MD simulations of ligand-bound and ligand-free aptamer: Molecular level insights into the binding and switching mechanism of the *add* A-riboswitch

MONIKA SHARMA,[1] GOPALAKRISHNAN BULUSU,[1,2] and ABHIJIT MITRA[1]

[1]Center for Computational Natural Sciences and Bioinformatics (CCNSB), International Institute of Information Technology (IIIT-H), Gachibowli, Hyderabad-500032, India
[2]Life Sciences R&D, TCS Innovation Labs Hyderabad, Tata Consultancy Services Limited, Madhapur, Hyderabad 500081, India

ABSTRACT

Riboswitches are structural *cis*-acting genetic regulatory elements in 5′ UTRs of mRNAs, consisting of an aptamer domain that regulates the behavior of an expression platform in response to its recognition of, and binding to, specific ligands. While our understanding of the ligand-bound structure of the aptamer domain of the adenine riboswitches is based on crystal structure data and is well characterized, understanding of the structure and dynamics of the ligand-free aptamer is limited to indirect inferences from physicochemical probing experiments. Here we report the results of 15-nsec-long explicit-solvent molecular dynamics simulations of the *add* A-riboswitch crystal structure (1Y26), both in the adenine-bound (CLOSED) state and in the adenine-free (OPEN) state. Root-mean-square deviation, root-mean-square fluctuation, dynamic cross-correlation, and backbone torsion angle analyses are carried out on the two trajectories. These, along with solvent accessible surface area analysis of the two average structures, are benchmarked against available experimental data and are shown to constitute the basis for obtaining reliable insights into the molecular level details of the binding and switching mechanism. Our analysis reveals the interaction network responsible for, and conformational changes associated with, the communication between the binding pocket and the expression platform. It further highlights the significance of a, hitherto unreported, noncanonical W:H *trans* base pairing between A73 and A24, in the OPEN state, and also helps us to propose a possibly crucial role of U51 in the context of ligand binding and ligand discrimination.

Keywords: molecular dynamics; purine riboswitch; aptamer; dynamic cross-correlation; ligand discrimination; molecular mechanism

INTRODUCTION

Riboswitches are *cis*-acting regulatory elements found in 5′ untranslated regions (UTRs) of many mRNAs of bacteria and fungi (Mandal and Breaker 2004a; Nudler and Mironov 2004; Soukup and Soukup 2004). Several riboswitches, specific to thiamine pyrophosphate (TPP) (Mironov et al. 2002; Winkler et al. 2002a), guanine (Mandal et al. 2003), adenine (Mandal et al. 2003; Mandal and Breaker 2004b), coenzyme B12 (Nahvi et al. 2002), flavin mononucleotide (FMN) (Mironov et al. 2002; Winkler et al. 2002b), glucosamine-6-phosphate (GlcN6P) (Winkler et al. 2004), glycine (Mandal et al. 2004), lysine (Grundy et al. 2003; Rodionov et al. 2003; Sudarsan et al. 2003), and S-adenosylmethionine (SAM) (Epshtein et al. 2003; McDaniel et al. 2003; Winkler et al. 2003; Fuchs et al. 2006; Gilbert et al. 2008), queuosine precursor, 7-aminomethyl 7-deazaguanine (pre-Q$_1$) (Roth et al. 2007), and Mg (Cromie et al. 2006; Dann et al. 2007) have been identified until now.

All known riboswitches are highly structured RNA elements that are composed of two distinct yet mutually interacting domains: a "ligand-binding" or aptamer domain, and an expression platform involved in gene regulation. Conformational changes due to ligand binding affect the folding patterns in the expression platform, enabling riboswitches to exercise ligand-dependent control on downstream gene expression for enzymes involved in the metabolism of the ligand (Serganov et al. 2004; Gilbert et al. 2008; Montange and Batey 2008).

Riboswitches distinguish between closely related ligands using different strategies such as complementarity of bases, as in purine riboswitches (Serganov et al. 2004), and the length of ligand as in TPP and SAM riboswitches (Montange and Batey 2006; Serganov et al. 2006; Thore et al. 2006). Although adaptive recognition strategies of nucleic acid aptamers have been studied for a long time (Hermann and Patel 2000), naturally occurring RNA aptamer domains attached to riboswitches have brought in newer dimensions (Buskirk et al. 2004; Lea and Piccirilli 2007)

In recent years, purine (guanine and adenine) riboswitches have attracted great attention from structural biologists and biophysicists alike (Batey et al. 2004; Serganov et al. 2004; Noeske et al. 2005, 2007; Wickiser et al. 2005b; Gilbert et al. 2006; Lemay et al. 2006; Rieder et al. 2007; Greenleaf et al. 2008; Lin and Thirumalai 2008), partly because they are among the simplest riboswitches, which contain the phylogenetically most conserved aptamer domains reported to date. Another compelling reason is that, in spite of high structural similarity of their aptamer domain, they are remarkable for their ligand selectivity and for the variety in their strategy for gene expression control (Mandal et al. 2003; Mandal and Breaker 2004a). On ligand binding, the *xpt* G-riboswitch operates by switching *OFF* gene expression by *transcriptional deactivation* (Mandal et al. 2003). In contrast, the *pbu*E (*ydhL)* A-riboswitch turns gene expression *ON* by *transcriptional activation* (Mandal and Breaker 2004b). Yet another control strategy is employed by the *add* A-riboswitch, the subject of the current investigation, which on ligand binding switches gene expression *ON* by *translational activation* (Serganov et al. 2004; Rieder et al. 2007). Thus, while the ligand-bound CLOSED state of the aptamer domains of these purine riboswitches are structurally very similar, ligand binding leads to the stabilization of a transcription terminator in the first case and to the stabilization of a transcription antiterminator in the second case. (In this work, we refer to the ligand-bound state of the aptamer as the CLOSED state and the ensemble of possible conformations of the ligand-free state as the OPEN state.) In the third case, the CLOSED state aptamer domain assists in unmasking of the Shine–Dalgarno sequence in the expression platform, and therefore helps in activating translation. In general, the transcription regulatory riboswitches involve longer stretches from the aptamer domain and function through primarily a kinetically controlled mechanism, while the translation regulatory *add* A-riboswitch involves a shorter stretch and operates through thermodynamic control (Rieder et al. 2007).

The Leontis–Westhof nomenclature (Leontis and Westhof 2001) is used here for naming base pairs using the following format: <base pair1><symbol(•/∘)><basepair2> <Edge1(W/H/S)>:<Edge2(W/H/S)> <Glycosidic bond orientation (cis/trans)>, where • denotes *cis* orientation; ∘ denotes *trans* orientation; and W, H, and S denote Watson, Hoogsteen, and sugar edges, respectively.

The crystal structure of adenine riboswitch reveals an intricate architecture common to the general class of purine riboswitches

Despite having different mechanisms for switching, crystal structures (Serganov et al. 2004; Gilbert et al. 2006, 2007; Lemay et al. 2006; Edwards and Batey 2009) of ligand-bound aptamer domain of natural purine riboswitches show remarkably conserved organization of different structural elements. All of them consist of three helices, P1, P2, and P3, forming a three-way junction, with the junction loops J1-2, J2-3, and J3-1 connecting them and with loops L2 and L3 capping P2 and P3, respectively (Fig. 1). In the *add* A-riboswitch, 1Y26 (Serganov et al. 2004), the junction loops J1-2 (three nucleotides: 22–24) and J3-1 (two nucleotides: 73 and 74) are smaller than J2-3 (eight nucleotides: 46–53). Loops L2 and L3 are each 7-nucleotides (nt) long: 32–38 and 60–66, respectively.

The complete binding pocket of the adenine bound *add* A-riboswitch, formed through stacked base triplets interconnecting the three junction loops J1-2, J2-3, and J3-1, with two base pairs of P1 (Fig. 1A), can be visualized in terms of three blocks:

1. A "roof" block, comprising of a triplet involving a water-mediated hydrogen bond between the canonical pair, U22–A52 and A73 (RT1), and a second triplet, A23∘G46–C53 (RT2) consisting of a canonical G46–C53 base pair and a noncanonical A23∘G46 S:W *trans* base pair, above it; whereas RT1 links together one base each from all the three junction loops, RT2 links J1-2 with J2-3.
2. A "floor" block, consisting of a triplet A21–U75•C50 (FT1) and a second triplet U20–A76•U49 (FT2) below it. The floor comprises of interactions between elements of helix P1 (U20, A21, U75, A76) and junction J2-3 (U49, C50), through two canonical base pairings, A21–U75 and U20–A76, forming triplets involving two noncanonical base pairs, U75•C50 S:W *cis* and A76•U49 S:W *cis*, respectively.
3. The "ligand-binding" block, consisting of elements of the three junction loops J1-2 (U22), J2-3 (U51 and U47), and J3-1 (U74). The major ligand-binding interactions with adenine are provided by U51, through what may be described as Ade∘U51 S:W *trans* interaction, and the discriminating host U74, which binds with adenine in a canonical manner. In addition, there are auxiliary binding interactions involving O2′ of U22, hydrogen bonding with N7 of the adenine, and O2 of U47, forming a bifurcated hydrogen bond with C8 and N9 of adenine. These additional interactions, though discounted earlier (Gilbert et al. 2006; Lemay et al. 2006), have recently been validated through ab initio quantum mechanical studies (Sharma et al. 2009).

