GTP triphosphate is black. The GTP nucleobase is blue, except in *D*, where CPK coloring is used.

81% of the total surface area of the monophosphate form of the GTP ligand (see Materials and Methods). The Sassanfar ATP aptamer buries 94% of the hydrophobic surface area and 85% of the total surface area of AMP.

Taking the biochemical and NMR structural data together, the Class I GTP aptamer does not make a larger number of specific contacts with the ligand than the Sassanfar ATP aptamer. Nor does the Class I aptamer exhibit a mode of interaction with the ligand that is qualitatively different from the Sassanfar aptamer. How, then, does the Class I aptamer achieve a K_d for GTP that is almost 2 orders of magnitude better than that of the Sassanfar aptamer for ATP in similar buffer conditions?

Given the similarities in ligand binding and recognition, we conclude that different stabilities for the overall folds of the Class I and Sassanfar aptamers must themselves be responsible for the differences in K_d. Although both aptamers have the same stem–loop–stem secondary structure organization (Figs. 1, 8), the Class I aptamer has four more invariant bases (11 vs. 7) within its recognition bulge and 10 more bits of total information content ($\sim$45 bits vs. 35 bits). The more informationally complex Class I aptamer has two platforms of stacked base-triples and several other base–base contacts that stabilize the core of the RNA. In contrast, the recognition bulge of the Sassanfar ATP aptamer has no base-triples and only two stabilizing base–base interactions (Dieckmann et al. 1996, 1997; Jiang et al. 1996).

Theophylline and FMN differ in size and structure from the ribonucleotides. For these reasons, direct comparison of the GTP aptamer with the theophylline and FMN aptamers may be less informative than comparison with the ATP aptamer (Brooijmans et al. 2002). Still, it is notable that the theophylline and FMN aptamers, which also have lower K_d's for their ligands than the ATP aptamer, have recognition bulges characterized by large numbers of conserved positions in base-triples and other stabilizing base–base interactions (Fan et al. 1996; Zimmermann et al. 1997; Hermann and Patel 2000).
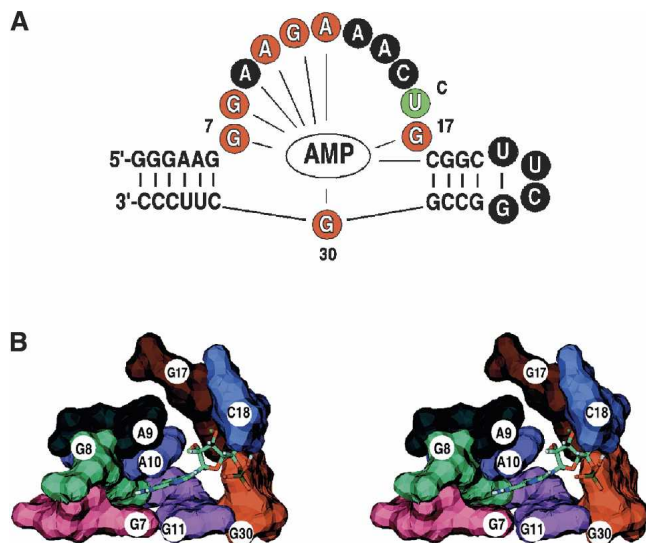
FIGURE 7. Sassanfar ATP aptamer (Sassanfar and Szostak 1993). (*A*) The secondary structure model is shown in the same manner as the Class I GTP aptamer in Figure 1. The recognition bulge nucleotides are colored according to sequence conservation. The lines indicate that intermolecular NOEs to the AMP ligand were detected for those positions (Dieckmann et al. 1996). (*B*) Stereo view of the molecular surface of the Sassanfar ATP aptamer binding pocket (PDB ID 1RAW) with the AMP ligand CPK colored and residues numbered as in *A*.

