Role of a Heterogeneous Free State in the Formation of a Specific RNA−Theophylline Complex†

Fiona M. Jucker, Rebecca M. Phillips,‡ Scott A. McCallum, and Arthur Pardi*

Department of Chemistry and Biochemistry, 215 UCB, University of Colorado, Boulder, Colorado 80309-0215

Received November 1, 2002; Revised Manuscript Received January 7, 2003

ABSTRACT: The helical regions of RNA are generally very stable, but the single-stranded and loop regions often exist as an ensemble of conformations in solution. The theophylline-binding RNA aptamer forms a very stable structure when bound to the bronchodilator theophylline, but the theophylline binding site is not stably formed in the absence of ligand. The kinetics for theophylline binding were measured here by stopped-flow fluorescence spectroscopy to probe the mechanism for theophylline binding in this RNA aptamer. The kinetic studies showed that formation of the RNA−theophylline complex is over 1000 times slower than a diffusion-controlled rate, and the high affinity of the RNA−theophylline complex arises primarily from a slow dissociation rate for the complex. A theophylline-independent rate was observed for formation of the theophylline−RNA complex at high theophylline concentration, indicating that a conformational change in the RNA is the rate-limiting step in complex formation under these conditions. The RNA−theophylline complex requires divalent metal ions, such as Mg^{2+}, to form a high-affinity complex, and there is a greater than 10000-fold reduction in affinity for theophylline in the absence of Mg^{2+}. This decrease in binding affinity in the absence of Mg^{2+} results primarily from an increased dissociation rate for the complex. The implications of an ensemble of conformations in the free state of this theophylline-binding RNA are discussed and compared with mechanisms for formation of protein−ligand complexes.

Many RNAs exist as an ensemble of conformations in solution (*1*−*3*). This arises because the helical regions of RNA are quite stable and RNA can readily form partially folded states, consisting of stable helical secondary structure but little or no tertiary structure (*4*−*6*). This is in contrast to most proteins where the individual α-helices and *β*-sheets are not stable on their own and protein domains often fold in a cooperative manner (*7*). The ability of individual secondary structure units to stably form in solution has important implications for RNA catalysis and recognition (*8*−*11*). For example, the lead-dependent ribozyme, the hammerhead ribozyme, and the hairpin ribozyme appear to exist as a distribution of conformational states with only a minor population in the active species, and a conformational change is required to populate the catalytically active species (*11*−*19*). Therefore, the rate of interconversion from an inactive to active state may represent the rate-limiting step for reaction (*20*, *21*). RNA recognition may have similar properties, where induced fit or conformational capture mechanisms are thought to be important for many RNAs that bind ligands with high affinity and specificity (*2*, *22*−*24*). For example, the HIV TAR[1] and RRE RNAs, the U1A binding RNA, and the S15 binding region of 16S rRNA form high-affinity complexes with their binding partners, yet these RNAs display rather unstructured or heterogeneous states in the absence of their binding partner (*2*, *22*, *23*). In an induced fit or conformational capture mechanism, the ligand selects a subset of the RNAs that are in conformations amenable to binding. Although this mechanism has been proposed to play an important role in RNA−ligand binding, there is presently little known about how the rates for interconversion between these different conformations affect the activity of RNA.

To probe how an ensemble of conformations in the free state of an RNA can affect the mechanism for RNA−ligand binding, we have studied the kinetics of an RNA aptamer binding to its high-affinity ligand theophylline. This RNA was identified by in vitro selection with the goal of developing a diagnostic tool to measure serum levels of the bronchodilator drug theophylline (Figure 1A) without interference from the structurally similar caffeine (*25*). This theophylline-binding aptamer is a small RNA that displays high affinity ($K_d = 0.2-0.4\ \mu M$) and specificity for its ligand. The ∼10000-fold discrimination between theophylline and caffeine provided the first example that RNA is capable of discriminating between small ligands of similar structure with specificity rivaling that of antibodies.

† This work was supported by NIH Grant AI 30726. We also thank the W. M. Keck Foundation for support of the Molecular Structure Program on the Boulder Campus.

* To whom correspondence should be addressed. E-mail: arthur.pardi@colorado.edu.

‡ Current address: Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0332.

[1] Abbreviations: EDTA, ethylenediaminetetraacetate; FRET, fluorescence resonance energy transfer; HIV, human immunodeficiency virus; HPLC, high-performance liquid chromatography; NMR, nuclear magnetic resonance; RRE, Rev response element; TAR, transactivation response region; UTR, untranslated region.