The elements of the binding pocket are organized through the intermediacy of two more blocks:
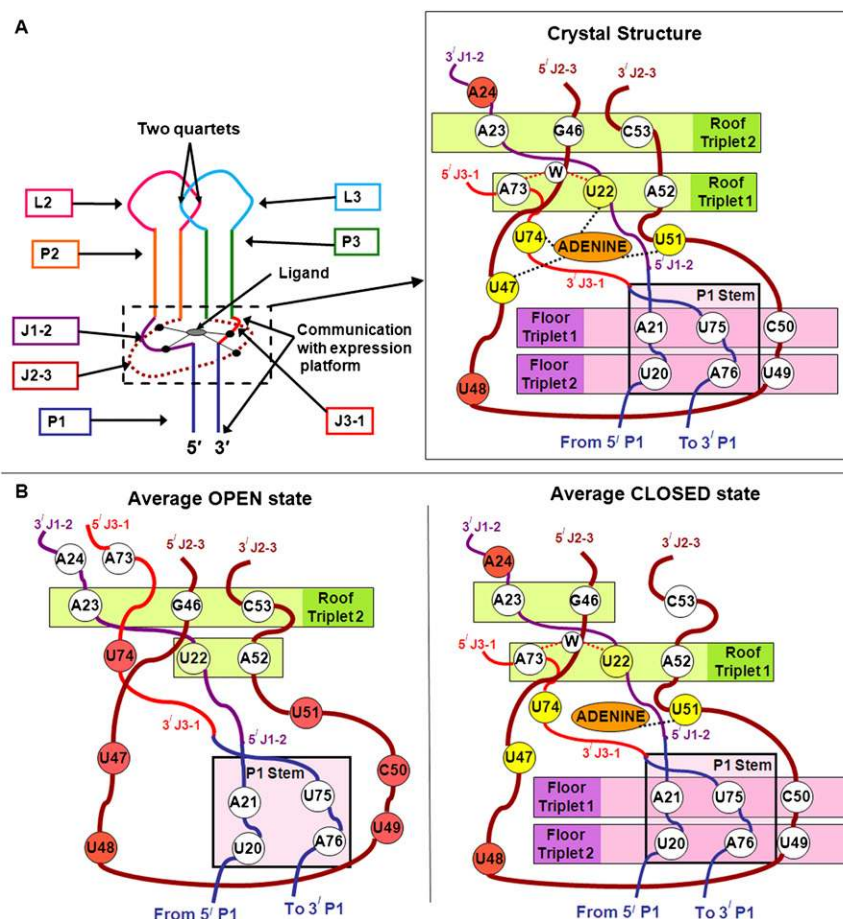
FIGURE 1. Schematic representation of the (*A*) crystal structure of aptamer domain of adenine riboswitch, with ligand-binding domain; (*B*) ligand-binding domain of average OPEN state structure and CLOSED state structure.

them. U62, the third unpaired L3 base, is also stacked out.

2. The "floor" triplets merging into the P1 stem constitutes the "Switching" block, since the 3′ strand of the P1 stem is involved in alternative base pairing with the Shine–Dalgarno sequence of the expression platform, in the absence of adenine binding (Mandal and Breaker 2004a,b; Nudler and Mironov 2004; Serganov et al. 2004; Soukup and Soukup 2004).

The Story So Far …

Detailed analysis of the crystal structures supplemented by comparative binding and activity analysis (Gilbert et al. 2006, 2007; Lemay et al. 2006; Lemay and Lafontaine 2007; Edwards and Batey 2009) of wild-type and site-directed mutants, of different purine riboswitches, have helped characterize the phylogenetically conserved nucleotides in terms of those that are important for recognition, discrimination, and binding of the ligand and differentiate them from those that are important only for the switching process (Lemay and Lafontaine 2007; Mulhbacher and Lafontaine 2007). Thus, in *add* A-riboswitch, it is fairly clear that the conservation pattern of the P1 nucleotides between position 77 and position 83 relates to the formation of the ligand-free OPEN state fold. They are not associated with the ligand-binding potential of the aptamer (Rieder et al. 2007). The emerging consensus regarding the essential binding related features may be listed as follows:

1. Y74 is the discriminating nucleotide for purine riboswitches.
2. Ligand is completely encapsulated by the aptamer in the CLOSED state.
3. The binding process is slow, slower than expected for a diffusion controlled association.
4. The aptamer is globally organized in terms of a structured and preformed dome domain and a locally disorganized binding pocket domain in the ligand-free OPEN state. On ligand binding, in particular, the five nucleotides 47–51 in the J2-3 loop, switches from a flexible OPEN state conformation to a rigid organization in the CLOSED state.

1. A "dome" over the "roof" forms a fourth block, comprising of the stem–loops P2–L2 and P3–L3; held together through kissing loop interactions between L2 and L3. The kissing interactions involve nucleotides A33, U34, G37 and G38 of L2 and C60, C61, A65, and A66 of L3, giving rise to two base quadruplet motifs as follows: (1) The first is comprised of two interloop base pairs, a canonical G37–C61, and a noncanonical base pair, U34∘A65 W:H *trans*. The quadruplet motif arises out of two noncanonical, intraloop auxiliary base pairings: C61•A65 S:W *cis* and U34•G37 S:S *cis*. (2) The interloop base pairings constituting the second quadruplet motif is centered around the noncanonical G38•A66 S:W *cis* pair with each of the bases respectively participating in the canonical G38–C60 and noncanonical A33•A66 W:H *cis* interactions. Of the three unpaired bases in L2, U32 nicely stacks over the noncanonical U31•U39 W:W *cis* loop closing pair of P2 and U36 and is completely flipped out. The third unpaired L2 base, A35, is aligned in a H:H *trans* manner with unpaired A64 of L3 and has A63, also from L3, stacking over

5. Binding leads to the pairing of four of these five nucleotides and links ligand binding with the stabilization of the P1 stem, which in effect, constitutes the switch.

6. The preformed dome domain, with kissing loop interactions between L2 and L3, can exist independent of the remaining part of the aptamer.

7. A minimal binding domain, with the L2–L3 interaction severely reduced or completely absent, can also function, albeit with significantly lower efficiency (Batey et al. 2004). The L2–L3 interaction possibly assists a side-by-side placement of the helices P2 and P3 and helps in the coaxial stacking of P1 and P3, thereby providing stability to a "binding-ready" conformation in the OPEN state and thus improve the binding kinetics (Stoddard et al. 2008).

8. The Mg^{2+} ions appear to be important for preorganization of the aptamer fold and for ligand binding. Recent NMR studies suggest that Mg^{2+} ions are necessary to force molecules from a conformational pool of the "free" aptamers into a single binding competent conformer (Ottink et al. 2007).

However, despite several recent experimental and theoretical studies directed toward gaining an understanding of the ligand-free OPEN state of the aptamer domain, as well as of the intermediate stages, during the course of binding and switching (Eskandari et al. 2007; Rieder et al. 2007; Greenleaf et al. 2008; Lin and Thirumalai 2008; Edwards and Batey 2009), there still exist several open questions that have not been unequivocally answered.

1. *What is the mechanism for ligand discrimination?* Purine riboswitches can be made to change their ligand specificity merely by changing the identity of Y74 (Mandal and Breaker 2004a)—but how? It cannot merely be based on, as is widely implicated, high interaction energy of canonical base pairing, since quantum mechanical studies show that a W:W *cis* base pair has higher interaction energy for G:U than for A:U. Also, for A:U, the Watson–Crick edge of uracil interacts, in *cis* orientation, much better with the sugar edge of adenine than with the Watson–Crick edge (Bhattacharyya et al. 2007).

2. *What is the extent of "disorganization" of the binding pocket vis-à-vis "preorganization" of the aptamer?* The free energy landscape of binding is defined not only by the binding enthalpy but also by the extent of entropic penalty, which is reduced in cases involving extensive preorganization of the binding competent conformation. At the same time, a high degree of preorganization would sterically impede the access of the ligand to the binding pocket.

3. *How many conformational states are involved in the binding and switching process?* How many conformations does the aptamer fold visit in the OPEN state? Is there such a state as a "binding competent" state? What role do Mg^{2+} ions play in this regard? The energetics of binding is not only determined by the interactions in the bound state but also by the energetic penalties associated with conformational changes and folding transition. How high are the energy barriers for going from "binding competent" to fully folded state? Does binding take place by conformation selection and association (low energy of activation to binding) or by induced fit (higher energy of activation to binding).

4. *How flexible is the CLOSED state?* The PDB records of 1Y26 show reduced B factors for the binding pocket nucleotides, as well as for the bases involved in the loop–loop interactions. The B factors are distinctly high for the unpaired residues in L2 and L3 and for P1 residues away from the binding pocket. While these observations provide some clues, a detailed assessment of the flexibility in the bound state can help understand enthalpy–entropy compensation and the entropic contributions to the free energy in the bound state.

5. *What is the initial ligand docking site?* It has been suggested that the binding process involves two steps: an initial low affinity association of the ligand with some "nonrandom" site of the aptamer, followed by slow binding with reorganization of the binding pocket (Buck et al. 2007). Whether Y74 itself is the initial docking site or is there another base where the ligand initially docks, is still an open question.

6. *What is the mechanism of the slow binding kinetics? Can it unequivocally explain available kinetic data?* Earlier studies on FMN and adenine-sensing riboswitches suggest that the ligand-binding rate is several orders of magnitude slower than what is expected from diffusion controlled association (Wickiser et al. 2005a,b). Also, assessment of the kinetics of binding by stopped-flow fluorescence measurements, using 2-aminopurine (2AP) as a reporter ligand for the C74U mutant, and 7-deazaguanine as a ligand for a U48–2AP reporter mutant of a G-riboswitch appear to rule out several possible binding mechanisms including a two-state model and one in which a bimolecular ligand association step is followed by a unimolecular folding step (Gilbert et al. 2006). Instead, the experimental data has been interpreted to support a two-step mechanism in which the ligand binds to riboswitch RNA that exists as multiple slowly interconverting conformers, of which only one is binding competent. Implicit in the formulation of this model are the following assumptions: involvement of a slow bimolecular step is necessary for explaining a second-order kinetics; U74, being the discriminatory base, is also involved in the bimolecular step, and that a slower unimolecular step precedes the bimolecular step. However, none of the above assumptions have unequivocal experimental basis.

It is worthwhile to summarize the models of the "binding-competent" state here.

