A Semiconserved Residue Inhibits Complex Formation by Stabilizing Interactions in the Free State of a Theophylline-Binding RNA†

Grant R. Zimmermann,‡ Thomas P. Shields,‡ Robert D. Jenison,§ Catherine L. Wick,‡ and Arthur Pardi*,‡

Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, Colorado 80309-0215, and NeXstar Pharmaceuticals Inc., 2860 Wilderness Place, Boulder, Colorado 80301

Received January 12, 1998; Revised Manuscript Received April 10, 1998

ABSTRACT: The theophylline-binding RNA aptamer contains a 15 nucleotide motif that is required for high-affinity ligand binding. One residue within this RNA motif is only semiconserved and can be an A or C. This residue, C27, was disordered in the previously determined three-dimensional structure of the complex, suggesting that it is dynamic in solution. ^{13}C Relaxation measurements are reported here, demonstrating that C27 is highly dynamic in the otherwise well-ordered RNA−theophylline complex. A synthetic complex with an abasic residue at position 27 was found to exhibit wild-type binding affinity ($K_d \sim 0.2$ μM), indicating that the base of residue 27 is not directly involved with theophylline binding. Surprisingly, the U27 and G27 RNAs were found to bind theophylline with low affinity (K_d values > 4 μM). NMR spectroscopy on the U27 RNA revealed the presence of an A7-U27 base pair in the free RNA that prevents formation of a critical base-platform structural motif and therefore blocks theophylline binding. Similarly, a protonated A7H$^+$-C27 base pair forms in the absence of theophylline at low pH, which explains the unusual pH dependence of theophylline binding of the C27 RNA aptamer. Thus the weak binding for various nucleotides at position 27 arises not from unfavorable interactions in the RNA−theophylline complex but instead from stable interactions in the free state of the RNA that inhibit theophylline binding.

A major challenge in interpreting the results of in vitro selection experiments for nucleic acids is trying to understand how sequence variations affect the function or structure of the molecule. For selections that have extensive sequence variation, the comparative analysis approach has proven extremely robust in determining the secondary structure of RNAs (*1, 2*). Thermodynamic folding algorithms can also be used to model secondary structures of RNA, but these programs do not generally identify a unique fold for the RNA (*3*). Interactive modeling and computational methods have been used to generate three-dimensional models of RNAs based on their primary structure, chemical probing, and phylogenetic data (*4*−*7*). In the past one had to rely primarily on these theoretical approaches to model the three-dimensional structure of RNA, but recently there has been an explosion in the number of RNA structures determined by X-ray or NMR techniques (*8*). This growth in the structural database for RNA not only has increased the number of known RNA folding motifs but also has improved our understanding of how these motifs assemble into biologically active global structures. However, it is important to realize that the three-dimensional structure alone does not necessarily lead to a clear understanding of the function of the molecule. Thus any three-dimensional model or structure is most valuable when combined with mutational, biochemical, and thermodynamic experiments that probe the activity of the RNA.

The theophylline-binding RNA aptamer (*9*) was identified by the SELEX[1] procedure (*10, 11*) for in vitro selection of functional RNAs. The results of this selection were unusual because only a single family of sequences was observed. This is very different from most selections for protein binding, where a variety of high-affinity RNA motifs are identified (*2, 12*−*15*). Furthermore, only one nucleotide in the consensus sequence of the theophylline-binding RNA showed any sequence variation, with this residue being only an A or C (*9*). The secondary structure of an RNA hairpin containing the consensus theophylline-binding motif (boxed region) is shown in Figure 1A. The three-dimensional structure of this RNA−theophylline complex was recently determined by solution NMR spectroscopy (*16*). An important feature of this RNA motif its ability to bind theophylline with high affinity (K_d of $\sim$0.2−0.8 μM), while simultaneously discriminating against the highly similar molecule caffeine (*9*). The K_d of this RNA aptamer for caffeine is $\sim$4 mM, which is a remarkable level of

† This work was supported by NIH Grant AI30726 and a NIH Research Career Development Award (AI01051) to A.P. and an NRSA Fellowship GM17074 to T.P.S. We also thank the Colorado RNA Center for their generous support of RNA research on the Boulder campus.

* Correspondence should be addressed to this author.
‡ University of Colorado at Boulder.
§ NeXstar Pharmaceuticals Inc.

[1] Abbreviations: CPMG, Carr−Purcell−Meiboom−Gill; HSQC; heteronuclear single-quantum correlation spectroscopy; NMR, nuclear magnetic resonance; NTP, nucleoside triphosphate; PAGE, polyacrylamide gel electrophoresis; rRNA, ribosomal RNA; SELEX, selective evolution of ligands by exponential enrichment; ssDNA, single-stranded DNA; TBDMS, *tert*-butyldimethylsilyl; TEA·HF/NMP, anhydrous triethylamine/hydrogen fluoride in *N*-methylpyrrolidinone.