Greater structural stability could lead to faster on-rates and lower K_d's by reducing the conformational heterogeneity in the prebound state of an aptamer (Szwajkajzer and Carey 1997) or otherwise increasing the propensity of the RNA to adopt a folded, active form (Schultes et al. 2005). If ligand dissociation requires an aptamer to partially unfold, greater structural stability could also produce slower off-rates and thus lower K_d's. Imino proton spectra collected in the absence and the presence of ligand show that neither the Class I nor the Sassanfar aptamers has a fully structured recognition bulge until the ligand binds (Fig. 2 in Dieckmann et al. 1996; Fig. 8). This is consistent with heterogeneous prebound states and induced-fit binding mechanisms (Hermann and Patel 2000).

For the Class I GTP and Sassanfar ATP aptamers the difference in on-rates accounts for much of the variation in K_d's; the Class I aptamer has an on-rate for GTP (2500 m^{-1} s^{-1}; S.C. Oestreich, J.M. Carothers, and J.W. Szostak, unpubl.) that is 45-fold faster than the Sassanfar aptamer on-rate for ATP ($\sim$55 m^{-1} s^{-1}; estimated from Gebhardt et al. 2000) in 5mM Mg^{2+}. The corresponding difference in off-rates is smaller; the Class I aptamer-GTP off-rate is 1.9×10^{-4} s^{-1} ($t_{1/2}$ = 60 min; S.C. Oestreich, J.M. Carothers, and J.W. Szostak, unpubl.) and the Sassanfar aptamer-ATP off-rate is $\sim$3 $\times$ 10^{-4} s^{-1} ($t_{1/2}$ = 32 min; estimated from Gebhardt et al. 2000). Thus, the additional information content in the Class I aptamer specifies a more stable structure that has both a significantly faster on-rate and a moderately slower off-rate for GTP than the Sassanfar aptamer has for ATP. Note that these results do not necessarily

mean that, compared with the Sassanfar ATP aptamer, the prebound conformation of the Class I GTP aptamer is more stable, or that the population of prebound structures is less heterogeneous. Rather, it is also possible that the additional complexity of the Class I aptamer simply enables the RNA to more rapidly become well-ordered upon encountering the ligand. As mentioned above, the same structural elements that result in a greater tendency to adopt the active conformation (increasing the on-rate) could simultaneously reduce the likelihood that the RNA will dissociate from the ligand (decreasing the off-rate).

For the set of 11 different GTP aptamers, 10 $\pm$ 5 more bits of information content are needed to specify RNA molecules which, in a given environment, exhibit 10-fold improvements in binding affinity (Carothers et al. 2004). The same correspondence between information and activity for two ribozymes with RNA–RNA ligation activity led to the hypothesis that common physical phenomena could underlie a general information-RNA function relationship (Carothers et al. 2004).
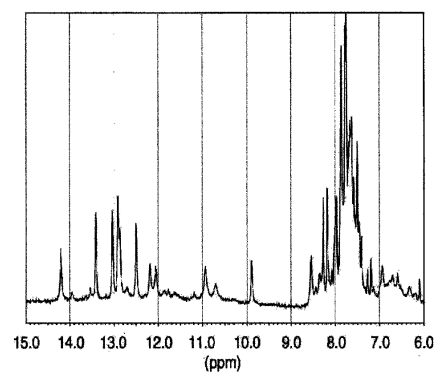
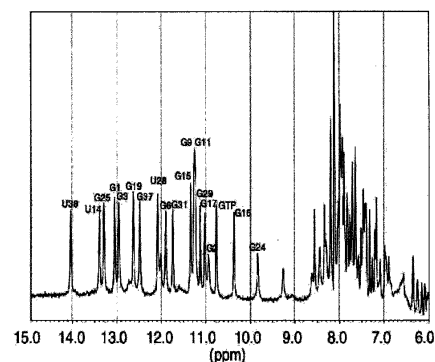**FIGURE 8.** 1D exchangeable and aromatic proton spectra for the Class I GTP aptamer. (*A*) Spectrum collected before GTP was added to the sample. (*B*) Spectrum collected after a 1.1-fold molar excess of GTP was added to the sample. The spectra were collected at 500 MHz in 95% H_2O/5% D_2O at 24°C. Residues with well-resolved imino proton resonances are identified. (See also Fig. 2 of Dieckmann et al. 1996.)