1. Single molecule FRET studies reveal that the folding pathway includes a discrete intermediate state (Lemay et al. 2006) at low Mg^{2+} concentrations. A similar suggestion was made based on association kinetics observed by Gilbert et al. (2006), who hypothesized that this intermediate may have the stems P1, P2, and P3 formed, L2 and L3 interact, and P1 and P3 stacked.

2. Distinct from FRET studies on single labeled RNA molecules (Lemay et al. 2006), ligand fluorescence studies on ensemble solutions (Eskandari et al. 2007) suggest the possibility of different distributions of local and global conformers for adenine and guanine riboswitches, respectively.

3. Noeske et al. (2007), based on multinuclear NMR spectroscopic studies, observed that, in the absence of Mg^{2+}, the ligand-free state of the RNA is conformationally heterogeneous, with the pool consisting largely of conformations with base-pairing patterns detrimental to ligand binding. These conformations become partially preorganized for ligand binding in the presence of Mg^{2+}. Similarly, Ottink et al. (2007) argue for a more structured aptamer domain in the binding-competent state and assign disorder only to five nt of J2-3. The experimental evidence in the paper corroborates their hypothesis that Mg^{2+} ions are required to drive several alternative conformations to a single stable binding efficient conformer. However, their assertion about the preexistence of both the roof triplets [A73(w)U22–A52 and A23∘G46–C53], based on data of only three nucleotides (U22, G46, and A52) may be accepted with caution.

4. Rieder et al. (2007) followed changes in fluorescence behavior of selected nucleotides in the binding pocket region as well as in the L2–L3 region to infer Mg^{2+} ion-induced preorganization and binding-induced tightening of the aptamer fold in both the regions. In addition, they highlight a possible role for A24 in this process.

5. Greenleaf et al. (2008) using single-molecule force spectroscopy and dual-trap optical tweezers technology, and Lin and Thirumalai (2008), using self-organized polymer model by Langevin dynamics in the overdamped limit, have carried out unfolding–refolding studies of adenine aptamers to infer the nature of their binding energy landscape. Both studies confirm the sequence of folding states as formation of P2/P3 followed by P1, implying that it is easiest to unfold P1. The experimental study suggests that the binding competent form may consist of several conformers, ranging from those with P1 unfolded to those resembling the fully folded aptamer; that the completely folded state is much stabler in the presence of adenine than in its absence, and that the binding-competent

form follows a higher activation energy pathway to convert into a far more stable fully folded form in the presence of adenine. However, these studies have been carried out in the absence of Mg^{2+} ions and may not be reliable for understanding in vivo activity.

It is generally understood that time-resolved NMR and single-molecule techniques, referred to above, are best suited for studying millisecond-scale processes in RNA dynamics (Al-Hashimi and Walter 2008). For faster processes, implicated in RNA folding dynamics in the picosecond and nanosecond scale, molecular dynamics studies can complement magnetic field-induced residual dipole coupling studies (mRDC) (Ottink et al. 2007). In order to address some of the questions highlighted above, we have therefore carried out molecular dynamics simulation studies on the aptamer domain of the *add* A-riboswitch, both in the ligand-free OPEN state as well as in the ligand-bound CLOSED state.

RESULTS AND DISCUSSION

General observations from trajectory analyses

Root-mean-square-deviation (RMSD) analysis of the two trajectories

The comparative plot of the RMSD versus time (Fig. 2) and the comparative average RMSD values (Table 1) reveal several features related to the structure and dynamics of the CLOSED and OPEN states.

1. Both the CLOSED and OPEN states are flexible, and as expected, the average RMSD value of the CLOSED state (RMSD 2.99 Å) is lower than that of the OPEN state (RMSD 4.12 Å), where the ligand, and its interactions with the binding pocket, is missing.

2. A large increase was observed in the RMSD values for the OPEN state in the initial 3 nsec of the simulation (Fig. 2A) and the values on the average, not only remain higher than that in the CLOSED state, the fluctuations in the RMSD values are also much higher in the OPEN state ($\sigma = 0.65$) than in the CLOSED state ($\sigma = 0.30$).

3. Though the average RMSD for the CLOSED state trajectory is significantly lower than that of the OPEN state, the value of 2.99 Å is large enough to indicate significant residual flexibility of the bound state.

4. When the RMSD plots corresponding to the CLOSED and OPEN state trajectories are superposed with respect to stretches 25 to 45 and 56 to 70, belonging to the roof and dome part of the aptamer fold, the regions above the bounding residues show an average value of RMSD of 1.99 Å ($\sigma = 0.14$) for the OPEN state and 2.37 Å ($\sigma = 0.16$) for the CLOSED state, while the remaining parts
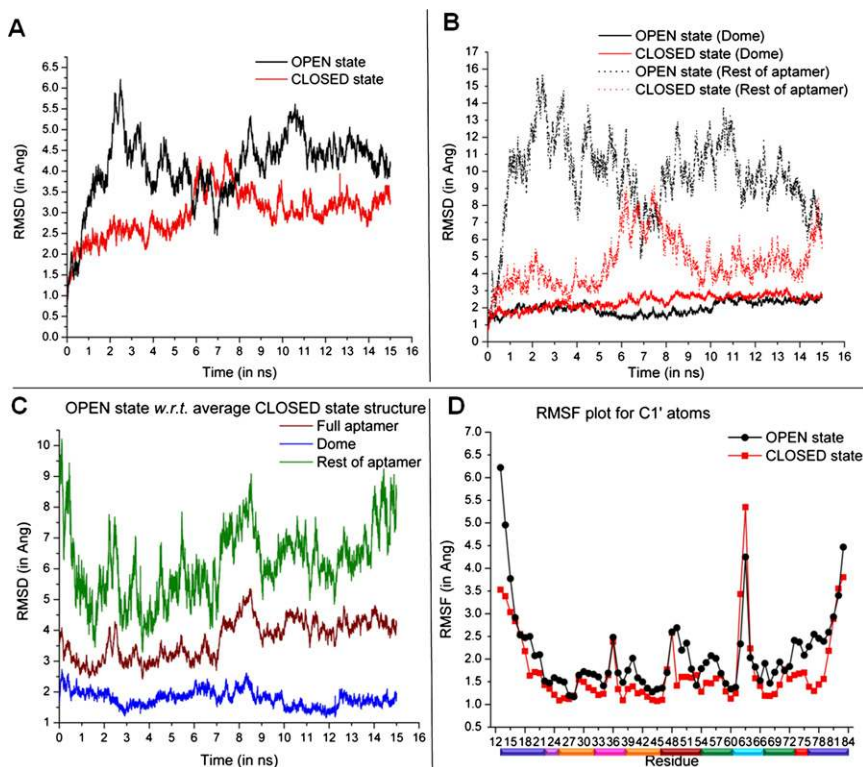
FIGURE 2. RMSD plots for (*A*) trajectory of OPEN and CLOSED state with respect to their initial structure; (*B*) trajectory of OPEN and CLOSED state for the dome region (stretches: 25–45 and 56–70) shown in solid line, and for the rest of the aptamer domain is shown in dashed line; (*C*) trajectory for OPEN state with respect to the average CLOSED state structure (full aptamer, dome region, and rest of the region shown in different colors); and (*D*) RMSF for the C1′ atoms of the nucleotides averaged over the time course of simulation. Structural regions are indicated by the color bar as per the scheme used in Figure 1A.

of the P3 helix, and the three junction loops forming the binding pocket. This is also apparent from the comparative RMSD plots for the two stretches (Fig. 2B).

Root-mean-square-fluctuation (RMSF) analysis of C1′ atoms of individual nucleotides

It can be seen from Figure 2D that, except for U62 and U63, all nucleotides have correspondingly lower RMSF values in the CLOSED state. In the CLOSED state, not only do the bases in the P1 helix region, particularly near the "binding pocket" end, but also the binding pocket nucleotides, defined by the three junction loops, show reduced RMS fluctuations. A causal relationship between these two observations is clearly indicated by the marked reduction in the RMSF values of ligand-bound U51 and of its neighbors C50 and U49, which are involved in the formation of the "floor" triplets that sequester the P1 base pairs: U75–A21 and A76–U20. These observations are in tune with chemical probing experiments that, for most J1-2 and J2-3 nucleotides, report extensive modification in the OPEN state and lack of it in the CLOSED state (Stoddard et al. 2008).

show an average value of RMSD of 9.79 Å (σ = 5.45) and 4.64 Å (σ = 2.09) for OPEN and CLOSED states. This suggests that major contributions toward increased OPEN state RMSD values come from the loosening up of the P1 helix, the two junction proximal base pairs

Similarly, the nucleotides U36, U48, U62, and U63, which remain flipped out and show markedly higher RMSF values than their neighbors during the CLOSED state simulations, have also been found to be flexible in experimental studies (Ottink et al. 2007; Stoddard et al. 2008).

TABLE 1. Average and standard deviation for various RMSD values

Reference structure	Computed state	Observed stretch of nucleotides	Mean (Å)	σ (Å)
Minimized structure from 1Y26				
	CLOSED	Complete aptamer	2.99	0.30
	CLOSED	Roof/dome "Residues 25–45 and 56–70"	2.37	0.16
	CLOSED	"Residues other than 25–45 and 56–70"	4.64	2.09
Minimized structure with ligand removed from 1Y26				
	OPEN	Complete aptamer	4.12	0.65
	OPEN	Roof/dome "Residues 25–45 and 56–70"	1.99	0.14
	OPEN	"Residues other than 25–45 and 56–70"	9.79	5.45
Average structure from CLOSED state trajectory				
	OPEN	Complete aptamer	3.71	0.39
	OPEN	Roof/dome "Residues 25–45 and 56–70"	1.79	0.07
	OPEN	"Residues other than 25–45 and 56–70"	6.16	1.34

Comparative analysis of fluctuations in backbone torsion and chi angles

Here, we present some general observations related to fluctuations in the torsion angles δ, χ, and ζ during both CLOSED and OPEN state simulations (see Supplemental Information SI1).