10.1021/bi027103+ CCC: $25.00
Published on Web 02/11/2003

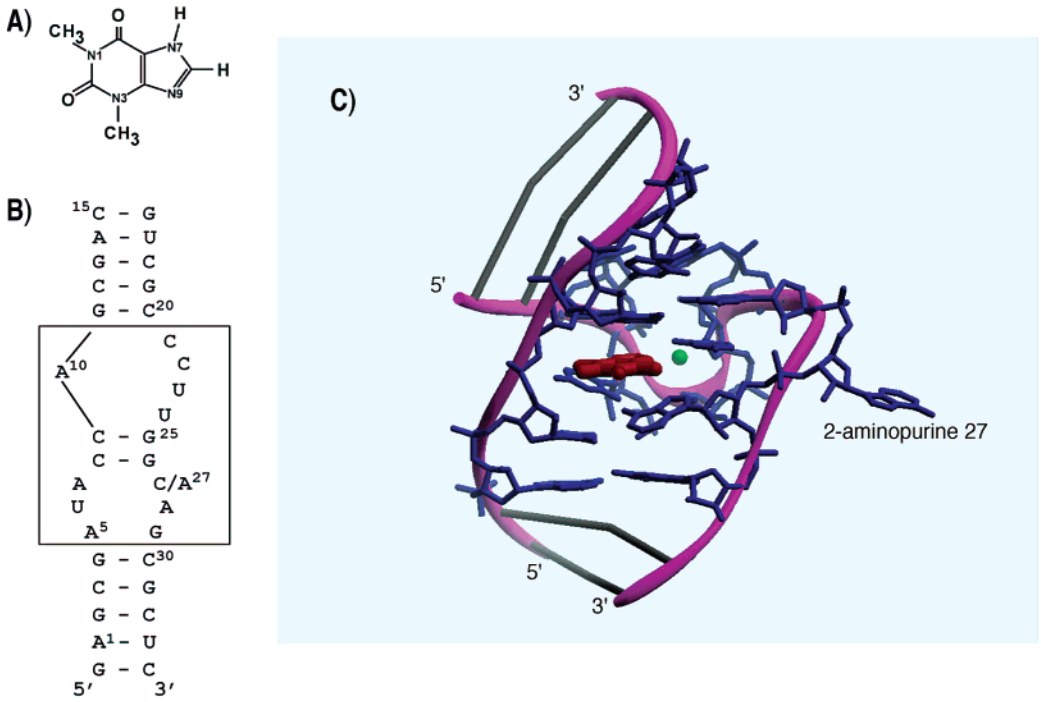

FIGURE 1: (A) Chemical structure of the bronchodilator theophylline. (B) Secondary structure of the theophylline-binding RNA. The nucleotides in the box are conserved and form the theophylline-binding pocket. (C) NMR solution structure of the conserved region of the RNA−theophylline complex (*26*). A Mn^{2+} site in the S-turn (*28*) is shown in green, and the theophylline is shown in red. The figure was produced with the RIBBONS program (*50*).

The structural basis for theophylline binding to the RNA aptamer has been well characterized by NMR and mutational analysis (*25*−*28*). The RNA aptamer contains a highly conserved 15-nucleotide theophylline-binding pocket surrounded by two helical regions (Figure 1B). The RNA−theophylline complex shown in Figure 1C consists of an intricate network of interlocking structural motifs, where the binding pocket is formed by an S-turn, a base-zipper motif, an A-platform, and several base triples and noncanonical base pairs (*26*). These motifs assemble into a complicated fold where many of the bases in the binding pocket participate in multiple motifs. Only one base in the conserved region, position 27, has no interactions with other nucleotides in the binding pocket and is freely exposed to solvent (Figure 1C). Mutational analysis revealed that this base can be either a cytidine, an adenine, or an abasic site without affecting binding affinity (*27*). For the studies here, residue 27 was replaced with the fluorescent adenine analogue 2-aminopurine to probe the kinetics of theophylline binding by stopped-flow fluorescence spectroscopy.

Previous NMR studies of the theophylline-binding RNA showed that the free RNA does not adopt a unique conformation and that the RNA only folds into a well-ordered structure in the presence of theophylline (*26*, *29*). The study here investigates the role of the heterogeneous free state in the formation of the RNA−theophylline complex. Kinetic characterization of theophylline binding revealed slow association kinetics, over 1000 times slower than a diffusion-controlled reaction. Since Mg^{2+} is required for formation of a high-affinity RNA−theophylline complex, the role of Mg^{2+} in the kinetics of theophylline binding was also investigated. The results indicate that Mg^{2+} increases the lifetime of the RNA−theophylline complex, consistent with previous NMR studies that showed a divalent metal binding site in the S-turn of the theophylline-binding pocket (*28*). The kinetic studies also showed that theophylline binding occurs by a two-step reaction. A conformational change is required prior to theophylline binding, where the unstructured free RNA is converted into a binding-competent species. This induced fit mechanism appears to be common in RNA−ligand interactions. The molecular properties of RNAs that contribute to this mechanism are discussed and contrasted with molecular recognition by proteins.

MATERIALS AND METHODS

Sample Preparation. HEPES and theophylline were obtained from Sigma; all other chemicals were obtained from Fisher, unless otherwise noted. The RNA oligonucleotides were obtained from Dharmacon and deprotected according to manufacturer's recommendation. The RNAs were purified over a Dionex NucleoPac PA 100 (9 × 250 mm) anion-exchange HPLC column. The column was heated to 85 °C, and a linear gradient was applied between (A) 20 mM Tris, pH 8.0, and 1 mM EDTA and (B) 20 mM Tris, pH 8.0, 1 M NH_4Cl, and 1 mM EDTA at a flow rate of 3 mL/min. Full-length RNA eluted between 400 and 600 mM NH_4Cl and was ethanol precipitated prior to desalting over consecutive NAP5 and NAP10 columns (Amersham Pharmacia Biotech AB). All buffer components used for sample preparation (except $MgCl_2$) were applied to Chelex resin (Sigma) to remove divalent cations. The RNA duplex was formed by annealing the two strands at a 1:1 ratio in 50 mM HEPES, pH 7.5, 100 mM NaCl, and 1 mM EDTA. The mixture was heated to 65 °C for 3 min followed by slow cooling. The annealed RNA duplex was exchanged into H_2O using Microcon YM3 centrifugal filter devices (Amicon Bioseparations) and purified over a NAP 5 column. For each set of stopped-flow experiments, the duplex RNA was diluted to

2 μM in reaction buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA for experiments in the presence of Mg^{2+} or 50 mM HEPES, pH 7.5, 100 mM NaCl, and 1 mM EDTA for experiments in the absence of Mg^{2+}) and incubated for 3 min at 37 °C followed by 30 min at room temperature, before the RNA was diluted to a final concentration of 0.2 μM.

Measurement of Equilibrium Constant. The equilibrium constant for theophylline binding was determined in 100 mM HEPES, pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM EDTA, at 25 °C from the equilibrium fluorescence of stopped-flow fluorescence experiments. Stopped-flow measurements of 2-aminopurine-modified RNA were performed on a SX.17MV stopped-flow spectrofluorometer (Applied Photophysics). The excitation wavelength was 318 nm, and the fluorescence intensity was monitored with a 335 nm longpass filter. A total of 4000 data points were collected for each experiment corresponding to times from 50 ms to 20 s. The dead time of the instrument was 2 ms. For each reaction 5−10 experiments were averaged, and baseline correction was applied to correct for instrument drift. Binding was initiated by rapid mixing of RNA and theophylline. The RNA concentration was 0.1 μM after mixing, and the theophylline concentration ranged from 0.5 to 15 μM after mixing. After the binding reaction reached equilibrium, the equilibrium fluorescence F_{eq} was measured, and the apparent K_d was determined by fitting to the equation:

$$F_{eq} = \frac{F_{max}[\text{theophylline}]}{K_d + [\text{theophylline}]} \quad (1)$$

where F_{max} is the maximum fluorescence possible when all of the RNA duplex is bound to theophylline and the equilibrium constant K_d = [theophylline][RNA_{free}]/[RNA_{bound}].