S0006-2960(98)00082-8 CCC: $15.00
Published on Web 05/30/1998

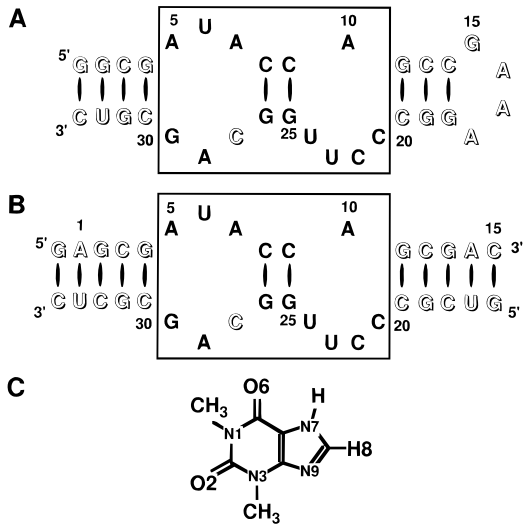

FIGURE 1: Sequence and secondary structure of the theophylline-binding RNAs are illustrated for the (A) hairpin and (B) duplex constructs. The boxed region encompasses the theophylline-binding motif identified in the in vitro selection (*9*). The residues in boldface type are absolutely required for theophylline binding and the residues in outlined type are not conserved or, in the case of position 27, are semiconserved. (C) The covalent structure of theophylline is shown.

discrimination because caffeine differs from theophylline only by the addition of a single methyl group at the N7 position. This level of specificity is 10-fold better than that exhibited by a monoclonal antibody used for the same purpose (*17*).

The only nonconserved residue in the theophylline-binding motif is residue 27 that can be an A or a C. Equilibrium constants measured for theophylline binding to a set of RNAs containing various nucleotides at position 27 demonstrated that this position was important for the ligand binding activity of this aptamer. The previously determined three-dimensional structures of the RNA−theophylline complex revealed a high degree of disorder for nucleotide C27 (*9*). These results suggested that this residue in the core of the theophylline binding pocket was dynamic in solution. To directly characterize the dynamic properties of the RNA, ^{13}C NMR relaxation experiments were performed here on isotopically labeled RNA−theophylline complexes containing A or C residues at position 27. Surprisingly, it was also found that whereas the ligand binding of the A27 RNA was relatively insensitive to pH, the binding of the C27 RNA was adversely affected at pH < 6.5. These thermodynamic and NMR studies are used in combination with the three-dimensional structure to define the structural and functional basis for the observed sequence variation in this theophylline-binding RNA aptamer.

EXPERIMENTAL PROCEDURES

RNA Preparation. All theophylline-binding RNA samples used in the NMR spectroscopy experiments were generated by in vitro transcription with T7 RNA polymerase (*18, 19*). Single-stranded DNA templates containing a T7 promoter and the desired theophylline-binding RNA sequence were generated by solid-phase synthesis. A synthetic ssDNA (18 nucleotides) complementary to the T7 promoter region of the template was added to create a double-stranded promoter, and transcription was performed with either $^{13}C/^{15}N$-labeled or unlabeled NTPs (*20*). The $^{13}C/^{15}N$-labeled NTPs were prepared from rRNA extracted from *Escherichia coli* grown on minimal medium containing 99% enriched ^{13}C-glucose and ^{15}N-ammonium sulfate as described previously (*19*). The RNA transcripts were resolved by denaturing PAGE, and full-length RNAs were eluted from gel slices by crushing and soaking. The heterogeneous length RNA transcripts were processed in trans by a hammerhead ribozyme to remove the 3′-terminal sequence (5′-GCAGGUCX$_n$-3′). The homogeneous length RNA shown in Figure 1A was then separated from the hammerhead ribozyme and the short 3′-terminal sequence by a second denaturing PAGE. The theophylline-binding RNAs were further purified by DEAE-Sephacel anion-exchange chromatography to remove contaminants associated with the PAGE purification. RNAs were then dialyzed extensively in the appropriate NMR buffer using a Centricon-3 concentrator (Amicon); either 20 mM sodium phosphate (pH = 6.8), 30 mM sodium chloride, and 2.0 mM magnesium chloride or 20 mM sodium succinate-d_4 (pH = 5.5), 30 mM sodium chloride, and 2.0 mM magnesium chloride. Theophylline (Figure 1C) was added to form a 1:1 complex as required.

Samples used in the determination of equilibrium binding constants for mutant theophylline-binding RNAs were synthesized by in vitro transcription as described above or were generated by solid-phase synthesis as indicated. Synthetic RNAs were prepared in two separate pieces that were then annealed to reconstitute the theophylline-binding core motif shown in Figure 1B. Mutant RNAs were prepared with ribo-C, deoxy-C, or deoxy-abasic (dSpacer, Glen Research Inc.) nucleotides at position 27. RNAs were generated (0.2 μmol scale) using 2′-*O*-TBDMS-protected phosphoramidites (Glen Research) on an ABI oligonucleotide synthesizer (*21*). Each RNA was removed from the solid support and the bases were deprotected (N4 of cytidines were acetyl-protected) by a 10 min incubation at 65 °C in 40% methylamine. RNAs were dried and resuspended in 250 μL of an anhydrous TEA·HF/NMP solution as described (*21*) and incubated at 65 °C for 1.5 h to remove the 2′-*O*-TBDMS protecting group. Fully deprotected synthetic RNAs were desalted on a size-exclusion NAP25 column (Pharmacia) and then purified by anion-exchange chromatography on a Waters HPLC system with a Dionex NucleoPac PA-100 column maintained at 88 °C. Purified RNAs were desalted in a Centricon-3 concentrator and then dialyzed into binding buffer (see below).