Analog binding studies (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.) demonstrated that the higher-affinity GTP aptamers do not make a larger number of specific contacts with the ligand than GTP aptamers with lower affinities. Those results favor a model where increasing the stability of the functional RNA, through a greater number of intra-aptamer contacts, is the simplest mechanism for improving activity. Comparison of the structural and informational properties of the Class I GTP aptamer and the Sassanfar ATP aptamer provides further evidence in support of this hypothesis. Stabilizing the active conformation of the RNA may also be the easiest way to improve ribozyme activities (Bergman et al. 2000). If so, the informational cost to improve activity could be generalizable because the difficulty of specifying further intra-RNA contacts should not depend on the function of the RNA.

It will be interesting to assemble other examples of highly active in vitro-selected aptamers and ribozymes and determine the physical mechanisms that contribute to the high levels of function. At some level of activity, it may become easier, or necessary, to improve functionality through an increase in the number of contacts with the substrate or ligand. Also, improvements of any kind could become incrementally more difficult to achieve (Szostak 2003). Either circumstance could lead to deviations from the information–function relationship that has been observed (Carothers et al. 2004). On the other hand, even though they are burdened with an evolutionary history, large complex biological RNA molecules have the same structural motifs found in small synthetic RNA molecules (Moore 1999; Hermann and Patel 2000). At least in this sense, large functional RNA structures are not fundamentally different from smaller structures, suggesting that the information–function relationship could be scale-invariant over a wide range. We are hopeful that future studies will reveal the extent to which a quantitative relationship between informational complexity and RNA functional activity can be generalized.

MATERIALS AND METHODS

RNA sample preparation

RNAs were transcribed using T7 RNA polymerase (Milligan and Uhlenbeck 1989). The uniformly labeled $^{13}C/^{15}N$ sample was prepared using nucleotide triphosphates enzymatically synthesized from nucleotide monophosphate precursors (Nikonowicz et al. 1992). A *cis*-acting hammerhead ribozyme was transcribed at the 3′ end of the RNA to generate uniform 3′ ends, for the $^{13}C/^{15}N$ uniformly labeled sample (Grosshans and Cech 1991). The other samples were transcribed in a similar manner without the *cis*-acting hammerhead ribozyme. Although hammerhead cleavage required a U at position 40, other constructs were made with a C at position 40 as shown in Figure 1 (the binding affinity for GTP is the same with U40 or C40). Specifically labeled samples were transcribed using individual $^{13}C/^{15}N$-labeled nucleotides purchased from Cambridge Isotope Laboratories, Inc. RNAs were

purified by denaturing PAGE and exchanged into NMR buffer using Centricon YM-10 spin-filters (Millipore, Inc). For isotropic experiments the NMR buffer was 75mM KCl, 10 mM potassium phosphate, 5 mM MgCl$_2$, 0.1 mM EDTA (pH 6.1). The concentration of KCl in the buffer was reduced from 200 mM in the original selection (Davis and Szostak 2002) to 75 mM in order to obtain narrower spectral line-widths. These lower salt conditions did not significantly raise the K_d of the complex (100 nM vs. 75 nM). Experiments involving exchangeable protons were performed in 95% H$_2$O/5% ^{2}H$_2$O. All other experiments were performed in 99.96% ^{2}H$_2$O (Sigma-Aldrich). Samples were heated to 80°C for 1 min in the presence of a ∼1.1:1 molar excess of GTP ligand followed by slow-cooling to room temperature.

NMR spectroscopy

Data were recorded on Varian Inova 750 MHz, Varian Inova 500 MHz, Varian Unity 400 MHz, and Bruker Avance 500 MHz spectrometers. Some of the experiments were performed using a Bruker Cryoprobe or a Varian Coldprobe. With the few exceptions noted below, RNAPack pulse sequences were employed (Lukavsky and Puglisi 2001).