Fluorescence-Detected Stopped-Flow Kinetics. Association rate constants were determined from similar stopped-flow experiments as above. The RNA concentration after mixing was 0.1 and 0.25 μM for micromolar and millimolar theophylline concentrations, respectively. Theophylline concentration in reaction buffer varied from 0.1 μM to 15 mM after mixing. All solutions were at least 10 mM below the solubility limit of theophylline (45 mM). All experiments were conducted at 25 °C, except for some experiments at high theophylline concentration that were also performed at 15 °C. The stopped-flow fluorescence data were fit to eq 2 using the KaleidaGraph program:

$$F = F_{eq}(1 - e^{-k_{observed}t}) + C \quad (2)$$

where F is the observed fluorescence as a function of time, F_{eq} is the equilibrium fluorescence, $k_{observed}$ is the observed rate constant for a given theophylline concentration, and C is the background fluorescence. This analysis (eq 2) does not include product dissociation during the experiment, since including the reverse reaction in the analysis resulted in the same values for $k_{observed}$ (data not shown).

To measure the dissociation rate constant, the RNA−theophylline complex was preformed with varying amounts of theophylline (0.1 μM 2-aminopurine RNA with 0.25, 0.50, and 1.0 μM theophylline after mixing) and chased with unmodified RNA (5.0 μM after mixing). The dissociation rate constant was fit to the equation:

$$F = F_0(e^{-k_{off}t}) + C \quad (3)$$

where F_0 is the starting fluorescence and k_{off} is the dissociation rate constant.

Data Simulations. Data simulations were used to calculate the concentrations of all RNA species from the rate law for eq 5, and the time dependence of the RNA−theophylline complex was compared with the experimental stopped-flow data. The FORTRAN program used for these simulations is included in Supporting Information and required input of the rate constants k_1, k_{-1}, k_2, and k_{-2} and the total concentration of theophylline and RNA. The rate constants k_1, k_{-1}, and k_2 were manually optimized to give the best visual fit to the experimental data. For a given theophylline concentration the simulation program samples a range of values for one rate constant while holding the other three rate constants fixed. Since there is no direct experimental information for k_{-1}, the reverse reaction of the conformational change, we first sampled a wide range for k_{-1} while holding k_1, k_2, and k_{-2} fixed at the experimentally derived values. Simulations were first performed at low theophylline concentrations to reproduce the experimental data. After determining the range of rate constants that fit these experimental data, simulations were performed at high theophylline concentrations. This iterative process was repeated until the simulations satisfied all experimental data, from 0.1 μM to 15 mM theophylline.

RESULTS

2-Aminopurine Fluorescence as a Probe of Theophylline Binding. The kinetics of theophylline binding were characterized by stopped-flow fluorescence spectroscopy. To probe theophylline binding, 2-aminopurine was introduced at the nonessential position 27 within the conserved binding region (Figure 1B,C). This base fluoresces when exposed to solvent, but the fluorescence is quenched when the base forms stacking interactions (*30*, *31*). Since the base of residue 27 is fully solvent exposed in the structure of the RNA−theophylline complex, we predicted that 2-aminopurine would fluoresce in the complex. Indeed, the fluorescence of the RNA−theophylline complex is equivalent to that of free 2-aminopurine (data not shown). Furthermore, there is no detectable fluorescence in the uncomplexed RNA, consistent with a stacked conformation of the 2-aminopurine in the free RNA, leading to the model in Figure 2. Figure 3A shows some of the stopped-flow data used to determine an equilibrium constant for theophylline binding. The observed fluorescence reaches different equilibrium levels as a function of theophylline concentration, and these values yield a K_d of 0.2−0.4 μM at 25 °C (Figure 3B), which is the same K_d previously determined for unmodified RNA by equilibrium filtration (*25*, *27*). Thus, the 2-aminopurine-modified RNA provides an excellent system for monitoring theophylline binding.

Kinetics of Theophylline Binding. To gain a better understanding for the mechanism of theophylline binding, the kinetics were probed by rapid mixing stopped-flow fluorescence spectroscopy. Data from some experiments performed at low theophylline concentration (0.1−15 μM) are shown in Figure 3A. These data were collected at high enough Mg^{2+} concentration (10 mM), where both the equilibrium constant

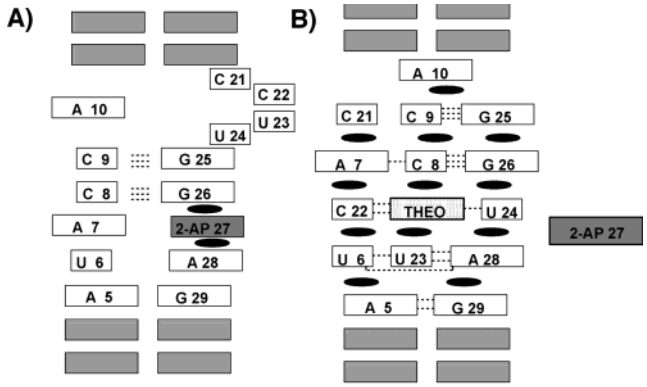

FIGURE 2: Schematic diagram illustrating the change in stacking interactions for 2-aminopurine in the free RNA and the RNA−theophylline complex. (A) A proposed model for the structure of the free RNA, where the 2-aminopurine at position 27 is in a stacked conformation, leading to quenching of the fluorescence. (B) In the RNA−theophylline complex the fluorescent base 2-aminopurine is protruding into the solvent, giving a fluorescent RNA (*26*).

and the rate of association for theophylline binding are unaffected by increasing Mg^{2+} concentration (data not shown). The apparent association rate constant k_{on} was obtained by fitting the data to a pseudo-first-order reaction mechanism:

$$\text{RNA} + \text{Theo} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{RNA·Theo} \qquad (4)$$

This simple model fits the data at micromolar theophylline concentrations. A plot of the observed rate constant versus theophylline concentration confirms pseudo-first-order binding kinetics (Figure 3C) and yields an apparent association rate constant k_{on} of $(1.7 \pm 0.2) \times 10^5\ M^{-1}\ s^{-1}$ at 25 °C (Table 1).