NMR Spectroscopy. All NMR data were collected on a 500 MHz Varian Inova spectrometer equipped with *z*-axis, pulsed-field gradients. One-dimensional ^1H spectra were collected in 90% H_2O/10% ^2H_2O at 15 °C using an 11-echo pulse sequence to suppress the solvent resonance (*22*). A total of 2048 complex points were collected with a spectral width of 12 000 Hz and with the excitation maximum at 12.8 ppm to optimally excite the imino proton region of the spectrum. Two-dimensional (^{13}C,^1H) HSQC−WATERGATE (*23*−*25*) spectra were acquired in 90% H_2O at 5 °C at several pH values. A total of 512 complex points were collected in the ^1H dimension (t_2) with a 6000 Hz spectral width, and 256 complex points were collected in the ^{13}C dimension (t_1) with a spectral width of 7500 Hz, resulting in a 7 h acquisition time. The aromatic ^1H,^{13}C cross-peaks were folded in the indirect dimension by setting the carbon

carrier to 86 ppm during the incremented delay and use of a half-dwell for the initial t_1 value (*26*). The carbon carrier was set to 114 ppm at all other times during the pulse sequence. The time domain data were zero-filled to 1024 and 512 complex points for proton and carbon, respectively. Carbon $T_{1\rho}$ values were measured with a 2D (^{13}C,^{1}H) nonconstant time version of a previously described pulse sequence (*27*). Separate 2D spectra were collected for 0, 4, 8, 16, 20, 32, 40, 48, 60, and 80 ms ^{13}C spin-lock times with a 2.1 kHz spin-lock field. The 8 and 48 ms spin-lock time points were collected twice to estimate the error in signal intensity for individual measurements. A total of 512 complex ^{1}H points were collected with a spectral width of 6000 Hz and 128 complex points were collected over 5000 Hz in the ^{13}C dimension. The carbon carrier was set to 95 ppm, which is between the C1′ and C5 regions of the ^{13}C spectrum, or at 145 ppm for the $T_{1\rho}$ measurements for the C8 resonances in the A27 RNA−theophylline complex. The time domain data for each 2D experiment were zero-filled to 1024 and 256 points in the ^{1}H and ^{13}C dimensions, respectively. The ^{13}C,^{1}H cross-peaks were manually picked in the 0 ms spin-lock time spectrum and the same "footprints" were used to integrate the cross-peaks in all 2D spectra. Volumes were normalized to the peak volume in the 0 ms spin-lock time spectrum and the $T_{1\rho}$ times were determined by fitting to an exponential decay. All of the NMR data were processed using the Felix 95.0 software package (Biosym Technologies).

Equilibrium Binding Experiments. The equilibrium binding constants were measured with ^{14}C- or ^{3}H-labeled theophylline by the equilibrium filtration method (*9*). Either 0.2 μM ^{14}C-labeled theophylline or 0.02 μM ^{3}H-labeled theophylline was combined with various concentrations of RNA (0.05−10 μM), heated to 65 °C for 5 min, and then allowed to cool slowly to 25 °C in binding buffer: 100 mM HEPES (pH 7.3), 50 mM NaCl, and 5.0 mM $MgCl_2$. The annealed complex (150 μL) was added to a Microcon-10 (Amicon) concentrator and spun at 13000g for 4 min. The fraction of theophylline bound was calculated at each RNA concentration by scintillation counting of a 25 μL aliquot from each side of the concentrator membrane. Data were fit to a quadratic binding equation for a 1:1 complex as previously described (*28*). The pH dependences of theophylline binding to the C27 and A27 RNA hairpins were measured with 0.2 μM ^{14}C-theophylline at 0.5 μM RNA from pH 5.0 to 8.0 in 100 mM MES, 50 mM NaCl, and 5.0 mM $MgCl_2$.

RESULTS

K_d Measurements for Theophylline−RNA Complexes. The in vitro selection for theophylline-binding RNAs identified a conserved core motif (Figure 1), but position 27 within this core was found to vary between A and C residues (*9*). The equilibrium constants for the C27, A27, G27, and U27 RNA hairpin constructs binding to theophylline were measured by equilibrium filtration and results are reported in Table 1. Binding constants were also measured for the RNA duplex constructs (Figure 1B) containing a C, deoxy-C, or deoxy-abasic nucleotide at position 27 as illustrated in Figure 2 and these K_d values are also given in Table 1. Surprisingly, the ribo-C, deoxy-C, and deoxy-abasic position 27 mutants all bind theophylline with equivalent affinities, within the error of the measurements.

Table 1: Binding Affinities of Position 27 Mutants of the Theophylline-Binding RNA

nucleotide at position 27	K_d[a] (μM)
Hairpin[b]	
C	0.4
A	0.8
G	> 4
U	> 20
Duplex[c]	
ribo-C	0.2
deoxy-C	0.2
deoxy-abasic	0.1

[a] Dissociation constants were measured at 25 °C in 50 mM HEPES, pH 7.3, 100 mM NaCl, and 5.0 mM $MgCl_2$ by the equilibrium filtration method (*9*). Results are the average of three independent determinations except for the G27 and U27 RNA hairpins, which are single point determinations due to the lower affinity of these RNAs. [b] The K_d values for the hairpin were measured with ^{14}C-labeled theophylline (0.2 μM) and various RNA concentrations (0.05−10 μM). [c] The K_d values for the duplex were measured with ^{3}H-labeled theophylline (0.02 μM) and various RNA concentrations (0.05−10 μM).

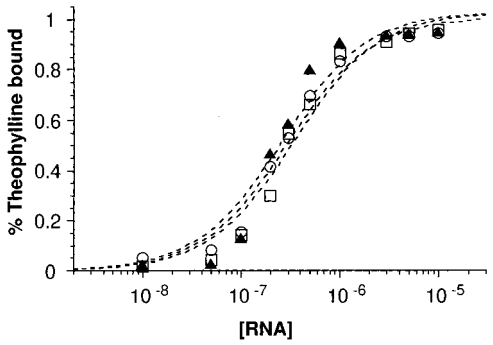

FIGURE 2: Binding curves for the position 27 mutants of the theophylline-binding RNA employing the duplex construct in Figure 1B. The fraction of theophylline bound as a function of RNA concentration at 25 °C, pH = 7.5, is plotted for the ribo-C (○), deoxy-C (□), and deoxy-abasic (▲) RNAs.