In both the states, the δ value for most of the nucleotides, oscillates around 84°, and corresponds to a C3′-endo pucker (Murray et al. 2003). Only 12 nucleotides, six of which are unpaired in the crystal structure (A35 and U36 from L2, U62, U63, A64 from L3 and U48 from J2-3), showed δ values corresponding to C2′-endo pucker. Two other nucleotides, U34 and A65, are involved in one of the quartets participating in L2–L3 interactions. All the four remaining nucleotides (U22, U47, U49, and C53) are directly involved in binding to the adenine ligand and/or in forming the roof/floor triplets. It is interesting to note that the average δ value for C53, which closes the J2-3 loop by pairing with G46, corresponds to an A-RNA like C3′-endo pucker (Fig. 3) in the OPEN state and to a C2′-endo pucker in the CLOSED state. This could mean that, as is also indicated by dynamic cross-correlation (DCC) analysis discussed later, C53 was better paired with G46 in the OPEN state, and that changes induced by ligand binding, loosen the G46–C53 base pairing, and switch the corresponding sugar to the C2′-endo conformation.

Binding induced changes in base pairing patterns and stabilities are also decipherable from an analysis of fluctuations in χ in the two states. For example, the χ value for

FIGURE 3. Dial plots showing variation of torsion angles χ, ζ, and δ during the simulations. In the circular plots, the radial axis signifies time, with the origin as 0 psec of the production run and progressing outward, and the angular axis represents the value of torsion angles with the range spanning as 0 to 2π in anticlockwise direction with 0 being at the *right* side of the plot. The histograms plotted at the periphery of the circular plots denote the frequency of the values spanned by torsion angles.

U51, found to be highly variable in the OPEN state, becomes far more focused (Fig. 3) in the CLOSED state. The mean value also undergoes a change from around −130° to −150°. This corresponds to the inward movement of U51, unpaired and dynamic in the OPEN state, to join U74 in binding with the adenine ligand in the CLOSED state. Similarly, changes in the χ values of U49 and C50, suggest that the floor triplets are formed only as a consequence of ligand binding. Interestingly, all of the six χ values, associated with the P1 helix base pairs A19–U77, U20–U76, and A21–U75, show noticeable reduction in fluctuation in the CLOSED state. This supports the hypothesis that, in the presence of the ligand, the P1 helix becomes stabler toward the binding pocket end.

This hypothesis is further supported by the fact that ζ values for the P1 nucleotides A19–A21 are noticeably less fluctuating in the CLOSED state than in the OPEN state. It may be noted that the torsion angle ζ is considered to be the most significant component contributing toward the conformational variability observed in RNA molecules (Schneider et al. 2004), and that several nucleotides show a shift in their respective ζ values in the two states. While the rationalization of all the changes may not be a straightforward task, the change in the degree of variability and/or the mean value, for all the J2-3 nucleotides (46–53), binding pocket side nucleotides of P3 (54–56 and 69–71), and the J1-2 nucleotide (21) highlight the transition of a flexible binding pocket in the OPEN state to a more structured and stable binding pocket and its adjoining helices, in the CLOSED state (Fig. 3).
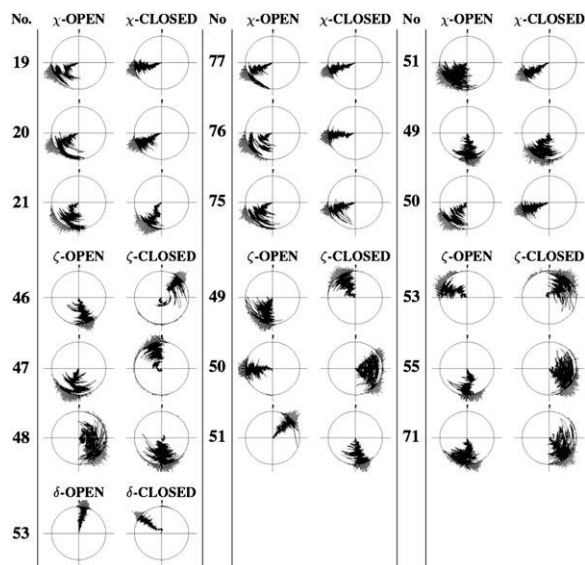
All versus all DCC analysis for pairs of bases in the CLOSED and OPEN states, respectively

The DCC maps for the two states, representing correlation coefficients (see Supplemental Information SI2) calculated as time average over the duration of the simulation, are shown in Figure 4. The whole range of correlation from +1 to −1 is represented by the variation in the color intensity of blue (positively correlated) through white (uncorrelated) to red (negatively correlated). The salient features of the maps are annotated as boxed regions 1 to 13 in the CLOSED (Fig. 4A) and OPEN state (Fig. 4B) maps.

Both of the maps show very similar patterns of blue islands indicating positively correlated fluctuations of nucleotides implicated in the formation of the aptamer fold. This supports the existence of a preformed aptamer fold in the OPEN state. A closer scrutiny however, brings out several distinctive features indicative of changes associated with ligand-binding and consequent stabilization of the aptamer fold:

1. The difference in correlation pattern in region 5 highlights the "CLOSED state" disruption in the intraloop correlation of the J2-3 nucleotides.
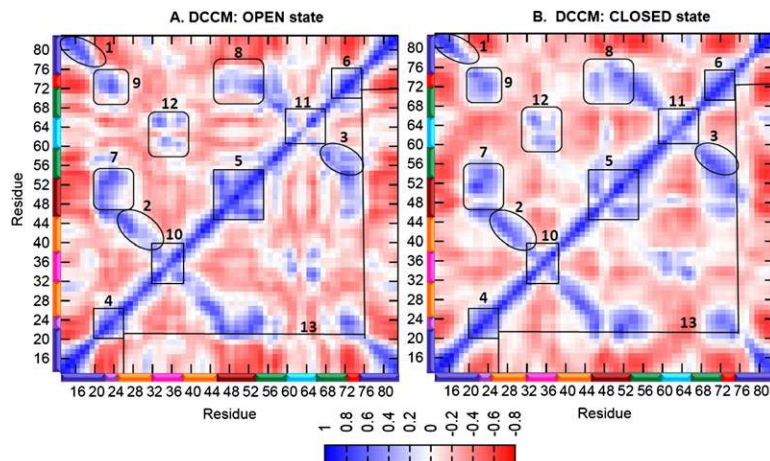
FIGURE 4. Dynamic cross-correlation maps calculated as time averaged for the centroid of base pairs over the simulation. Regions 1, 2, and 3 indicate the expected positive correlation between nucleotide pairs constituting the helices P1, P2, and P3, respectively. Regions 4, 5, and 6 indicate intrajunction self-correlations of J1-2, J2-3, and J3-1, respectively. Regions 7, 8, and 9 reveal the nature of inter junction correlations, namely, J1-2/J2-3, J2-3/J3-1, and J3-1/J1-2, respectively. Since the junction nucleotides are expected to move along with the neighboring nucleotides in their adjoining helices, the regions are marked in the figure by appropriately including portions of the latter. The correlation pattern in the intraloop nucleotide dynamics are indicated in the regions 10 (L2) and 11 (L3), while the L2–L3 interloop correlations are shown as region 12. Region 13 shows the correlation of P1 nucleotides with the rest of the aptamer fold. The color bar along the axes indicates the structural elements to which residues belong, with color coding as in Figure 1A.

2. The differences seen in the regions 7, 8, and 9, indicate the difference in the manner in which the three junction loops interact with each other in the two states.

3. Regions 11 and 12, respectively defining the correlations within the loop L3 and between loops L2 and L3, highlight the contrasting behavior of U62 and U63 in the two states. Both of them show a relative lack of correlation with their L3 neighbors, and selective correlation with L2 partners, only in the OPEN state.

4. Of particular significance is region 13 (a large L-shaped box), which shows predominantly negative correlation of most of the P1 nucleotides with the rest of the aptamer fold. The reduction in the number of negatively correlated P1 nucleotides is indicative of the sequestration effect of ligand induced formation of the floor triplets in the CLOSED state.

While these boxed regions describe the essentially positive correlations within and between different structural regions, the other regions of the DCC maps also reveal interesting correlation behavior. Thus, in both the states, nearly all the junction loop nucleotides are positively correlated with each other (regions 7–9), and negatively correlated with stems P2 and P3 and loops L2 and L3. Interestingly, the loops themselves are also negatively correlated with the helices. A notable exception appears to be with U48 of J2-3, which, in the CLOSED state, is seen to correlate poorly with other junctions and yet correlate positively with loop L3.

Comparison of the average representations of the OPEN and CLOSED trajectories in the context of interactions observed in the crystal structure

Analysis with respect to the average representation of trajectories

In order to understand the switching mechanism, we need to evaluate the differences in the atomic level interactions in the two states and follow-up on the functionally significant changes associated with ligand binding. However, the only point of reference so far, for understanding detailed atomic level interactions, is the crystal structure of the ligand-bound aptamer. The interactions defining the functionally important structural blocks in the crystal structure are not necessarily preserved during the CLOSED state simulation because of its significant residual flexibility. The most striking observation is that even the interaction between the adenine ligand and its discriminating binding pocket host U74 is severely compromised during a large part of the CLOSED state simulation. How then can we compare and contrast base pairing interactions, which are missing in the OPEN state trajectory and at the same time not persistent during the CLOSED state simulation? To address this issue, we have computed the time-averaged representations of the OPEN and CLOSED state trajectories, and have carried out our analysis with respect to these average structures as static representations of the two states in solution phase. The relevant features of our analysis are shown in Figures 5 and 6.