The dissociation rate constant was determined by rapid mixing of excess free unmodified RNA with the fluorescent RNA−theophylline complex and monitoring the decrease in fluorescence. The data fit a single-exponential yielding a dissociation rate constant k_{off} of $0.07 \pm 0.02\ s^{-1}$, and this dissociation rate constant was independent of theophylline concentration (data not shown). The ratio of $k_{off}/k_{on} \sim 0.4\ \mu M$ is in excellent agreement with the K_d for theophylline binding (0.2−0.4 μM).

The experiments at low theophylline concentrations (0.1−15 μM theophylline) showed a linear dependence of the observed rate with theophylline concentration. However, at

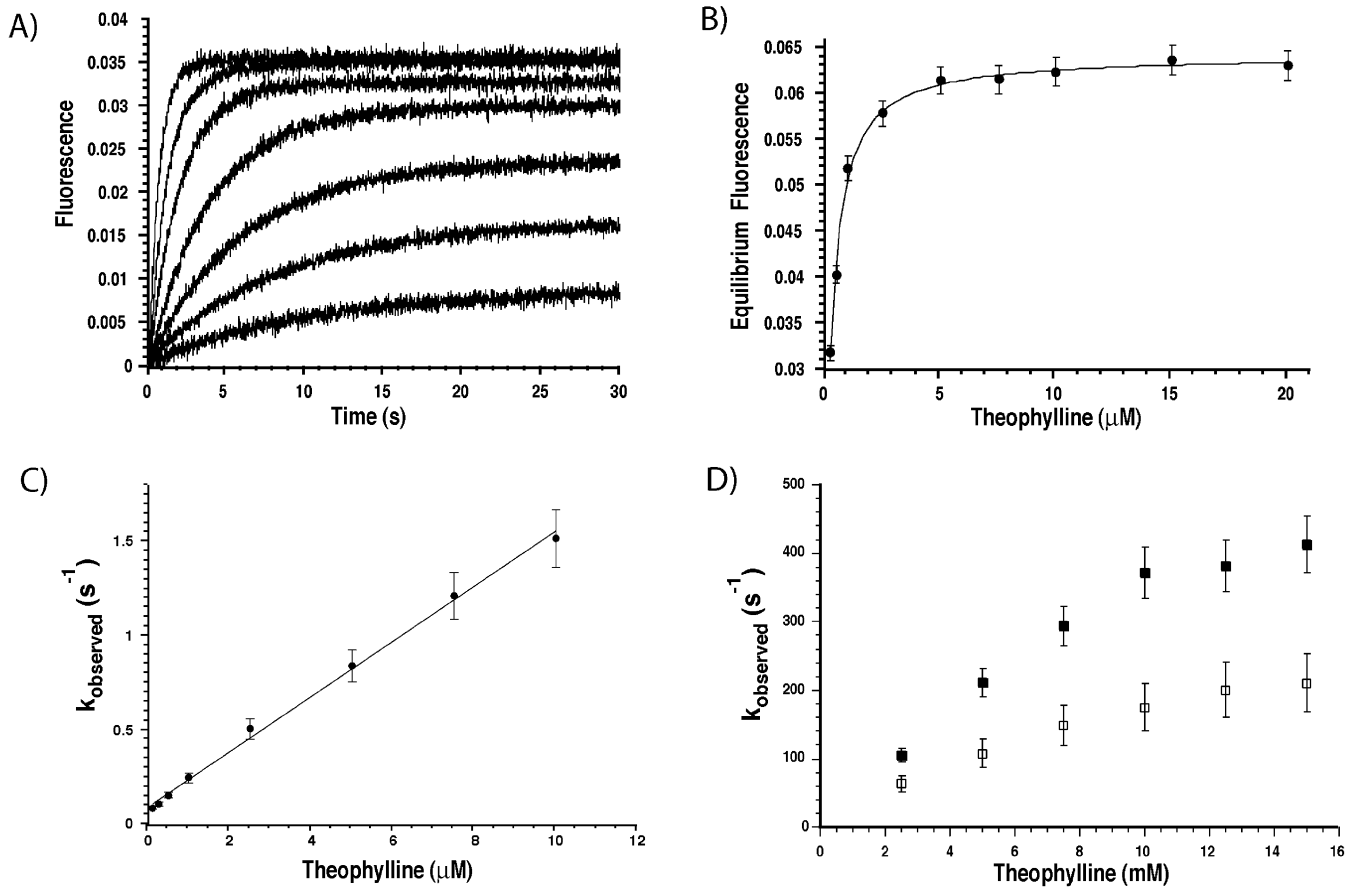

FIGURE 3: (A) Stopped-flow fluorescence spectroscopy was used to monitor the kinetics of theophylline binding. Examples of the data are shown for different theophylline concentrations (0.1, 0.25, 0.5, 1.0, 2.5, 5.0, and 10 μM). The equilibrium fluorescence was used to determine the equilibrium constant, and the early time point data were used to determine the observed rate constant. (B) The equilibrium constant for theophylline binding was determined from the equilibrium fluorescence of 2-aminopurine-substituted RNA obtained at a variety of theophylline concentrations. An apparent K_d for theophylline binding of 0.2−0.4 μM at 25 °C was obtained from these data. (C) The observed rate for complex formation, $k_{observed}$, is shown as a function of theophylline for micromolar theophylline concentrations where the observed rate is first order in respect to theophylline. The slope corresponds to the association rate constant k_{on}, and the *y*-intercept represents k_{off}. (D) At high theophylline concentrations the observed rate for complex formation starts to become independent of theophylline concentration. The data for 15 and 25 °C are shown as open and filled squares, respectively.