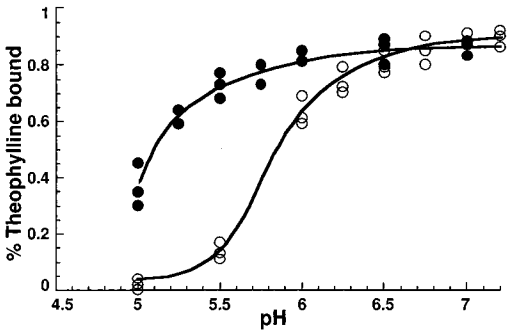

FIGURE 3: Fraction of theophylline bound as a function of pH, plotted for A27 (●) and C27 (○) theophylline-binding RNAs. The fraction of theophylline bound was determined at 25 °C by the equilibrium filtration method as described under Experimental Procedures. These data show the unusual pH dependence of theophylline binding for the C27 RNA (see text).

pH Dependence of Theophylline Binding. The C27 and A27 RNAs bind theophylline at pH = 7.3 with the same K_d values (0.4−0.8 μM) within the error of the measurements (Table 1). However, the C27 and A27 theophylline-binding RNAs exhibit very different pH dependences for ligand binding (Figure 3). Theophylline binding by the A27 RNA is relatively unaffected until the pH drops below ∼5.5, whereas binding to the C27 RNA is inhibited below pH ∼6.5.

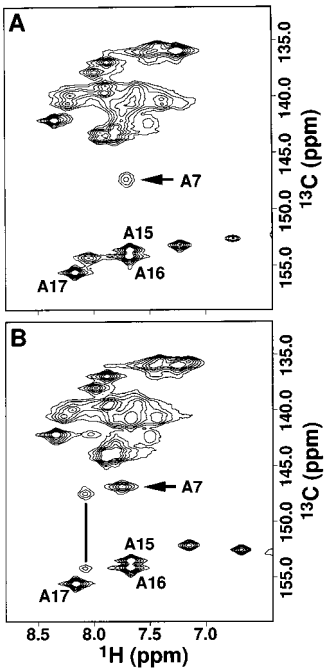

FIGURE 4: Aromatic region of the 2D (^{13}C,^{1}H) HSQC spectra of the C27 theophylline-binding RNA in the absence of ligand at 5 °C at two different pH values. (A) At pH = 6.8, the A7 C2 resonance is shifted upfield from the other C2 resonances and is in intermediate exchange, indicating that this adenine is at least partially protonated at this pH (*29*). (B) At pH = 5.5, the C2 resonance for A7 is a single peak, indicating that this adenine is now fully protonated at this pH. In addition, a second adenine is now in slow exchange between protonated and unprotonated forms at this pH (vertical line). The C2 resonances of the adenine residues in the GAAA tetraloop (A15, A16, and A17) were not protonated under these conditions and were easily assigned by comparison with the spectra of the RNA−theophylline complex (*16*).

Two-dimensional (^{13}C,^{1}H) HSQC spectra of the C27 theophylline-binding RNA were collected in the absence of ligand at pH 6.8 and 5.5 and the aromatic regions of these spectra are compared in Figure 4. Several adenine C2 resonances are shifted significantly upfield at low pH, which is diagnostic of protonated A residues in the core of the theophylline-binding RNA (*29*). Three of the seven adenine residues (A15, A16, and A17) in the C27 theophylline-binding RNA were readily assigned because they reside in the GAAA tetraloop and do not change chemical shift between the free and bound states of the RNA. At pH 6.8, A7 in the theophylline-binding core of the RNA was found to be in intermediate exchange between a protonated and unprotonated state (Figure 4A). Upon reducing the pH to 5.5, this A7 C2 resonance sharpens as it becomes completely protonated. At this pH another adenine residue in the core is in slow exchange between a protonated and unprotonated state as evidenced by the presence of two C2 resonances as shown in Figure 4B. These data provide strong evidence for N1-protonated adenine residues in the core of the theophylline-binding RNA in the absence of ligand at low pH.

^{13}C Relaxation Measurements. As shown in Figure 5, the $T_{1\rho}$ relaxation times for various carbon resonances in C27 are much longer than those measured for any other residue. The ribose C1′ carbon of C27 has a $T_{1\rho}$ of 38 ms, compared to the C1′ resonances of residues in other regions of the RNA, which all have $T_{1\rho}$ values of 14−25 ms (see Table 2).

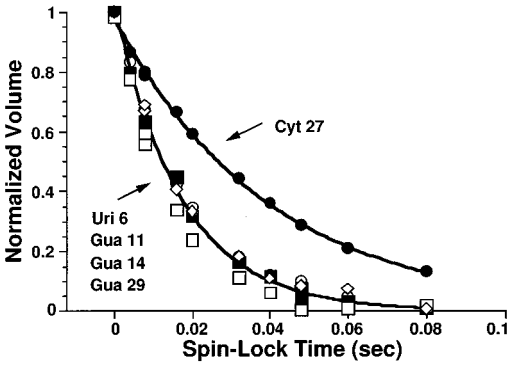

FIGURE 5: The $T_{1\rho}$ ^{13}C relaxation curves, plotted for some of the C1′ resonances in the RNA−theophylline complex. Two-dimensional (^{13}C, ^{1}H) NMR experiments employing a ^{13}C spin-lock were used to measure the $T_{1\rho}$ values (*27*). Peak volumes were normalized to the 0 mixing time point and are plotted versus the ^{13}C spin-lock time. The very slow relaxation of the C27 C1′ resonance indicates that this residue is highly dynamic relative to the other residues in the RNA−theophylline complex. Similar results were obtained for the C4′ and C5 resonances of residue C27 and for residue 27 in the A27 RNA (see Table 2).