The ribose resonances of the aptamer were assigned using 3D HCCH-TOCSY (Kay and Xu 1993), 3D HCCH-TOCSY, and 3D HCCH-COSY experiments where the sample was 3 mM uniformly labeled $^{13}C/^{15}N$ aptamer and 3.3 mM unlabeled ligand. Intranucleotide aromatic assignments were made using 2D HCN-H1′ and 2D HCN-H6/H8 experiments. Adenine aromatic H2 protons were assigned with a 2D aromatic-optimized HCCH-TOCSY experiment. Exchangeable resonances for bases in the stems were assigned from 2D NOESY spectra. The resonances of the UUCG tetraloop could be assigned from previously published chemical shifts (Fan et al. 1996; Doreleijers et al. 2003). Other exchangeable protons were assigned from base-optimized 2D HNC-TOCSY-CH spectra.

Sequential nucleotide assignments were made using 3D HCP and 3D ^{13}C-edited NOESY-HSQC data. To complete and confirm the assignments we generated several segmentally labeled samples. Two such $^{13}C/^{15}N$ G samples were comprised of two-stranded versions of the aptamer where one strand consisted of residues 1–20 and the other of residues 25–41. Two other samples were made with $^{13}C/^{15}N$ A, where one of the As in a two-base varying position, either A32 or A35, was substituted with a G.

2D ^{13}C-edited HCCH-COSY and 2D ^{13}C-^{1}H HSQC-CT experiments with 1 mM $^{13}C/^{15}N$-labeled GTP ligand and 0.9 mM unlabeled aptamer were sufficient to assign the ribose resonances of the ligand. The aromatic and exchangeable protons of the ligand were readily assigned with the same sample from 2D ^{13}C-^{1}H HSQC and 2D ^{15}N-^{1}H HSQC spectra.

Aromatic base C-H residual dipolar couplings (RDCs) were extracted by subtracting couplings measured in isotropic conditions from those measured in partially aligned conditions. The C-H couplings were obtained from splittings detected in a ^{1}H-decoupled 2D ^{13}C-^{1}H CT-HSQC spectrum and a 2D ^{13}C-^{1}H CT-TROSY spectrum, both recorded with a ^{1}H frequency of 500 MHz. The C-H splitting is equal to the difference between the ^{1}H-decoupled resonance and the narrow TROSY component, scaled by a factor of two. Isotropic measurements were made with 0.8 mM uniformly labeled $^{13}C/^{15}N$ aptamer and 0.9 mM unlabeled GTP ligand. Liquid

crystal conditions for the partially aligned experiment were obtained by dissolving 0.5 mM uniformly labeled ^{13}C/^{15}N aptamer, 0.55 mM unlabeled ligand and buffer into Pf1 phage (ASLA Biotech, Ltd.) that had been dialyzed into NMR buffer (Hansen et al. 1998); the final Pf1 phage concentration was ~18 mg/mL. The pH of the buffer was increased to 6.8 in these experiments to avoid precipitating the phage (the binding affinity of the aptamer for GTP was not affected by the change in pH).

The 2D and 3D NOESYs, 2D DQF-COSY (see below), 3D HCP, and all experiments involving exchangeable protons were performed at 15°C. The 2D adenine aromatic-optimized HCCH-TOCSY was performed at 30°C. Dipolar couplings were measured in isotropic and liquid-crystal conditions at 30°C. All other experiments were conducted at 25°C.

Spectra were processed using NMRPipe (Delaglio et al. 1995) and visualized using SPARKY 3 (Kneller and Kuntz 1993).

Structural restraints

Interproton distances were obtained from the following experiments: 2D NOESY-WATERGATE (100 msec mixing time), 3D ^{13}C-edited ribose-focused NOESY-HSQC (100 msec), 3D ^{13}C-edited aromatic focused NOESY-HSQC (100 msec), 3D ^{13}C-edited aromatic focused NOESY-HSQC (200 msec), and a 3D ^{15}N-edited NOESY-HSQC (100 msec). Peak volumes were calibrated using the pyrimidine H5-H6 and amino-imino cross peaks as references (Zimmermann et al. 1997). For strong peaks the upper bound was set at 3.3 Å. For medium peaks the upper bound was 4.5 Å. For weak peaks the upper bound was 6 Å. The upper distance for weak peaks involving exchangeable protons was 7 Å.