Table 1: Summary of Experimental and Simulated Kinetic and Thermodynamic Parameters for the Theophylline−RNA Complex at 25 °C

	K_d[a] (μM)	k_1 (s^{-1})	k_{-1} (s^{-1})	k_2 (M^{-1} s^{-1})	k_{-2} (s^{-1})
results derived from stopped-flow experiments in 10 mM Mg^{2+}	0.3 ± 0.1	500 ± 100	nd	$(1.7 \pm 0.2) \times 10^5$	0.07 ± 0.02
range of parameters obtained from simulations of stopped-flow data in 10 mM Mg^{2+}	0.2−0.6	400−500	300−1000	$(2-5) \times 10^5$	nd
results derived from stopped-flow experiments in the absence of Mg^{2+}	~7000[b]	$\sim 50 \pm 15$	nd	$\geq 5 \times 10^3$	nd
range of parameters obtained from simulations of stopped-flow data in the absence of Mg^{2+}[c]	nd	nd	30−800	$(5-500) \times 10^3$	≥20

[a] Equilibrium constants for theophylline binding. [b] The K_d in the absence of Mg^{2+} was obtained from NMR titration experiments (*29*). [c] The simulations for no Mg^{2+} data were performed for a given range of k_2 values, assuming a lower limit for k_2 from the experimentally derived lower limit of $\geq 5 \times 10^3$ M^{-1} s^{-1} and an upper limit for k_2 obtained from the fastest rate consistent with the simulations for the Mg^{2+} data.

higher theophylline concentrations, the observed rate deviates from linearity, becoming almost independent of theophylline at the highest concentrations (≥10 mM) as seen in Figure 3D. These data indicate that at high theophylline concentration a different step in the reaction becomes rate limiting with a rate constant of 500 ± 100 s^{-1} at 25 °C. Since the kinetic data do not fit the simple one-step binding mechanism over the full range of theophylline concentration, a two-step mechanism had to be invoked:

$$\begin{array}{ccc} \mathbf{RNA^{active} + Theo} & \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} & \mathbf{RNA{\bullet}Theo} \\ k_1 \uparrow \downarrow k_{-1} & & \\ \mathbf{RNA^{inactive}} & & \end{array} \qquad (5)$$

The first step in this mechanism represents a conformational change in the RNA, from an ensemble of inactive species, $RNA^{inactive}$, into binding-competent species, RNA^{active}. Complex formation is fast at high (millimolar) theophylline concentrations, and the conformational change in the RNA becomes rate limiting for theophylline binding (Figure 3D). Thus only at low theophylline concentrations does the experimentally determined $k_{observed}$ equal k_2[theophylline] (eq 5). However, the dissociation rate constant k_{off} (eq 4) is equivalent to k_{-2} under all conditions (eq 5), because both $RNA^{inactive}$ and RNA^{active} are nonfluorescent species (see below). A summary of the kinetic parameters directly determined from the experimental data is given in Table 1.

Two-Step Binding Mechanism Is Confirmed by Simulations. Simulations were performed to test whether the experimental data fit the two-step mechanism in eq 5 over the full range of theophylline concentrations (0.1 μM to 15 mM). Stopped-flow data were simulated by calculating the time-dependent concentrations of the three RNA species according to the rate law for eq 5 (Figure 4). The simulation program requires input of the individual rate constants and the total concentrations of RNA and theophylline. The rate constants were iteratively adjusted to give the best fit to the experimental stopped-flow data at a given theophylline concentration. This process was repeated for 0.1 μM to 15 mM theophylline concentrations until a self-consistent set of rate constants was obtained. A k_{-2} of 0.07 s^{-1} was used in all the simulations since this rate constant was directly determined from the experiments with excess unmodified RNA. Simulations for the other stopped-flow data gave a range of rate constants: k_1 from 400 to 500 s^{-1}, k_{-1} from 300 to 1000 s^{-1}, and k_2 from 2×10^5 to 5×10^5 M^{-1} s^{-1} (Table 1).

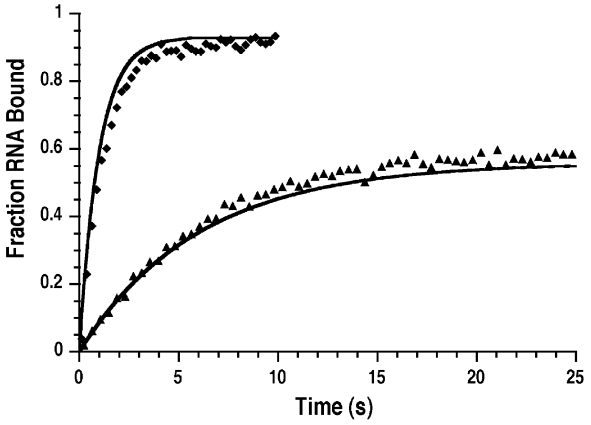

FIGURE 4: Simulations were used to confirm the proposed binding mechanism. The experimental data for 0.5 and 5 μM theophylline are shown as triangles and diamonds, respectively, where only a small percentage of each data set is plotted. The simulation for each theophylline concentration is shown as a solid line. The following rate constants were used for the simulations: $k_1 = 500$ s^{-1}, $k_{-1} = 300$ s^{-1}, $k_2 = 3 \times 10^5$ M^{-1} s^{-1}, and $k_{-2} = 0.07$ s^{-1}.

The equilibrium constant K_1 for the first step in eq 5 was determined as the ratio of k_1 and k_{-1}, where k_1 was estimated from both experiments and simulations and k_{-1} was obtained from the simulations. The K_1 of 0.5−1.6 for conversion of $RNA^{inacitve}$ to RNA^{acitve} indicates that 33−62% of the RNA is in a conformation amenable to binding theophylline. No fluorescence was detected in the free RNA, and the difference in fluorescence between free and fully bound RNA is equal to the fluorescence observed for an equal concentration of free 2-aminopurine (data not shown). This indicates that no fluorescent RNA species is present in significant concentration in the free RNA, implying that RNA^{active} is a nonfluorescent species or is being quenched by fast interconversion with another nonfluorescent species.

Since RNA^{acitve} is significantly populated, a burst in the reaction kinetics would occur if complex formation were very fast compared to the conformational change between $RNA^{inactive}$ and RNA^{acitve}. Indeed, a burst is observed in the simulations but only at very high theophylline concentrations (>50 mM), much higher than the experimental range. However, the simulations show no burst under any of the experimental theophylline concentrations. This is illustrated in Figure S1 in Supporting Information, which shows the experimental and simulated binding kinetics for 7.5 mM theophylline.