Table 2: Carbon $T_{1\rho}$ Relaxation Times for the C27 RNA− and A27 RNA−Theophylline Complexes[a]

	$T_{1\rho}$ (ms)		
residue	C1′	C4′	C5/C8
C3			14 (17)
U6	19 (23)		(21)
C8			(21)
G11	18 (22)		(27)
G14	18		
A18	(22)		(22)
G19			(24)
C20			15 (24)
C22	23 (27)		
U23	25 (30)	(34)	23 (26)
U24	24 (27)		16 (24)
G25	(20)		
G26		18 (29)	(24)
C/A27	38 (39)	52 (47)	42 (53)
G29	14 (24)		
U32			27 (23)
C33	(28)	25	

[a] The $T_{1\rho}$ relaxation times for well-resolved cross-peaks were measured at 25 °C as described in the text. The values for the C27 RNA−theophylline complex are listed without parentheses, and the values for the A27 RNA−theophylline complex are given in parentheses.

This difference in relaxation times indicates that residue C27 is reorienting rapidly relative to the other RNA residues in the complex (*30*−*32*). The ^{13}C $T_{1\rho}$ relaxation times for some of the well-resolved C4′ and C5 resonances were also measured and the $T_{1\rho}$ times for C27 were again substantially longer than residues on the rest of the RNA (Table 2). Unfortunately we are not able to measure the ^{13}C relaxation times for the theophylline ligand because it is not ^{13}C-labeled. However the C1′ of U6, which is in the U6-U23-A28 triple that forms the floor of the theophylline binding pocket (*9*), has a $T_{1\rho}$ relaxation time similar to those measured for the residues in the stems, consistent with the binding pocket having similar dynamics to the helical regions of the RNA. The proton T_1 and T_2 relaxation times were measured by nonselective 1D inversion recovery and CPMG spin−echo pulse sequences (*32*), respectively, on the C27 RNA−theophylline complex at 25 °C (not shown). These data also

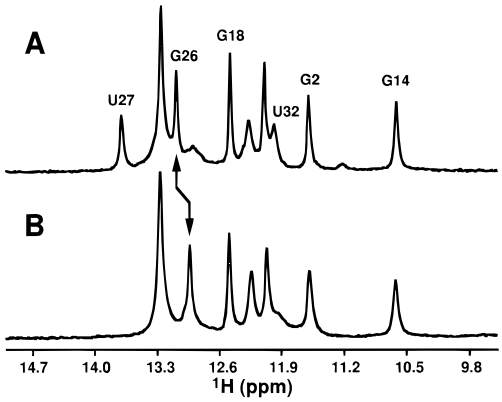

FIGURE 6: One-dimensional ^{1}H NMR spectra of the (A) U27 and (B) C27 theophylline-binding RNAs acquired in the absence of ligand at 15 °C (pH 6.8). A new peak at 13.7 ppm in the U27 RNA spectrum (relative to the spectrum of the C27 RNA) arises from the imino proton of U27, which exchanges slowly with solvent in this RNA. The arrow indicates the G26 imino proton resonance that shifts slightly between the U27 and C27 RNAs.

support the conclusion that residue 27 is highly dynamic within the RNA−theophylline complex. The ^{13}C $T_{1\rho}$ relaxation times were also measured for some well-resolved resonances in the ^{13}C/^{15}N-labeled A27 RNA−theophylline complex (see Table 2). Similar to the C27 RNA, the ^{13}C $T_{1\rho}$ relaxation times for A27 in this RNA were substantially longer than the relaxation times measured in other residues.

NMR of the U27 Theophylline-Binding RNA. The theophylline-binding RNA with a U at position 27 has a K_d of >20 μM (Table 1) at pH 7.5, which is >50-fold weaker binding than the wild-type C27 RNA. Because U27 resides in a symmetrical internal loop across from A7 (Figure 1), it is possible that these two residues can form a Watson−Crick base pair. To look for evidence for this base pair, the U27 theophylline-binding RNA was analyzed by 1D and 2D NMR spectroscopy. The 1D imino proton spectra of the C27 and U27 theophylline-binding RNAs in the absence of ligand are plotted in Figure 6. The imino proton resonances in the stem and hairpin loop regions of the free U27 RNA were assigned by NOE connectivities and by comparison with the complete resonance assignments of the C27 RNA−theophylline complex (*9*). The only significant differences between the spectra of the U27 (Figure 6A) and C27 (Figure 6B) RNAs are the appearance of a new resonance at 13.7 ppm (U27), a downfield shift for the resonance at 13.0 ppm (G26), and a new very weak resonance at 11.2 ppm. The chemical shift of the peak at 13.7 ppm indicates a U imino proton in a Watson−Crick base pair (*33*), and both intra- and inter-base pair NOEs (not shown) confirm the formation of this A7-U27 base pair. The weak resonance at 11.2 ppm has not been unambiguously assigned but may represent a small population of the U27 theophylline-binding RNA in which U6 is engaged in a non-Watson−Crick interaction with A28.