Base-pairing interactions in the crystal structure, which do not persist during the MD simulation of the CLOSED state, can be identified easily in terms of highly distorted or even disengaged base pairs in the CLOSED state average structure. These base pairs, weakened due to binding-related conformational changes, are important since they confer some flexibility that contributes to the stability of the bound state in terms of entropic benefits. Thus, while the comparison of the OPEN state average structure with the crystal structure helped us identify essential binding-induced changes in atomic level interactions, the comparison between the crystal structure and the CLOSED state average structure indicated the nature of conformational penalties that accompany ligand binding.

In order to independently evaluate the changes in the strength of interactions, we also calculated the mean and standard deviation values of energies of relevant base clusters and substructural blocks for geometries observed

during the simulations using CHARMM27 force field. Table 2 shows a block-wise listing of the solvent accessible surface area (SASA) for relevant bases, DCC for base pairs, and average energies. The more stable interactions show larger negative mean values for average energies while the standard deviation (σ) values additionally tell us about the nature of global strain affecting the local interactions. Better base pairing is also indicated by higher DCC coefficient and, usually, lower SASA values.

Comparison and validation of average structures with respect to experimental observations

The RMSD values of the average structures follow the same trend as the average RMSD values (Table 1) of the two trajectories. Figure 6A, below, shows the superposition of the backbone traces of the minimized crystal structure with the average OPEN (RMSD: 4.01 Å) and CLOSED (RMSD: 2.51Å) state structures. More importantly, the RMSD of the OPEN state trajectory when calculated with respect to the average CLOSED state structure, shows that the average RMSD and standard deviation of the roof and dome portion (1.79 Å, $\sigma = 0.07$) are significantly lower than those for the rest of aptamer domain (6.16 Å, $\sigma = 1.34$) (Fig. 2C; Table 1). These features are highlighted in the superposition of backbone traces of the OPEN and the CLOSED state average structures (Fig. 6B) with respect to the roof–dome region above the stretches delineated by the residues U25/A45 and A56/U70 shown as blue and red sticks for the OPEN and CLOSED states, respectively. An RMSD of only

0.89 Å for the regions above the bounding residues is seen, while the rest of the aptamer has an RMSD of 4.65 Å.

Thus, a comparison of the two average structures reveals that the major contributions to the difference in the RMSD values of the average conformations are due to the loosening up of P1 helix, the junction proximal base pairs of the P3 helix and the junction loops forming the binding pocket. This picture of globally structured and locally unstructured OPEN state transitioning to a locally structured CLOSED state within the same global structure, conforms to experimental conclusions and validates the average structures (Lemay and Lafontaine 2007; Ottink et al. 2007; Rieder et al. 2007; Stoddard et al. 2008).

Variation in the interaction patterns of the components of binding pocket

We have also compared the interactions within the average CLOSED and OPEN state structures, blockwise, with those identified in the crystal structure. In the binding pocket of the crystal structure, the adenine is hydrogen bonded to U22, U47, U51, and U74; it is encapsulated by two roof triplets (RTs) and two floor triplets (FTs) (Fig. 1A). The schematic representation of the average OPEN and CLOSED state structures is shown in Figure 1B.

Roof triplet interactions

The water-mediated interaction constituting RT1 [U22–A52(w)A73] is not present in the OPEN state. In the CLOSED state simulation, it is found to be present for most

TABLE 2. Significant interactions with respective SASA (in Å^2), DCC coefficient, and average energy (kcal/mol) values

|	SASA (Å^2)			Paired base (OPEN)					Paired base (CLOSED)				
							Average energy					Average energy	
Block	Base ID	OPEN	CLOSED	Base ID	SASA (Å^2)	DCC	(kcal/mol)	(σ)	Base ID	SASA (Å^2)	DCC	(kcal/mol)	(σ)
P1	U20	172	166	A76	165	0.46	−105.8	8.2	A76	122	−0.20	−108.3	7.5
	A21	131	113	U75	229	0.33	−102.5	8.3	U75	143	0.83	−106.6	7.5
J1-2	U22	110	73	A52	170	0.82	−111.3	6.1	A52	106	0.89	−114.1	5.6
	A23	199	191	G46	128	0.71	−117.4	7.3	G46	119	0.75	−120.6	6.3
	A24	109	99	A73	242	0.86	−26.05	5.8	A73	191	0.57	−14.72	6.1
J2-3	G46	128	119	C53	144	0.81	−207.9	5.2	C53	164	0.35	−189.7	10.7
	U47	117	82	NA	NA	NA	NA	NA	Ade	18	0.53	−170.8	6.1
	U48	259	292	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	U49	224	152	A76	165	0.26	−99.1	5.7	A76	122	0.49	−100.7	6.2
	C50	189	170	U75	229	0.25	−162.1	5.8	U75	143	0.63	−165.9	6.7
	U51	238	144	NA	NA	NA	NA	NA	Ade	18	0.84	−174.5	6.4
	A52	170	106	U22	110	0.82	−111.3	6.1	U22	73	0.89	−114.1	5.6
	C53	144	164	G46	128	0.81	−207.9	5.2	G46	119	0.35	−189.7	10.7
P3	C54	133	158	G72	173	0.79	−207.3	5.7	G72	178	0.81	−205.2	5.9
	A55	183	204	U71	171	0.72	−111.2	5.9	U71	205	0.64	−106.8	6.1
J3-1	A73	242	191	A24	109	0.86	−26.05	5.8	A24	99	0.57	−14.72	6.1
	U74	215	191	NA	NA	NA	NA	NA	Ade	18	0.55	−168.2	5.8

Note: "NA" indicates not applicable.
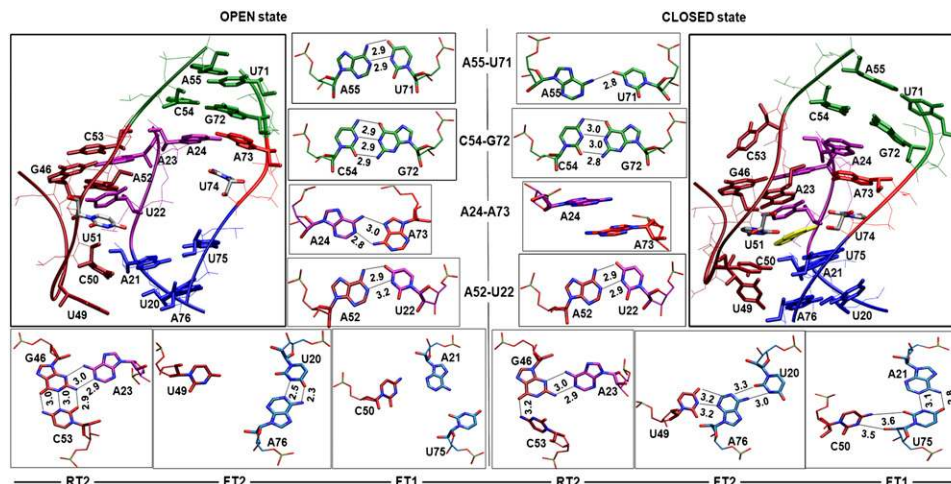
FIGURE 5. Comparison of specific significant interactions in average OPEN and CLOSED state structures. Note change in base-pairing geometry and interactions of C53 and A55, in contrast with retention of the same for A52 and C54. Hydrogen-bond distances are given in angstroms.

part of the trajectory, albeit with intermittent disruptions. The U22–A52 pair of RT1, however, while being stable in both the structures; is more compact in the average CLOSED state structure as indicated by interaction energy, DCC coefficient, and SASA values (Fig. 5; Table 2). A73, the third base of the triplet, as observed in the average structure and as borne out by higher interaction energy and higher DCC coefficient value (Table 2), is paired with A24 in the OPEN state but stacks below it in the CLOSED state. Interestingly, despite being paired in the OPEN state and unpaired in the CLOSED state, SASA values show both A24 and A73 to be more solvent exposed in the OPEN state. A similar variation in SASA value for A24 is also indicated in experimental studies (Rieder et al. 2007). While lower SASA for the two bases can easily be explained in terms of their mutual stacking in the CLOSED state, as we shall see later, their high OPEN state SASA has functional implications in terms of accessibility of U74 in the OPEN state.

The roof triplet RT2 (A23∘G46–C53) is persistent throughout the simulation in OPEN state. However, it is not persistent in the CLOSED state, though it is present in the crystal structure. This is reflected in terms of better average energy of the base triplet in the OPEN state (−218.0 kcal/mol) compared with that in CLOSED state (−205.2 kcal/mol) (Table 3). Examination of the CLOSED state average structure reveals a large distortion in the geometry of the canonical base pair, G46–C53 (Fig. 5), and indicates that this destabilization of RT2 is because of binding related conformational strain acting upon the pair. The role of this base pair in the destabilization of RT2 is also borne out by average energy, DCC coefficient, and SASA values in Table 2. In fact, the plot of average energy versus time, for G46–C53 during the CLOSED state simulation, indicates that the base pair gets distorted after 4 nsec of the simulation time (Supplemental Information SI3). On the other hand, the A23∘G46 pair of RT2 remains

properly paired in both states, while being more compact in the CLOSED state (Table 2).

FT interactions

The average energy values (Table 3) for both FT1 (A21–U75•C50) and FT2 (U20-–A76•U49) show better values for the CLOSED state (FT1: −181.4 kcal/mol; FT2: −200.0 kcal/mol) than for the OPEN state (FT1: −178.9 kcal/mol; FT2: −197.2 kcal/mol). Of the four base pairs constituting these two triplets, the only pair present in the crystal structure as well as in both the average structures, is the P1 base pair U20–A76 of FT2 (Fig. 5). All the other three base pairs, namely, the P1 base pair A21–U75 and the two J2-3–P1-linking base pairs U75•C50 of FT1 and A76•U49 of FT2, though present both in the crystal structure as well as in the average CLOSED state structure, are absent in the average OPEN state structure (Fig. 5) These inferences are substantiated by the average energy, DCC, and SASA values (Table 2), and point toward the existence of a molecular level connection between J2-3 linking with P1 to form the floor triplets and P1 helix stabilization.