Magnesium Ion Dependence of Theophylline Binding. Mg^{2+} is required for high-affinity theophylline binding and plays an important role in stabilizing the RNA−theophylline complex (*25*, *28*). It has been shown that the affinity for

theophylline increases over 10000-fold in the presence of Mg^{2+} (*29*). To investigate the origin of the dramatic effect of Mg^{2+}, the kinetics for complex formation were also measured in the absence of Mg^{2+}. Due to the weak binding of theophylline in the absence of Mg^{2+} only a very limited set of kinetic data could be obtained. Since conditions could not be reached where the second step in eq 5 is rate limiting, only a lower limit for $k_2 \geq 5 \times 10^3$ M^{-1} s^{-1} was obtained (Table 1). The simulations rendered a lower limit for the dissociation rate constant $k_{-2} \geq 20$ s^{-1}. This implies an upper limit for the lifetime of the complex ($1/k_{-2}$) of ≤ 50 ms, compared to a lifetime of 14 s with Mg^{2+}, corresponding to a ≥ 280-fold increase in lifetime in the presence of Mg^{2+}.

DISCUSSION

The molecular interactions that contribute to high affinity and specificity of ligand−target interactions can be probed by structural and thermodynamic studies. However, these studies generally give limited information on the mechanism of complex formation. Thus kinetic studies are often used to complement the structural and thermodynamic data by providing information on the rate-limiting step, intermediates, or alternate pathways in a reaction. The theophylline-binding RNA aptamer has been extensively characterized by binding studies, mutational analysis, and NMR spectroscopy, providing a wealth of information on the structure and the stability of the complex as well as on the function of individual bases in the RNA (*25*−*28*). However, these studies did not provide data on the mechanism for theophylline binding to the RNA. In the study here, the kinetics of complex formation are investigated using stopped-flow spectroscopy. The results show that the RNA−theophylline complex forms with slow kinetics relative to diffusion controlled binding, and a conformational change occurs prior to theophylline binding.

A Conformational Change Is Required for Theophylline Binding. Proteins often bind small ligands by a simple one-step reaction mechanism where the kinetics are first order in ligand concentration (*7*, *32*). In many of these systems the free protein is in a conformation amenable to bind the ligand. This simple one-step binding mechanism was observed for the RNA−theophylline complex at low theophylline concentration. However, the kinetic data do not fit this mechanism at all theophylline concentrations, and a more complex two-step binding mechanism is required to fit the data (eq 5). The first step in this mechanism is theophylline independent and involves a conformational change from a binding inactive ($RNA^{inacitve}$) to a binding active RNA (RNA^{acitve}), and the second step is formation of the RNA−theophylline complex. As seen in Figure 3D, the observed rate starts to plateau above 10 mM, which means that at high theophylline concentrations the conformational change from $RNA^{inacitve}$ to RNA^{acitve} becomes rate limiting. In contrast, at lower theophylline concentration the second step, the formation of the RNA−theophylline complex, is rate limiting, and binding appears to be first order in theophylline concentration (Figure 3C). This RNA−theophylline system is rather unusual because the kinetics for complex formation could be measured over a 10^5 range of theophylline concentrations, and the requirement for a two-step binding mechanism was only observed at the higher theophylline concentrations.

Previous NMR studies on the theophylline-binding RNA showed that the two helical stems, but not the theophylline-binding site, were stably formed in the absence of theophylline (*26*, *29*). These NMR spectra showed that base pairs and base triples in the theophylline-binding site, which are observed in the RNA−theophylline complex, are not stably formed without theophylline. The kinetic results here also support a model where the RNA forms an ensemble of conformations in the absence of theophylline. Thus both the NMR and kinetic data are consistent with a major structural rearrangement of the core prior to formation of the RNA−theophylline complex. The mechanism in eq 5 shows a single rate constant for structural rearrangement of $RNA^{inactive}$ to RNA^{active}; however, this rate constant, k_1, likely represents an average rate constant for the conversion of many different conformers to RNA^{active}.

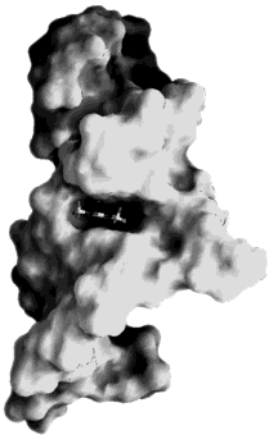

FIGURE 5: The binding pocket for theophylline in the RNA−theophylline complex determined by solution NMR is illustrated (*26*). The RNA is shown as a surface representation and theophylline as a stick model. The figure was produced with the GRASP program (*51*).

The rate constants for theophylline binding also indicate that RNA^{active} is not in an optimal conformation for theophylline binding. The apparent association rate constant k_2 for theophylline binding ($\sim 2 \times 10^5$ M^{-1} s^{-1}) is over 1000 times slower than diffusion-controlled binding (*7*). For an ideally preformed binding pocket, a fast association rate constant near the diffusion limit is predicted, and rates that approach diffusion control are seen in some protein−ligand interactions (*32*). The slow association rate constant k_2 suggests that the theophylline-binding pocket is not ideally formed in RNA^{acitve}, and conformational selection, desolvation of the RNA or the ligand, or an additional conformational rearrangement reduces the association rate constant from the diffusion-controlled rate (*32*).

Figure 5 shows the NMR solution structure of the RNA−theophylline complex with a surface representation of the RNA, and theophylline in a stick model, and emphasizes the large pocket required for theophylline binding (*26*). Theophylline stacks with two base triples that form the upper and lower face of the binding pocket (Figure 2B). Having such a large solvent-exposed surface area for these bases in the free RNA is energetically unfavorable; therefore, it is not surprising that the RNA changes conformation to optimize base stacking in the absence of the ligand. Thus optimizing stacking interactions between the bases, and reducing solvent exposure, likely drives the free RNA to a conformation different than in the complex. The fluorescence data also support this model. The 2-aminopurine shows high

fluorescence when it is exposed to solvent, but the fluorescence is efficiently quenched when stacking with other bases (*30*, *31*, *33*). The absence of 2-aminopurine fluorescence in the free RNA demonstrates that this base is not fully exposed to solvent in the free RNA and the RNA must undergo structural rearrangement to form the RNA−theophylline complex.