DISCUSSION

The theophylline-binding RNA motif consists of 14 conserved residues (boxed region of Figure 1) and one semiconserved residue at position 27 that can be an A or C (*9*). There are several possibilities why G or U residues were not observed in any of the theophylline-binding RNA aptamers. Nucleotides are conserved in order to maintain a specific interaction required for structure or function or to eliminate unfavorable electrostatic or steric interactions in the RNA. For example, there is an amino group on the major-groove face of the base at residue 27 in both the C27 and A27 RNAs, and this amino group could form a specific hydrogen bond required for ligand binding. The absence of a hydrogen-bond donor on the major-groove face of G and U might therefore lower binding affinity due to the loss of an important intermolecular interaction. However, there are examples where neither maintaining a specific interaction nor eliminating an unfavorable interaction appears to explain the observed sequence variations (*34*). In such cases certain nucleotides may stabilize an alternate inactive conformation of the free RNA, therefore inhibiting formation of the active species (*35*). A goal of this study was to understand how various nucleotides at position 27 affect theophylline binding in this RNA aptamer.

The three-dimensional structure of the theophylline-binding RNA was recently determined by solution NMR and revealed a well-ordered binding pocket for theophylline (*16*). Many of the core nucleotides form multiple interlocking interactions, which helps explain why these residues were so highly conserved in the in vitro selection (*16*). A striking result illustrated in Figure 7 is that residue C27 is completely disordered in the ensemble of structures calculated from the NMR data. No internucleotide NOEs were observed for protons on residue C27 and this absence of distance constraints is consistent with this residue being highly dynamic in solution (*16*). However, the *absence* of NOEs between residue C27 and the rest of the RNA does not

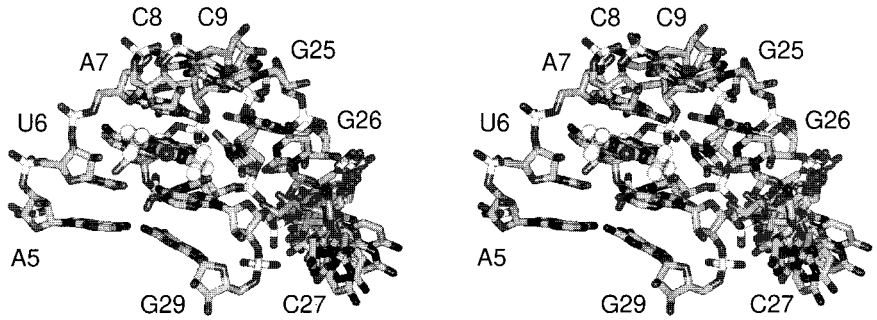

FIGURE 7: Stereoview of the ligand-binding core of the RNA−theophylline complex (*16*). One hundred structures were calculated by using the NMR-derived structural data as constraints, and the 10 lowest energy structures from this calculation were selected as a representative family. The position of C27 in each member of this family is shown relative to the other residues in the core of the refined average structure. The observed disorder for C27 is consistent with this residue being highly dynamic in the RNA−theophylline complex. The theophylline is depicted in ball-and-stick representation.

directly demonstrate unusual dynamics for this residue. ^{13}C relaxation measurements provide a means for directly probing rapid (subnanosecond) motions in biomolecules (*30*−*32*); therefore, $T_{1\rho}$ ^{13}C relaxation times were measured for various carbons in the RNA−theophylline complex. As illustrated in Figure 5 and Table 2, the $T_{1\rho}$ ^{13}C relaxation time for the C1′ on residue 27 is much longer than the relaxation times for other C1′ resonances in the RNA−theophylline complex. This long relaxation time indicates that this C1′−H1′ vector is reorienting much more rapidly than other C−H vectors in the molecule, directly demonstrating the dynamic nature of C27. The $T_{1\rho}$ ^{13}C relaxation times for the C4′ and C5 resonances are also much longer than the relaxation times for well-resolved resonances of other residues in the complex (Table 2). Figure 7 clearly shows that the base of C27 is disordered in the solution structures, which is confirmed by the $T_{1\rho}$ relaxation time for the C5 resonance. What is rather surprising is that the both the base (C5) and sugar (C1′ and C4′) carbons have long relaxation times, demonstrating that all these C−H vectors are quite dynamic. Thus the backbone of this nucleotide could be undergoing a crankshaft-type motion (*36*), which would allow a single nucleotide to change its orientation without affecting the conformation of neighboring residues. In the original selection both C and A nucleotides were observed at residue 27; therefore, we also measured $T_{1\rho}$ ^{13}C relaxation times in the A27 RNA−theophylline complex. Similar to the data for C27, A27 also has longer relaxation times than other residues in the RNA, again demonstrating that the sugar and base at 27 are highly dynamic (see Table 2). Thus, even though the adenine base has much greater propensity for forming stacking interactions than a cytidine, the base and sugar of A27 in the RNA−theophylline complex are much more dynamic than any other residue in the RNA.