The pattern of NOEs for stem bases 1–6, 18–20, 25–27, and 36–41 was consistent with A-form geometry. Watson-Crick (W-C) hydrogen-bonding restraints (Zimmermann et al. 1997) were employed at these positions (Kitamura et al. 2002). No hydrogen-bonding restraints were used for any position within the binding bulge or for the noncanonical base pairs that cap the stems (A7:A35 and G17:U28). A loose set of standardized intranucleotide and sequential internucleotide distance restraints were applied to the stem bases (positions 1–6, 18–20, 25–27, and 36–41) as follows: The upper distance restraint for intranucleotide H1′ to H6/H8 was 6.0 Å; for intranucleotide H2′ to H6/H8, the upper restraint was 2.8 Å; for intranucleotide H3′ to H6/H8, the upper restraint was 3.3 Å. For sequential imino protons, the upper distance was 4.2 Å. The upper distance between imino protons and sequential amino protons was 6.0 Å. Only experimentally derived restraints were used for the non-W-C pairs at the end of the outer stem (A7-A35) and inner stem (G17-U28). The pattern of NOEs for the UUCG tetraloop (positions 21–24) was consistent with previously published UUCG tetraloop structures (Allain and Varani 1995; Fan et al. 1996; Doreleijers et al. 2003). A set of 28 distance restraints was employed for the UUCG tetraloop.

DQF-COSY and TOCSY spectra were used to semiquantitatively determine ^{3}J–H1′–H2′ couplings, in order to assign sugar puckers as 2′ or 3′ endo (Allain and Varani 1995). Sugar pucker was loosely constrained using the four torsion angles: nu1, nu2, nu3, and nu4 (Kitamura et al. 2002). Bases that showed large couplings consistent with at least partial 2′ endo sugar pucker were G9, U10, G12, U13, U22, C23, U30, G31, and A35. Their sugar torsion angles were restrained to a range that encompassed 2′ and 3′ sugar puckers to

allow for the possibility that they fluctuate between the two conformations. In cases of ambiguity (e.g., from spectral overlap), the sugar puckers were not restrained. All other bases that had small ^{3}J–H1′–H2′ couplings were restrained to 3′ endo.

G12 and U13 showed very strong ^{3}J–C2′–P couplings and very weak ^{3}J–C4′–P couplings in a 3D HCP spectrum. Because both G12 and U13 also have 2′ endo sugar puckers, ε was restrained to *gauche-* for these two positions (Allain and Varani 1995).

G24, G29, and A35 had unusually strong intranucleotide H1′ to H8 NOEs. Accordingly, χ was restrained to *syn* for these bases (Allain and Varani 1995; Zimmermann et al. 1997). All other positions in the aptamer were restrained to χ *anti* except in those cases where χ could not be unambiguously determined. Preliminary structure calculations showed the GTP ligand in a *syn* conformation even though its χ angle had not been restrained. Consistent with a χ *syn* glycosidic angle, we detected a strong GTP H1′–H8 intranucleotide NOE. Guanosine nucleotides are known to preferentially adopt the *syn* conformation (Saenger 1984). In addition, the aptamer does not bind inosine (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.), a chemical analog of GTP with a χ *anti* conformation (Saenger 1984). Accordingly, the GTP ligand was restrained to χ *syn* in the final structure calculations.