Junction loop J2-3 interactions

Of the eight bases, 46–53, constituting the J2-3 loop, the bases G46, A52, and C53 are components of the roof triplets, and their interactions in both the states have already been evaluated in context. The remaining stretch of five bases, 47–51, which are unpaired in the OPEN state, get involved in the binding and switching process in the CLOSED state.

An examination of the OPEN state DCC coefficients of this stretch of unpaired bases shows the following:

1. All of them have fairly high DCC coefficients (0.53–0.71) with A21 of P1, where the lowest value was for U51 and the highest value was for U47.
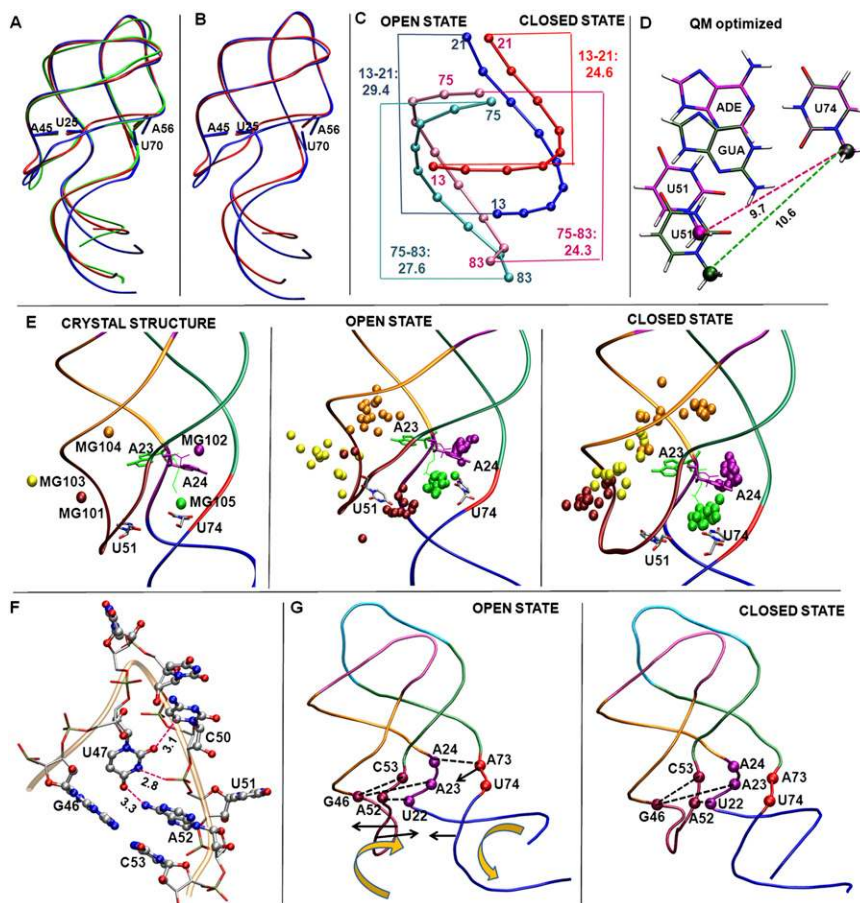
FIGURE 6. (*A*) Superposition of the average structures of OPEN (in blue color) and CLOSED (in red color) states with that of the crystal structure (in green color). (*B*) Superposition of average OPEN (in blue) and CLOSED (in red) state structures. (*C*) Lengthening of the P1 helix in the average structure of the OPEN state compared with that of the CLOSED state. (*D*) Superposed QM optimized geometries of triplets U74–Guanine–U51 (in green color) and U74–Adenine–U51 (in magenta color). C1′–C1′ distances are shown between U51 and U74. (*E*) Mg^{2+} ions positions captured at each 1-nsec snapshot for the OPEN and CLOSED states. (*F*) Tertiary contacts of the U47 shape the dynamics of the J2-3 backbone and mediate an essentially flipped out geometry of U51. (*G*) Anchoring of P2 and P3 to keep the roof and dome above the binding pocket in the preformed fold in the open state.

2. All of them, apart from U48 and U51, are also similarly correlated with U22 (DCC coefficients 0.58–0.84) with the highest value also for U47.

3. All of them, apart from U48 and U51, are well correlated with the other bases of J2-3. A DCC value >0.5 for U48 was observed only with U49 (0.61) and C50 (0.56), while that for U51 was observed only with C53 (0.52).

4. In contrast, U47 has noticeably high DCC score with all the other J2-3 bases: U46 (0.80), U49 (0.67), C50 (0.78), A52 (0.85), and C53 (0.74).

These observations lead us to two major conclusions regarding the disposition of J2-3 backbone in the OPEN state.

1. The dynamics of both U48 and U51 are less correlated with the rest of the J2-3 bases. This implies that U48 and

U51 mostly remain flipped out from the rest of the loop bases and are more accessible to the ligand. Since U51 is directly involved with ligand binding, it follows that it might also provide an initial docking site for the ligand. The flipped out geometry of U51 is also supported by its SASA value of 238 Å^2.

2. The dynamics of U47 is strongly correlated with the rest of J2-3, including the base pairs U22–A52 and G46–C53, and possibly plays a significant role in shaping the overall dynamics of the J2-3 backbone.

In order to investigate a possible correlation between the two conclusions above, we examined the molecular level interactions mediated by U47 (Fig. 6F) in the average OPEN state structure. As can be seen from the figure, the strong hydrogen bonding interaction between N3 of U47 with O1P of U51 indeed appears to mediate the flipped out geometry of the latter. The flipping out of U51 also appears to be assisted by the two tertiary contacts between O2 and O4 of U47 with C50 and A52, respectively. The hydrogen bonding pattern of U47 in the average OPEN state structure thus explains its crucial role in shaping J2-3 dynamics as inferred from DCC studies.

Ligand interactions within the binding pocket

In the crystal structure, adenine as ligand is bound to the O2′ of U22 (of J1-2), O2 of U47 and U51 (of J2-3), and U74 (of J3-1). The obvious absence of these interactions, in the average OPEN state structure, is indicated by the higher SASA values of

TABLE 3. Energetics of base clusters

Structural blocks	Paired bases	OPEN state		CLOSED state	
		Energy (kcal/mol)	σ	Energy (kcal/mol)	σ
Roof Triplet 1	A23 ∘ G46–C53	−218.0	8.0	−205.2	9.1
Floor Triplet 1	A21–U75•C50	−178.9	9.8	−181.4	9.4
Floor Triplet 2	U20–A76•U49	−197.2	7.8	−200.0	7.8
P1 helix (−FTs)	13–19/77–83	−1107.5	19.1	−1112.8	20.8
L2–L3 complex	32–38/60–66	−901.1	17.9	−904.2	18.6

these residues (Table 2). Interestingly, U51 is more solvent accessible (238 Å^2) than U74 (215 Å^2) in the OPEN state, owing to U74 stacking between A73 and U75 (Fig. 5).

In the CLOSED state, though the interactions of adenine with U51 are persistent, those with U22 and U47, and even with U74, are neither present in the average CLOSED state structure nor are they persistent during the simulation. This is also suggested by the higher interaction energy and higher DCC coefficient values for Ade∘U51 compared with that of Ade-U74 in the CLOSED state (Table 2).

Junction proximal base pairs of P3

Of the two junction proximal base pairs of P3, C54–G72, and A55–U71, which are well paired in the crystal structure, the latter becomes significantly less stable on switching from the OPEN to the CLOSED state. This is also reflected in the average structures and supported by average energy values, DCC coefficients, and SASA values (Table 2). The C54–G72 pair also appears more compact, and though, marginally so, is stabler in the OPEN state. These observations are suggestive of binding related conformational changes in the junction proximal end of the two P3 strands that pivot around C54–G72 and affect the stability of A55–U71 (Fig. 5).

Loop–loop interactions

The two base quartets involved in mediating the loop–loop interactions in the crystal structure are conserved in both the average structures. The average energy of the L2–L3 cluster shows that the CLOSED state (−904.2 kcal/mol) is slightly more stable than OPEN state (−901.1 kcal/mol) (Table 3). Since these interactions are responsible for holding together the roof–dome architecture from the top, and the roof–dome domain shows very little RMSD between OPEN and CLOSED states, it is unlikely that they have any major role in the switching process. Experimental observations also suggest that the loop–loop interactions are important but not essential for the functioning of purine riboswitches (Lemay et al. 2006; Stoddard et al. 2008)

Variation of P1 helix length

The 5′-to-3′ distances of backbone phosphorus of the strands of the helix for the average structure of OPEN state (strand1 [13–21]: 29.4 Å; strand2 [75–83]: 27.6 Å) are significantly higher than that for the average structure of the CLOSED state (strand1 [13–21]: 24.6 Å; strand2 [75–83]: 24.3 Å) (Fig. 6C). Average energy values related to interactions between P1 bases, as observed for the two junction proximal base pairs, as a part of the two floor triplets (Tables 2, 3), as well as for the rest (13–19 and 77–83) of the P1 helix (Table 3: OPEN: −1107.5 kcal/mol; CLOSED: −1112.8 kcal/mol), indicate a stabilization of P1 in the CLOSED state. It also suggests that this stabilization, which constitutes the switching, is related to the shortening of the

helix. The shortening, accompanied by an anticlockwise twisting (Fig. 6G) of the helix required for the formation of the two floor triplets, is in turn, related to ligand binding.