Many RNAs have been shown to undergo a conformational change upon ligand binding, and this has been referred to as conformational capture, induced fit, or ligand-induced folding (*2*, *22*−*24*). The free states of a number of in vitro selected RNAs are heterogeneous, and this may also be common for natural RNAs. For example, the U1A protein recognizes a hairpin motif in the 3′ UTR of its own messenger RNA with high affinity and specificity (*34*, *35*). NMR studies have shown that a functionally important internal loop in the RNA is disordered and highly flexible in the absence of U1A protein, and significant structural reorganization occurs in the RNA (and protein) upon complex formation (*35*, *36*). Although the recognition surface is much larger in the U1A system compared to the RNA−theophylline complex, the behavior of the free RNA is similar. The three-way junction of the S15 binding site in 16S ribosomal RNA is another example of an RNA that undergoes a conformational change upon ligand binding (*37*). A local structural rearrangement in the junction leads to a large change in the global conformation by reorientation of the three helices (*38*).

Magnesium Ions Stabilize the RNA−Theophylline Complex. NMR titration experiments previously showed a K_d of ~7 mM for theophylline binding in the absence of Mg^{2+} (*29*). This greater than 10000-fold decrease in affinity in the absence of Mg^{2+} demonstrates that Mg^{2+} plays a key role in stabilizing the RNA−theophylline complex. Interestingly, the NMR chemical shifts are very similar for the RNA−theophylline complex in the presence and absence of Mg^{2+}, indicating that the structure of the RNA−theophylline complex does not dramatically change with Mg^{2+} (*29*). Mn^{2+} ions can functionally replace Mg^{2+} for high-affinity theophylline binding, and Mn^{2+}-induced NMR line broadening data were used to identify a divalent metal ion binding site in the theophylline-binding RNA (*28*). This metal binding site is located on the outside of the S-turn formed between U22 and G26, and the divalent metal is in position to effectively stabilize the short phosphate−phosphate interactions in this sharp turn (Figure 1C). The kinetics for theophylline binding were measured in the presence and absence of Mg^{2+}, and the results show that most of the rate constants for theophylline binding are affected by Mg^{2+} (Table 1). A very large effect of Mg^{2+} is observed for the dissociation of the RNA−theophylline complex, with lifetimes ($1/k_{-2}$) of 14 s and ≤50 ms with and without Mg^{2+}, respectively. Misra and Draper have proposed a model for the role of Mg^{2+} in RNA folding where specific metal ions preferentially bind to the folded structure in RNAs (*39*, *40*). The NMR and kinetic data are consistent with this model where the metal binding site in the S-turn only forms in the RNA−theophylline complex. Thus the very long lifetime for the theophylline−RNA complex with Mg^{2+} suggests a specific role for Mg^{2+} in the complex, underscoring the importance of the metal binding site in the S-turn observed by NMR (*28*).

Mg^{2+} is critical for the correct folding and function of many RNAs (*39*−*42*). Mg^{2+} binding sites have been observed in crystal structures of a variety of RNAs including tRNA, P4−P6 of the *Tetrahymena* group I intron, the hammerhead ribozyme, and the hairpin ribozyme (*43*). For example, analysis of the S15 binding site of 16S rRNA indicated that Mg^{2+} stabilizes the folded state of the three-way junction (*37*), and recent studies on this system suggest that the rate of both folding and unfolding is accelerated by Mg^{2+} (*44*). In the S15 system the primary effect of Mg^{2+} is on the association rate constant, whereas in the theophylline-binding RNA the largest effect of Mg^{2+} is on the dissociation rate constant (Table 1), and both of these mechanisms lead to a more stable complex in the presence of Mg^{2+}.

Implications of a Heterogeneous Free State for Ligand Binding. Many enzymes exhibit a highly structured prefolded substrate binding site, and the entropic cost for ordering the polypeptide chain is paid in folding of the free state (*45*). The favorable interactions gained upon substrate binding then contribute to the affinity for the ligand. However, an increasing number of examples of association between more flexible molecules are being characterized (*45*−*47*). Such a mechanism may be a common feature of molecular recognition of RNAs. Many RNAs, such as HIV TAR, HIV RRE, the S15 binding region of 16 S rRNA, and many aptamers, form highly ordered structures in the presence of their ligand; however, the free RNA exists as an ensemble of conformations (*2*, *22*, *23*). The bulk of the secondary structure elements are formed without the ligand, but the conformations of the loops, bulges, or tertiary structures observed in the complexes are not present in the free RNAs.

One major difference between proteins and RNA is the stability of their secondary structures. Proteins gain their stability from the cooperative formation of the tertiary structure, and the individual α-helices and β-sheets are usually not stable on their own (*6*, *7*). In contrast, RNA secondary structure elements display very high stability, where even small hairpins stably form (*4*−*6*, *49*). This leads to a very hierarchical folding mechanism, with the secondary structure being formed first, followed by addition of tertiary contacts (*5*, *49*). Whereas globular protein domains often fold by an all-or-none process, RNAs can form stable secondary structure without committing to a unique tertiary fold (*6*).

Another characteristic of RNA is the aromatic character of the nucleic acid bases, which often leads to different stacking interactions in the free and ligand-bound state. RNA is built from large rigid bases; thus, to reposition a base may require extensive concerted motions of the sugar−phosphate backbone. A free state with a high degree of conformational freedom offers many different starting conformations, and the ligand selects conformations that are amenable to binding. Having a heterogeneous population of RNA conformers in the free state may actually represent an efficient mechanism for RNA−ligand binding, if the ligand binding site is not stable in the absence of the ligand.

The theophylline−RNA aptamer interaction offers an ideal system to study the effect of a heterogeneous free RNA on the formation of an RNA−ligand complex. Since theophylline is a rigid aromatic structure, the RNA must conform to the theophylline structure to provide a binding site. The binding pocket for theophylline is formed by eight nucleotides, and most of the intermolecular contacts involve

stacking and hydrogen-bonding interactions with the bases on the RNA (*26*). Formation of this unique complex requires unusual conformations of the RNA backbone, such as the S-turn in Figure 1C. Both NMR and kinetic data show that this RNA proceeds from a heterogeneous free state to a highly ordered complex upon theophylline binding, where the high degree of conformational freedom in the free RNA allows for exploration of optimal stacking and hydrogen-bonding interactions in the complex. Thus this RNA−theophylline system represents another example of what appears to be a common feature of RNAs, where the free RNA exists in an ensemble of conformational states and a stable conformation is only formed upon ligand binding.