Structure calculations

Structure calculations were performed with XPLOR-NIH version 2.9.2 (Schwieters et al. 2003). The restraints used in the calculations are summarized in Table 1. We employed a protocol derived from the one described by Kitamura et al. (2002). Briefly, 100 starting structures with random torsion angles were subjected to high temperature restrained molecular dynamics with NOE and base-pairing distance constraints (70 psec at 1000 K). This was followed by a second stage of high-temperature rMD using the Shake protocol (15 psec at 1000 K). The structures were then cooled to 300 K and subjected to energy minimization; dihedral angle constraints were introduced and their scale gradually increased. Eleven soft (20 kcal mol^{-1}) base-pair planarity restraints were introduced during 20 psec of further rMD at 300 K, followed by energy minimization. In the final stage the Lennard-Jones potential was turned on and the structures were energy minimized once again. The force constants used in the structure calculations were bond length = 1000 kcal mol^{-1} Å^{-2}, bond angle = 500 kcal mol^{-1} rad^{-2}, NOE = 50 kcal mol^{-1} Å^{-2}, dihedral angle = initial 5 kcal mol^{-1} rad^{-2} to final 400 kcal mol^{-1} rad^{-2}. The 20 structures with the lowest energies and no NOE violations >0.2 Å and no dihedral angle violations >5° were selected for further refinement with a set of aromatic C-H residual dipolar couplings.

Aromatic C-H residual dipolar couplings (RDCs) corresponding only to well-restrained nucleotides (seven couplings in stem P1, seven couplings in stem P2, and 11 couplings in the recognition bulge) were incorporated during 20 psec of rMD at 300 K as susceptibility anisotropy restraints (SANI), followed by energy minimization. The SANI force constant was increased from 0.005 kcal mol^{-1} to 1.000 kcal mol^{-1} in 0.005 kcal mol^{-1} steps. Values for the magnitude of the axial component (A_a) and rhombicity (R) of the alignment tensor were estimated from grid searches (Tjandra and Bax 1997) where A_a and R were arrayed. The values for A_a and R obtained from grid searches were comparable with those we computed (using MODULE 1.0) from least-squares minimization

of the target function with coordinates for the minimized average of the 10 lowest energy structures (Dosset et al. 2001) before refinement with the RDCs. The optimal values for the alignment tensor, scaled to C-H, were $A_a(\times 10^{-4}) = 23 \pm 1$ and $R = 0.17 \pm 0.03$. Initially we had treated the stems separately from one another to allow for the possibility of independent motion for each region (Leeper and Varani 2005). However, we found that the differences in the optimal A_a/R values for each region were not statistically significant. Before refinement with the RDCs the rmsd of the 10 lowest energy structures was 3.85 Å. After refinement with the RDCs the rmsd of the 10 lowest energy structures was 3.38 Å (see Table 1). None of the final 10 structures have any NOE violations >0.2 Å, dihedral angle violations >5°, or RDC violations >5 Hz.

Binding assays

Apparent dissociation constants (K_d's) were determined using the spin-filtration method (Jenison et al. 1994) as previously described (Davis and Szostak 2002) with trace levels of ^{32}P-GTP and titration of the RNA concentration over the range where 50% of the GTP was bound. The off-rate was measured in a similar manner except that excess unlabeled GTP was added after the RNA and ^{32}P-GTP were equilibrated to prevent the aptamer from rebinding dissociated ^{32}P-GTP. Aliquots were removed and processed by spin filtration at regular intervals to measure the fraction of the initially bound ^{32}P-GTP that had dissociated as a function of time (S.C. Oestreich, J.M. Carothers, and J.W. Szostak, unpubl.).

Solvent accessible surface area

We used the Parameter Optimized Surfaces Web server (Fraternali and Cavallo 2002; http://ibivu.cs.vu.nl/programs/popswww/) to compute the solvent accessible surface area of the GMP portion of the GTP ligand in the presence and the absence of the Class I aptamer. We focused on the GMP moiety because the triphosphate of GTP is known not to be important for high-affinity binding to the Class I aptamer (J.M. Carothers, S.C. Oestreich, and J.W. Szostak, in prep.) and to make more direct comparisons with the Sassanfar ATP aptamer (Sassanfar and Szostak 1993) whose NMR structure (PDB ID 1RAW) was previously solved in the presence of AMP (Dieckmann et al. 1996).

Coordinates and restraints

Structure coordinates for the averaged structure have been deposited in the RCSB Protein Data Bank (PDB ID 2AU4). Proton chemical shifts and experimental distance restraints will be available in the BioMagResBank upon publication.

ACKNOWLEDGMENTS

REFERENCES

Adami, C. 2004. Information theory in molecular biology. *Physics of Life Reviews* **1:** 3–22.