The role of magnesium ions

In the crystal structure, there are five Mg^{2+} ions present, either bound (MG102, MG105) or positioned deep within the core involving junctions and helices P2 and P3 (MG101, MG104) or on the surface (MG103). Two of these Mg^{2+} ions, MG102 and MG105, are penta-hydrated, and in both states show minimal mobility with respect to their initial positions due to their interaction with the oxygen atoms of phosphate backbone of nucleotides A24 and A23, respectively (Fig. 6E). In the OPEN state, these two Mg^{2+} associated bases play crucial roles in anchoring the junction proximal ends of helices P2 and P3. MG102 keeps the A23 residue in position for its interaction with G46, and MG105 keeps A24 in a position that it can pair with A73. This observation, along with the fact that fluorescence experiments (Rieder et al. 2007) imply a major role for Mg^{2+} toward reducing the solvent accessibility of A24 in the OPEN state, points toward a defining role of Mg^{2+} ions in stabilizing the preformed architecture of the OPEN state. Experimental ^{1}H-NMR studies (Ottink et al. 2007) on the guanine-free G-riboswitch also indicate the importance of Mg^{2+} ions in proper folding of the aptamer domain even in the absence of a ligand. It may be mentioned here, that Mg^{2+} ions are known to stabilize even intrinsically unstable base-pairing geometries, as in the case of the Levitt base pair in tRNA (Oliva et al. 2007).

MG101 is positioned within the junction J2-3, nearer to the adenine ligand and retains its initial position with less mobility in the CLOSED state. From OPEN state trajectory, it is observed that MG101 diffuses out of the junction J2-3 and settles inside the binding pocket. MG103 and MG104, however, are not involved in any interactions and show greater mobility.

Relating the conformational dynamics of the junction loops to the binding and switching process

J1-2 dynamics

In the crystal structure, A73 is a part of the roof triplet RT1 and A24 stacks between C54 of P3 and A73. This stacking, which helps in the encapsulation of the ligand bound to U74, is also seen in the average CLOSED state structure. In contrast, A73 is paired with A24 in a W:H *trans* geometry, in the OPEN state. This hikes up the junction loop J1-2, increasing the accessibility of U74.

J2-3 dynamics

The average structures along with the trajectory analysis confirm that the nucleotides involving the binding pocket

are disordered in the OPEN state, where the P1 helix is also less stable. The correlated flapping movement of the J2-3 loop (DCC and RMSF analyses), which supports the earlier suggestions (Gilbert et al. 2006; Ottink et al. 2007) that the ligand approaches the host U74 from the side of the J2-3 junction loop, is particularly significant. The role played by U47 in shaping the overall J2-3 dynamics and in forcing U51 in a solvent-exposed flipped out geometry is also noteworthy (Fig. 6F). Here, we list out other significant roles of J2-3 residues in the binding and switching process:

1. In the OPEN state, in the absence of ligand and interactions with A73, the U22–A52 pair follows the unrestricted flapping movement of the 47–51 stretch of the junction loop J2-3. This is indicated by significantly higher SASA for both the bases in the average OPEN state structure and decreased average interaction energy between the two bases in the OPEN-state trajectory. The resultant dynamics of U22 on the one hand and the dynamics of the ligand-free U74 on the other, explains the destabilization (reduced DCC coefficient) of P1 base pair U75–A21, constituted by their respective neighbors.
2. The G46–C53 base pair along with A23 of J1-2, forming the second roof triplet RT2, is the loop closing pair for J2-3; in the OPEN state it helps to anchor P2 and P3 stems side by side (Fig. 6G).
3. C50 and U49 get strongly latched on to the A21–U75 and U20–A76 pairs of P1, respectively, to secure the floor triplets, thus stabilizing the P1 helix in the CLOSED state (Tables 2, 3).
4. It appears that for U49 and C50 to sequester the P1 helix, U48 needs to flip out. It may be hypothesized that proper positioning of U51, along with its companion U47, as a consequence of ligand binding, cause this flip. Apart from the fact that this seems to constitute an essential component of binding-related conformational change, U48 does not appear to play any role in the binding process (Rieder et al. 2007). However, mutational analysis (Gilbert et al. 2007) has shown that, if nucleotide 48 has "Watson–Crick complementarity" with the nucleotide at position 74, then the aptamer almost completely loses its binding capacity. Understandably, such a mutation could initiate the collapse of the binding pocket, through the formation of a 48:74 pair, and thereby block access of the actual ligand to the position 74.

Putative crucial roles of U51

A close look at the structural changes during the CLOSED state simulation shows that the binding pocket is spacious enough to allow the adenine to move around, exploring the whole cavity, and its interactions with U22 and U47 seem to get disrupted after a run of 7–8 nsec. Even the hydrogen bonds between adenine and U74 are maintained only intermittently during the simulation. Only the strong pairing between adenine and U51 keeps pulling the latter away, thus breaking the labile hydrogen bonds between U51···U47 and U22···Ade. DFT results (Sharma et al. 2009) highlight the importance of U47 and U51, both in the actual binding process as well as in the stabilization of the binding pocket and the floor triplets, and hence, in the stabilization of the P1 helix. In terms of hydrogen bonding energy, only 26% of the total binding energy is due to the adenine–U74 interaction. Surprisingly, even U22 contributes about 20% to the binding. The relative importance of U47 and U51 to the binding of ligand is clear from their combined contribution of 54% of the total. Our ab initio studies thus support the observation contrasting the transient nature of the U74–Ade interactions with the significantly stabler U51∘Ade interactions. It thus appears likely that U51 plays crucial roles in the binding process.

U51 may play an important role in A/G discrimination

That the functioning of the purine riboswitches and their A/G specificity is crucially dependent on the identity of Y74 appears to be fairly well established (Mandal and Breaker 2004a). The question is, "what is the mechanism?" Is it that the recognition of the purine ligand, through Watson–Crick pairing with Y74, leads to a cascade of conformational changes that ultimately lead to the "switch?" Quantum chemical evaluation of A:U and G:U W:W *cis* interactions indicate that guanine binds more strongly with uracil than adenine does (Bhattacharyya et al. 2007). Hence, the question still remains as to what happens after the docking of the purine ligand to Y74 that ensures A/G discrimination. Our studies indicate that U51 may hold the key to the answer to this question.

The W:W *cis* G:U base pair, though stronger than canonical A:U (Bhattacharyya et al. 2007), has a wobble geometry and the positioning of guanine as ligand, compared with adenine, with respect to U74 is very different. In the context of the A-riboswitch, this would mean that the C1′–C1′ distance for the pair U51 and U74, in the ligand-bound state, should also be very different for guanine and adenine as ligands, respectively. In order to get an approximate idea of the variation in the geometries and distances, we generated a hypothetical model describing the quantum chemically optimized geometry of the triple U74–Gua∘U51 and superposed the U74 with that of the triple U74–Ade∘U51. As expected, the C1′–C1′ distances (10.6 Å for the former as against 9.7 Å for the latter) varied by nearly 10% (Fig. 6D). It may be mentioned here that the C1′ atoms of U51 and U74, in the models were replaced by methyl groups, and computations were carried out using methodology described elsewhere (Sharma et al. 2009).

The hypothesis is that proper ligand is necessary to appropriately place U74 and U51 so that the U75•C50 and

A76•U49 pairs can form and sequester the two floor triplets. The idea behind this argument is very similar to the concept of isostericity, as used to explain covariation (Leontis et al. 2002) of base pairs in evolutionarily related RNA molecules. This hypothesis blends well with available data and also explains why U51 is one of the conserved residues in all purine riboswitches, e.g., in the G-riboswitch where essentially only U74 is changed to C74. We can thus ascribe this to the two triples U74–Ade∘U51 and C74–Gua∘U51 being isosteric, while the discrimination of adenine riboswitch against guanine can be put to the fact that the triple U74–Gua∘U51 is not isosteric with the triple U74–Ade∘U51.

This isostericity hypothesis gets additional support from the fact that P1 stabilization is associated with 98% encapsulation of the ligand. Apart from the involvement of U47, U51, and U74, this also involves U22, which in the CLOSED state, forms a water-mediated hydrogen bond with A73 and also maintains pairing with A52, in order to complete the "roof" triplet RT1. This tying up of U22 possibly has an important role to play in damping the dynamics of A21 of P1 and thus help in stabilizing the P1 helix. Replacement of the triple U74–Ade∘U51 with nonisosteric triple U74–Gua∘U51 could make it difficult for U22 to utilize its O2′ to form hydrogen bonding with the N7 of guanine and also to participate in forming RT1.

U51 may be "capturing" the ligand and bringing it into the binding pocket, toward U74

Another question is: Although the binding kinetics is slower than what is predicted for diffusion control mechanisms, is the approach of the ligand to the binding pocket diffusion controlled or are there more active mechanisms available for "fetching" the ligand? The suggested two step-binding process involving an initial low-affinity association of the ligand with some "nonrandom" site of the aptamer, followed by slow binding with reorganization of the binding pocket (Buck et al. 2007), supports the second possibility. Here we propose a plausible explanation. In the OPEN state, U51, a constituent of the dynamic 47–51 stretch of J2-3 is flipped out of its stacking with C50 and A52 (Fig. 6F) and displays large fluctuations in its χ value (Fig. 3). Thus, it freely samples the solvent medium and appears to be a strong candidate as the initial docking site for the ligand.

Both adenine as well as guanine can bind strongly with the Watson–Crick edge of U51, in several ways; the natural way is from the Watson–Crick side in a *cis* fashion, whereby, adenine would form a canonical base pair and guanine would form a wobble base pair. Another type of *cis* pairing, with "free" adenine or guanine, is from their "sugar edge" side. In the course of its flapping movement, it can capture and carry the ligand to U74 for recognition, possibly more efficiently than that in a diffusion-controlled manner. Sev-

eral pointers support the possibility of the proactive role of U51 toward ligand recognition and binding.