ACKNOWLEDGMENT

We thank Drs. Susy Kohout for help with the stopped-flow instrument, Robert Batey and Jamie Williamson for useful discussions, and Grant Zimmermann for providing the GRASP figure.

SUPPORTING INFORMATION AVAILABLE

The Fortran program used for simulations and one figure with simulated and experimental data for complex formation at high theophylline concentration. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Uhlenbeck, O. C. (1995) *RNA 1*, 4−6.
2. Williamson, J. R. (2000) *Nat. Struct. Biol. 7*, 834−837.
3. Woodson, S. A. (2000) *Nat. Struct. Biol. 7*, 349−352.
4. Brion, P., and Westhof, E. (1997) *Annu. Rev. Biophys. Biomol. Struct. 26*, 113−137.
5. Tinoco, I., and Bustamante, C. (1999) *J. Mol. Biol. 293*, 271−281.
6. Draper, D. E. (1996) *Trends Biochem. Sci. 21*, 145−149.
7. Creighton, T. E. (1993) *Proteins, Structures and Molecular Principles*, 2nd ed., Freeman & Co., New York.
8. Rook, M. S., Treiber, D. K., and Williamson, J. R. (1998) *J. Mol. Biol. 281*, 609−620.
9. Thirumalai, D., and Woodson, S. A. (2000) *RNA 6*, 790−794.
10. Woodson, S. A. (2000) *Cell. Mol. Life Sci. 57*, 796−808.
11. Walter, N. G., Harris, D. A., Pereira, M. J. B., and Rueda, D. (2001) *Biopolymers 61*, 224−242.
12. Wang, S. L., Karbstein, K., Peracchi, A., Beigelman, L., and Herschlag, D. (1999) *Biochemistry 38*, 14363−14378.
13. Blount, K. F., and Uhlenbeck, O. C. (2002) *Biochemistry 41*, 6834−6841.
14. Hoogstraten, C. G., Wank, J. R., and Pardi, A. (2000) *Biochemistry 39*, 9951−9958.
15. Wedekind, J. E., and McKay, D. B. (1998) *Annu. Rev. Biophys. Bioeng. 27*, 475−502.
16. Wedekind, J. E., and McKay, D. B. (1999) *Nat. Struct. Biol. 6*, 261−268.
17. Treiber, D. K., and Williamson, J. R. (1999) *Curr. Opin. Struct. Biol. 9*, 339−345.
18. Zhuang, X. W., Bartley, L. E., Babcock, H. P., Russell, R., Ha, T. J., Herschlag, D., and Chu, S. (2000) *Science 288*, 2048−2051.
19. Butcher, S. E. (2001) *Curr. Opin. Struct. Biol. 11*, 315−320.
20. Lafontaine, D. A., Norman, D. G., and Lilley, D. M. J. (2001) *EMBO J. 20*, 1415−1424.
21. Walter, N. G. (2001) *Methods 25*, 19−30.
22. Frankel, A. D., and Smith, C. A. (1998) *Cell 92*, 149−151.
23. Leulliot, N., and Varani, G. (2001) *Biochemistry 40*, 7947−7956.
24. Weeks, K. M. (1997) *Curr. Opin. Struct. Biol. 7*, 336−342.
25. Jenison, R. D., Gill, S. C., Pardi, A., and Polisky, B. (1994) *Science 263*, 1425−1429.
26. Zimmermann, G. R., Jenison, R. D., Wick, C. L., Simorre, J.-P., and Pardi, A. (1997) *Nat. Struct. Biol. 4*, 644−649.
27. Zimmermann, G. R., Shields, T. P., Jenison, R. D., Wick, C. L., and Pardi, A. (1998) *Biochemistry 37*, 9186−9192.
28. Zimmermann, G. R., Wick, C. L., Shields, T. P., Jenison, R. D., and Pardi, A. (2000) *RNA 6*, 659−667.
29. Zimmermann, G. R. (1997) Ph.D. Thesis, Department of Chemistry, University of Colorado, Boulder, p 201.
30. Millar, D. P. (1996) *Curr. Opin. Struct. Biol. 6*, 322−326.
31. Jean, J. M., and Hall, K. B. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 37−41.
32. Fersht, A. (1984) *Enzyme Structure and Mechanism*, Freeman & Co., New York.
33. Rachofsky, E. L., Osman, R., and Ross, J. B. A. (2001) *Biochemistry 40*, 946−956.
34. Allain, F. H. T., Howe, P. W. A., Neuhaus, D., and Varani, G. (1997) *EMBO J. 16*, 5764−5774.
35. Kranz, J. K., and Hall, K. B. (1999) *J. Mol. Biol. 285*, 215−231.
36. Allain, F. H. T., Gubser, C. C., Howe, P. W. A., Nagai, K., Neuhaus, D., and Varani, G. (1996) *Nature 380*, 646−650.
37. Batey, R. T., and Williamson, J. R. (1998) *RNA 4*, 984−997.
38. Orr, J. W., Hagerman, P. J., and Williamson, J. R. (1998) *J. Mol. Biol. 275*, 453−464.
39. Misra, V. K., and Draper, D. E. (2002) *J. Mol. Biol. 317*, 507−521.
40. Misra, V. K., and Draper, D. E. (1998) *Biopolymers 48*, 113−135.
41. Pyle, A. M. (1996) *Met. Ions Biol. Syst. 32*, 479−520.
42. Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) *RNA 8*, 307−323.
43. Feig, A. L., and Uhlenbeck, O. C. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 287−319, Cold Spring Harbor Laboratory, Plainview, NY.
44. Kim, H. D., Nienhaus, G. U., Ha, T., Orr, J. W., Williamson, J. R., and Chu, S. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 4284−4289.
45. Schreiber, G. (2002) *Curr. Opin. Struct. Biol. 12*, 41−47.
46. Dyson, H. J., and Wright, P. E. (2002) *Curr. Opin. Struct. Biol. 12*, 54−60.
47. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) *Biochemistry 41*, 6573−6582.
48. Wikstrom, A., Berglund, H., Hambraeus, C., van den Berg, S., and Hard, T. (1999) *J. Mol. Biol. 289*, 963−979.
49. Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr. (2000) *Nucleic Acids: Structures, Properties and Function*, University Science Books, Sausalito, CA.
50. Carson, M. (1997) *Methods Enzymol. 277*, 493−505.
51. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.

BI027103+