Adami, C. and Cerf, N.J. 2000. Physical complexity of symbolic sequences. *Physica D.* **37:** 62–69.

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

Bergman, N.H., Johnston, W.K., and Bartel, D.P. 2000. Kinetic framework for ligation by an efficient RNA ligase ribozyme. *Biochemistry* **39:** 3115–3123.

Bondensgaard, K., Mollova, E.T., and Pardi, A. 2002. The global conformation of the hammerhead ribozyme determined using residual dipolar couplings. *Biochemistry* **41:** 11532–11542.

Brooijmans, N., Sharp, K.A., and Kuntz, I.D. 2002. Stability of macromolecular complexes. *Proteins* **48:** 645–653.

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

Carothers, J.M. and Szostak, J.W. 2006. In vitro selection of functional oligonucleotides and the origins of biochemical activity. In *Aptamers and synthetic catalytically active oligonucleotides: Identification and applications* (ed. S. Klussmann), pp. 3–28. Wiley-VCH, Berlin, Germany.

Carothers, J.M., Oestreich, S.C., Davis, J.H., and Szostak, J.W. 2004. Informational complexity and functional activity of RNA structures. *J. Am. Chem. Soc.* **126:** 5130–5137.

Chen, X., McDowell, J.A., Kierzek, R., Krugh, T.R., and Turner, D.H. 2000. Nuclear magnetic resonance spectroscopy and molecular modeling reveal that different hydrogen bonding patterns are possible for G·U pairs: One hydrogen bond for each G·U pair in r(GGCGUGCC)(2) and two for each G·U pair in r(GAGUGCUC)(2). *Biochemistry* **39:** 8970–8982.

Davis, J.H. and Szostak, J.W. 2002. Isolation of high-affinity GTP aptamers from partially structured RNA libraries. *Proc. Nat. Acad. Sci.* **99:** 11616–11621.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293.

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

Dieckmann, T., Butcher, S.E., Sassanfar, M., Szostak, J.W., and Feigon, J. 1997. Mutant ATP-binding RNA aptamers reveal the structural basis for ligand binding. *J. Mol. Biol.* **273:** 467–478.

Doreleijers, J.F., Mading, S., Maziuk, D., Sojourner, K., Yin, L., Zhu, J., Markley, J.L., and Ulrich, E.L. 2003. BioMagResBank database with sets of experimental NMR constraints corresponding to the structures of over 1400 biomolecules deposited in the Protein Data Bank. *J. Biomol. NMR* **26:** 139–146.

Dosset, P., Hus, J.C., Marion, D., and Blackledge, M. 2001. A novel interactive tool for rigid-body modeling of multi-domain macromolecules using residual dipolar couplings. *J. Biomol. NMR* **20:** 223–231.

Eaton, B.E., Gold, L., and Zichi, D.A. 1995. Let's get specific: The relationship between specificity and affinity. *Chem. Biol.* **2:** 633–638.

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

Feig, A.L. and Uhlenbeck, O.C. 1999. The role of metal ions in RNA biochemistry. In *The RNA world*, 2d ed. (eds. R.F. Gesteland et al.), pp. 287–319. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Fraternali, F. and Cavallo, L. 2002. Parameter optimized surfaces (POPS): Analysis of key interactions and conformational changes in the ribosome. *Nucleic Acids Res.* **30:** 2950–2960.

Gebhardt, K., Shokraei, A., Babaie, E., and Lindqvist, B.H. 2000. RNA aptamers to S-adenosylhomocysteine: Kinetic properties, divalent cation dependency, and comparison with anti-S-adenosylhomo-cysteine antibody. *Biochemistry* **39:** 7255–7265.

Gevertz, J., Gan, H.H., and Schlick, T. 2005. In vitro RNA random pools are not structurally diverse: A computational analysis. *RNA* **11:** 853–863.

Grosshans, C.A. and Cech, T.R. 1991. A hammerhead ribozyme allows synthesis of a new form of the Tetrahymena ribozyme homogeneous in length with a 3′ end blocked for transesterification. *Nucleic Acids Res.* **19:** 3875–3880.