The dynamic nature of residue 27 supports the hypothesis that this nucleotide is not important for theophylline binding. To directly test this prediction, the equilibrium constants for theophylline binding were measured for RNAs with different nucleotides at position 27. Figure 2 shows the equilibrium filtration measurements for RNAs containing ribo-C, deoxy-C, and deoxy-abasic nucleotides at residue 27, with the calculated equilibrium constants given in Table 1. It is especially striking that the abasic RNA binds theophylline with the same affinity as the other RNAs, because this means that the base of residue 27 is not directly involved with theophylline binding. These results predict that any base should be readily accommodated at position 27 in the RNA−theophylline complex. This surprising result is in conflict with the results of the in vitro selection, where G and U residues were not observed at position 27. The measured K_d values for the U27 and G27 RNAs confirm that these RNAs bind theophylline with lower affinities than the C27 and A27 RNAs (Table 1). This then raises the possibility that contacts made by position 27 in the free state of the RNA are inhibiting ligand binding. Residue 27 is directly across from the conserved A7 in the symmetric internal loop of the theophylline-binding RNA (Figure 1). A7 is a critical residue in the RNA−theophylline complex because it forms a base-platform (*37*) with C8 to create the roof of the ligand-binding pocket as shown in Figure 7 (*16*). A stable interaction between A7 and the nucleotide at position 27 would preclude formation of the A7-C8 base-platform motif

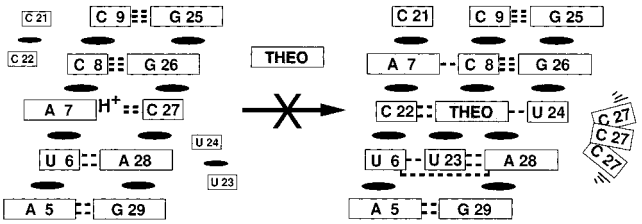

FIGURE 8: Schematic representation of the structural interactions in the free state of the RNA that mediate the pH dependence of ligand binding of the C27 theophylline-binding RNA. At low pH, an A7H$^+$-C27 base pair (left) blocks formation of the A7-C8 base-platform motif that caps the theophylline-binding site (right).

and therefore block the theophylline-binding site. This hypothesis provides a rationale for why G and U residues were not observed at position 27; U27 could form a Watson−Crick base pair with A7, and a G at position 27 could form a wobble (*38*) or sheared (*39*, *40*) GA pair to block formation of the binding pocket.

The NMR data in Figure 6 on the U27 theophylline-binding RNA support this model where an A7-U27 base pair in the free RNA blocks formation of the ligand-binding pocket. Comparison of the 1D spectra for the U27 and C27 RNAs reveals an additional resonance (13.7 ppm) in the free state of the U27 theophylline-binding RNA, which is assigned to the imino proton of U27. This A7-U27 base pair likely affects the conformation and environment of the C8-G26 pair above (Figure 1), because the G26 imino proton resonance is slightly downfield in the U27 RNA when compared to the C27 RNA (Figure 6). These results support the model in which interactions between G27 or U27 and A7 across the symmetric internal loop (Figure 1) stabilize the free state of the theophylline-binding RNA and therefore decrease the binding affinity of these aptamers.

The inhibitory role of the A7-N27 interaction is further supported by the unusual pH dependence of the C27 RNA. Recent studies have demonstrated that some A-C "mismatches" in internal loops in RNA actually form protonated AH$^+$-C base pairs (*29*, *41*, *42*). These protonated base pairs have been shown to bridge the junction between a helical stem and an internal loop in several RNAs (*29*, *41*, *42*). The secondary structure model for the C27 RNA aptamer in Figure 1, shows that an A7H$^+$-C27 base pair could form at the junction of a two base pair stem and the lower internal loop of this RNA. To test for formation of this A7H$^+$-C27 base pair, equilibrium filtration experiments were performed on the C27 and A27 RNAs as a function of pH as shown in Figure 3. The pH dependences for the C27 and A27 RNAs show that whereas theophylline binding to the C27 RNA aptamer is inhibited below 6.5, binding to the A27 RNA is only disrupted below pH 5.5. This difference indicates that some group in the C27 RNA is being protonated at a higher pH than in the A27 RNA, indicating that this group has a higher pK_a in the C27 than the A27 RNA. We have previously shown how ^{13}C NMR can be used to identify adenine residues with elevated pK_as (*29*). An adenine residue with a protonated N1, such as in an AH$^+$-C base pair, exhibits a C2 resonance that is shifted upfield by 8.2 ppm (*29*). The 2D (^{13}C,^1H) HSQC spectra of the C27 theophylline-binding RNA at pH 6.8 and 5.5 provide direct evidence for protonation of the N1 on A7 (see Figure 4), most likely arising from formation of an A7H$^+$-C27 base pair. Figure 8 shows

a schematic representation of how protonation of A7 at low pH could lead to formation of this A7H$^+$-C27 pair, which stabilizes the internal loop and blocks theophylline binding. These NMR data therefore support a model in which an A7H$^+$-C27 base pair prevents formation of the ligand-binding pocket, resulting in the observed pH dependence of the theophylline-binding activity of the C27 RNA.

CONCLUSION

The NMR and equilibrium binding data presented here indicate that interactions between residues 7 and 27 in the free state of the theophylline-binding RNA motif affect the ligand-binding activity of this aptamer. This important role for residue 27 is surprising because ^{13}C relaxation measurements show that this residue is highly dynamic in the RNA−theophylline complex. Furthermore, it was demonstrated that the base of residue 27 is not involved in any inter- or intramolecular interactions in the complex because an RNA with an abasic residue at this position exhibits wild-type binding affinity. The unusual pH dependence of the C27 RNA arises from formation of a protonated A7H$^+$-C27 in the free RNA, which inhibits formation of the RNA−theophylline complex. Since residue A7 is required for formation of the critical base-platform motif with C8 to help create the theophylline-binding site (*16*), formation of an A7H$^+$-C27 base pair in the absence of theophylline explains the unusual pH dependence of the ligand binding of the C27 RNA. Thus the effects of residue 27 on ligand binding are mediated from the free state of the RNA. When interpreting the effects of nucleotide substitutions on the equilibrium binding constant of a complex, it is tempting to assume that the observed changes arise from loss of a specific interaction or a steric clash in the *bound* form of the complex. These results for residue 27 in the theophylline-binding RNA aptamer provide a striking reminder that the *free* state of the species participating in the complex also contributes to the binding energy and must therefore be considered in the interpretation of any mutational analysis.