Hansen, M.R., Mueller, L., and Pardi, A. 1998. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5:** 1065–1074.

Hermann, T. and Patel, D.J. 2000. Adaptive recognition by nucleic acid aptamers. *Science* **287:** 820–825.

Hoffmann, B., Mitchell, G.T., Gendron, P., Major, F., Andersen, A.A., Collins, R.A., and Legault, P. 2003. NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. *Proc. Natl. Acad. Sci.* **100:** 7003–7008.

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

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

Kay, L.E. and Xu, G.A. 1993. Gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H2O samples of proteins. *J. Mag. Reson. B* **101:** 333–337.

Kitamura, A., Muto, Y., Watanabe, S., Kim, I., Ito, T., Nishiya, Y., Sakamoto, K., Ohtsuki, T., Kawai, G., Watanabe, K., et al. 2002. Solution structure of an RNA fragment with the P7/P9.0 region and the 3%-terminal guanosine of the tetrahymena group I intron. *RNA* **8:** 440–451.

Kneller, D.G. and Kuntz, I.D. 1993. UCSF Sparky: An NMR display, annotation and assignment tool. *J. Cell. Biochem.* **53:** 254.

Latham, M.P., Brown, D.J., McCallum, S.A., and Pardi, A. 2005. NMR methods for studying the structure and dynamics of RNA. *Chembiochem* **6:** 1492–1505.

Lau, M.W.L., Cadieux, K.E.C., and Unrau, P.J. 2004. Isolation of fast purine nucleotide synthase ribozymes. *J. Am. Chem. Soc.* **126:** 15686–15693.

Leeper, T.C. and Varani, G. 2005. The structure of an enzyme-activating fragment of human telomerase RNA. *RNA* **11:** 394–403.

Lorsch, J.R. and Szostak, J.W. 1996. Chance and necessity in the selection of nucleic acid catalysts. *Acc. Chem. Res.* **29:** 103–110.

Lukavsky, P.J. and Puglisi, J.D. 2001. RNAPack: An integrated NMR approach to RNA structure determination. *Methods* **25:** 316–332.

Milligan, J.F. and Uhlenbeck, O.C. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* **180:** 51–62.

Moore, P.B. 1999. Structural motifs in RNA. *Annu. Rev. Biochem.* **68:** 287–300.

Nagaswamy, U., Larios-Sanz, M., Hury, J., Collins, S., Zhang, Z., Zhao, Q., and Fox, G.E. 2002. NCIR: A database of non-canonical interactions in known RNA structures. *Nucleic Acids Res.* **30:** 395–397.

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

Rimmele, M. 2003. Nucleic acid aptamers as tools and drugs: Recent developments. *Chembiochem* **4:** 963–971.

Saenger, W. 1984. *Principles of nucleic acid structure.* Springer-Verlag, New York.

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

Schneider, T.D., Stormo, G.D., Gold, L., and Ehrenfeucht, A. 1986. Information content of binding sites on nucleotide sequences. *J. Mol. Biol.* **188:** 415–431.

Schultes, E.A., Spasic, A., Mohanty, U., and Bartel, D.P. 2005. Compact and ordered collapse of randomly generated RNA sequences. *Nat. Struct. Mol. Biol.* **12:** 1130–1136.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. 2003. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160:** 65–73.

Sullenger, B.A. and Gilboa, E. 2002. Emerging clinical applications of RNA. *Nature* **418:** 252–258.

Szép, S., Wang, J., and Moore, P.B. 2003. The crystal structure of a 26-nucleotide RNA containing a hook-turn. *RNA* **9:** 44–51.

Szostak, J.W. 2003. Functional information: Molecular messages. *Nature* **423:** 689.

Szwajkajzer, D. and Carey, J. 1997. Molecular and biological constraints on ligand-binding affinity and specificity. *Biopolymers* **44:** 181–198.

Tjandra, N. and Bax, A. 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278:** 1111–1114.

Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.P., and Pardi, A. 1997. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat. Struct. Biol.* **4:** 644–649.