ACKNOWLEDGMENT

We thank Anne Gooding for synthesis of RNA.

REFERENCES

1. Gutell, R. R., and Damberger, S. H. (1996) in *Ribosomal RNA and Group I Introns* (Green, R., and Schroeder, R., Eds.) pp 15−32, Chapman & Hall, New York.
2. Davis, J. P., Janjic, N., Javornik, B. E., and Zichi, D. A. (1996) *Methods Enzymol. 267*, 302−314.
3. Jaeger, J. A., Turner, D. H., and Zuker, M. (1990) *Methods Enzymol. 183*, 281−306.
4. Major, F., Turcotte, M., Gautheret, D., Lapalme, G., Fillion, E., and Cedergren, R. (1991) *Science 253*, 1255−1260.
5. Gautheret, D., Major, F., and Cedergren, R. (1993) *J. Mol. Biol. 229*, 1049−1064.
6. Louisemay, S., Auffinger, P., and Westhof, E. (1996) *Curr. Opin. Struct. Biol. 6*, 289−298.
7. Gaspin, C., and Westhof, E. (1995) *J. Mol. Biol. 254*, 163−174.
8. Uhlenbeck, O. C., Pardi, A., and Feigon, J. (1997) *Cell 90*, 833−840.
9. Jenison, R. D., Gill, S. C., Pardi, A., and Polisky, B. (1994) *Science 263*, 1425−1429.
10. Tuerk, C., and Gold, L. (1990) *Science 249*, 505−510.
11. Ellington, A. D., and Szostak, J. W. (1990) *Nature 346*, 818−822.
12. Szostak, J. W. (1992) *Trends Biochem. Sci. 17*, 89−93.
13. Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995) *Annu. Rev. Biochem. 64*, 763−797.
14. Conrad, R. C., Baskerville, S., and Ellington, A. D. (1995) *Mol. Divers. 1*, 69−78.
15. Tuerk, C. (1997) *Methods Mol. Biol. 67*, 219−230.
16. Zimmermann, G. R., Jenison, R. D., Wick, C. L., Simorre, J.-P., and Pardi, A. (1997) *Nat. Struct. Biol. 4*, 644−649.
17. Poncelet, S. M., Limet, J. N., Noel, J. P., Kayaert, M. C., Galanti, L., and Collet-Cassart, D. (1990) *J. Immunoassay 11*, 77−88.
18. Milligan, J. F., and Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.
19. Nikonowicz, E. P., Sirr, A., Legault, P., Jucker, F. M., Baer, L. M., and Pardi, A. (1992) *Nucleic Acids Res. 20*, 4507−4513.
20. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8789.
21. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.
22. Sklenár, V., and Bax, A. (1987) *J. Magn. Reson. 75*, 378−383.
23. Bodenhausen, G., and Ruben, D. J. (1980) *Chem. Phys. Lett. 69*, 185−189.
24. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
25. Mori, S., Abeygunawardana, C., Johnson, M. O., and Vanzijl, P. C. M. (1995) *J. Magn. Reson. B 108*, 94−98.
26. Bax, A., Ikura, M., Kay, L. E., and Zhu, G. (1991) *J. Magn. Reson. 91*, 174−178.
27. Yamazaki, T., Muhandiram, R., and Kay, L. E. (1994) *J. Am. Chem. Soc. 116*, 8266−8278.
28. Gill, S. C., Weitzel, S. E., and von Hippel, P. H. (1991) *J. Mol. Biol. 220*, 307−324.
29. Legault, P., and Pardi, A. (1994) *J. Am. Chem. Soc. 116*, 8390−8391.
30. Bruschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M., and Ernst, R. R. (1992) *J. Am. Chem. Soc. 114*, 2289−2302.
31. Peng, J. W., and Wagner, G. (1994) *Methods Enzymol. 239*, 563−596.
32. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, CA.
33. Wijmenga, S. S., Mooren, M. M. W., and Hilbers, C. W. (1993) in *NMR of Macromolecules* (Roberts, G. C. K., Ed.) pp 217−288, Oxford University Press, Oxford, England.
34. McKay, D. B. (1996) *RNA 2*, 395−403.
35. Uhlenbeck, O. C. (1995) *RNA 1*, 4−6.
36. Heath, P. J., Gebe, J. A., Allison, S. A., and Schurr, J. M. (1996) *Macromolecules 29*, 3583−3596.
37. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1696−1699.
38. Wu, M., and Turner, D. H. (1996) *Biochemistry 35*, 9677−9689.
39. Li, Y., Zon, G., and Wilson, W. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26−30.
40. Chou, S. H., Cheng, J. W., and Reid, B. R. (1992) *J. Mol. Biol. 228*, 138−155.
41. Legault, P., and Pardi, A. (1997) *J. Am. Chem. Soc. 119*, 6621−6628.
42. Cai, Z., and Tinoco, I., Jr. (1996) *Biochemistry 35*, 6026−6036.

BI980082S