1. Although U74 has its Watson–Crick edge exposed in a docking ready position, and an opening is maintained due to the presence of the A24–A73 base pair, it is stacked between A73 and U75 and does not appear to be easily accessible.
2. U51 with its Watson–Crick edge facing away from the main body of the aptamer (Fig. 5) is more solvent exposed, and, being a part of the dynamic J2-3 loop, has greater probability of meeting with the ligands in the solvent surrounding the aptamer.
3. Since the interaction energy of its W:S *cis* base pairing with free adenine is more than that for W:W *cis* base pairing, our model provides with the possibility of U51 being the initial docking site for the ligand.
4. Once U51 has adenine docked properly in a W:S *cis* geometry, U51 can swivel around with the ligand, for U74 to bind.

A molecular level understanding for the mechanism of the switching process

Based on our analysis of MD simulation trajectories and the corpus of experimental data reported to date in the literature, we propose a multistep "model," for the ligand-binding induced transition from the OPEN state to the CLOSED state (Fig. 7) and its correlation with the OFF to ON transition in the riboswitch. It embodies a molecular level understanding of the conformational changes involved, in terms of the making and breaking of H-bonds and strengthening and/or weakening of base-pair interactions. In addition to the essential and well-established features of the binding and switching process, our model also addresses experimental observations that are either as yet unexplained or that at most have speculative explanations. It may be mentioned here that since the simulations were carried out in the presence of Mg^{2+}, our essential conclusions imply a similar environment. The representation of the Mg^{2+} free state (Fig. 7) and its difference from the binding competent state has been extrapolated from available experimental data.

The OPEN state

The OPEN state structure, corresponding to the average structure obtained from the ligand-free trajectory, is labeled the "Binding Competent" in Figure 7. It is preorganized in the global sense involving a dome above the binding pocket and is locally disorganized at the binding pocket. The following interactions give shape to this preorganization and ensure ligand accessibility to the key binding nucleotides U51 and U74.
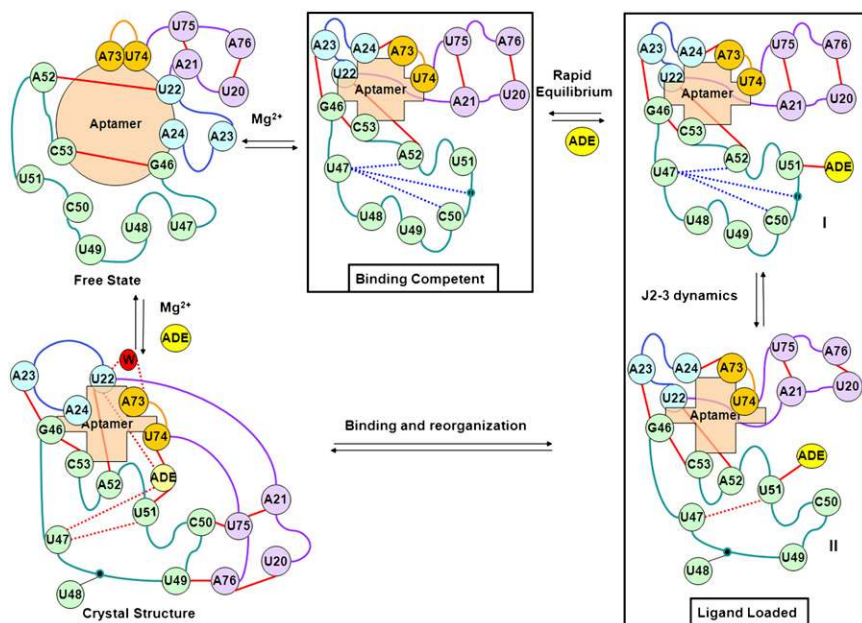
FIGURE 7. Schematic representation of our multistate model for the binding and switching process. The preformed roof–dome architecture is not included. The nucleotides corresponding to the three junction loops, and the functionally important P1 nucleotides are shown using four different colors, respectively. The red circle represents a water molecule in a water-mediated hydrogen bond. The yellow oval represents the adenine ligand. The backbone inherits colors from the colors of the residues it connects up, while base-pairing interactions are shown in red. The black dot on the backbone indicates the phosphate group. Blue dotted lines represent weak tertiary interactions in the OPEN state, while the red dotted lines indicate tertiary interactions evolving during the process of binding. The terminal residues of the junction loops are shown having contact with the aptamer. A change in shape of the aptamers represents conformational changes.

1. Two base quartets accounting for the L2–L3 interaction (not shown in the figure), constitute a dome that holds the P2 and P3 helices together from the top.
2. Four base-pairing interactions: G46–C53, A23–G46, U22–A52, and A24∘A73, anchoring the junction/binding site proximal ends of the P2 and P3 helices, aids the side-by-side placement of P2 and P3 stems as well as the coaxial stacking of P1 and P3. In this, the three nucleotides 22–24 of J1-2 play an important role.
3. The base pair G46–C53 brings the ends of the loop J2-3 together, while A23∘G46 completes the roof triplet RT2 tying this with J1-2.
4. The third base pair U22–A52 links a third J2-3 base with J1-2, while the fourth "anchoring" base pair, A24∘A73, allows solvent accessibility to U74 by keeping the junction loops J1-2 and J3-1 apart from each other.
5. Phosphate oxygen atoms of A23 and A24, found to be coordinating to two Mg^{2+} ions, respectively, underline the importance of Mg^{2+} ions toward maintaining a "binding friendly" conformation by stabilizing the base pairs A23∘G46 and A24∘A73.
6. The five remaining J2-3 bases, 47–51, constituting a flexible backbone stretch, flaps around together. U47 interacts strongly with the phosphate of U51 and also with

C50 and A52. This self-correlated flapping motion, which is fairly uncorrelated with the rest of the aptamer, while primarily contributing toward the local disorganization of the binding pocket, is essential for ligand access inside it.

Correlating ligand binding with switching

In the OPEN state, the P1 helix is prone to melting. The switch involves ligand-binding-induced stabilization. The molecular interactions correlating the ligand binding with the switching may be identified as follows: (1) The high mobility of U51 increases the mobility of its neighbor A52, which in turn, makes its strongly associated partner U22 highly dynamic. The fluctuations of U22, thus enhanced, along with the fluctuations of unpaired U74, affect the stability of the junction proximal base pair A21–U75 of P1, and are possibly responsible for the high melting propensity of P1 in the OPEN state. (2) Availability of the ligand sets off two sets of coordinated conformational changes in the aptamer, resulting in the near complete encapsulation of the ligand and in the simultaneous stabilization of the P1 helix. These changes are:

1. *J2-3 Dynamics:* to enable U51 to come closer to ligand-bound U74, the five unpaired J2-3 nucleotides, 47–51, undergo a directed conformational change. U47 frees up C50 and A52 and joins U51 in ligand binding while A52 maintains a stable base pair with U22, which also binds to ligand through its sugar O2′. Thus, U22 and U47 jointly assist in stabilizing the base pair A23∘G46. Essentially, this implies a unimolecular conformational rearrangement where existing tertiary interactions (blue dotted lines) involving U47 give way to new interactions (Fig. 7, red dotted lines), which involve a complete flipping out of U48 and swiveling in of U51. Simultaneously U49 and C50 come closer to U75 and A76, respectively, into a position favorable for the floor triplet formation.
2. *Binding and reorganization:* to enable U74 to come closer to the ligand-bound U51, A73 has to break its pairing with A24 to come down and stack below the latter. Thus, in the CLOSED state A73 forms the water-mediated hydrogen bond with U22 of the U22–A52 pair to complete the roof triplet RT1. In the process, the nucleotides U75 and A76 are brought to pair with C50

and U49, respectively, and thus, complete the two floor triplets that stabilize the junction proximal P1 base pairs.

Stepwise outline of the binding and switching mechanism

Thus, our model for the binding and switching mechanism (Fig. 7) involves essentially two steps, though a three-step mechanism cannot be ruled out. The first is a fast reversible bimolecular step involving the docking of the ligand on to U51 of the aptamer, which is in a binding competent conformation in the presence of Mg^{2+} ions. The resultant ligand loaded I structure may undergo the two sets of conformational changes outlined above, in a concerted manner through one single slow unimolecular step, to give rise to the final bound form. Alternatively, a slow unimolecular equilibration of the ligand loaded structure I with structure II, involving J2-3 dynamics, may subsequently be followed up with a second slow unimolecular step involving binding and reorganization.

The average structure in the OPEN state, which corresponds to the binding competent form, provides a molecular level explanation for its rapid equilibrium with the ligand loaded form. In the OPEN state, U51 is more solvent exposed than U74 and, being a part of the dynamic J2-3 loop, has far greater probability of meeting with the ligands in the solvent surrounding the aptamer. Most importantly, since the interaction energy of U51 W:S *cis* base pairing with free adenine is more than that for W:W *cis* base pairing, an initial docking of the ligand onto U51 in the correct geometry does not necessitate any conformational changes. Thus, the process of ligand loading involves low activation energy and essentially constitutes a fast bimolecular step with the concentration of the ligand determining the fraction of ligand loaded aptamer at equilibrium.

The making and breaking of the hydrogen bonding interactions, as observed in the two sets of conformational changes, account for the slow binding kinetics in terms of the slow unimolecular second step. While our model does not specify whether the two sets of changes are sequential or are in concert, it is imperative that the binding related changes are stabilized only if the ligand is present and the subsequent conformational change, driven by the recognition of the ligand by U74, is completed. Thus, the conformational change indicated as J2-3 dynamics does not by itself lead to a stable state.

The A/G discriminative ability of purine riboswitches

Once adenine has docked properly with its Watson–Crick edge, U51 can then swivel around with the ligand, for U74 to bind. Only when the correct ligand in the proper orientation is brought toward U74, the latter complete the binding process. Guanine, in place of adenine as ligand, would not be able to retain isostericity with the canonical

A:U W:W *cis* base-pairing geometry, and the binding process is unlikely to be completed properly. This would account for the highly discriminative ligand specificity.

Significance of the preformed architecture

The importance of the L2–L3 interaction to form the dome, in stabilizing the binding-competent OPEN state structure and thereby improve the binding kinetics, has been proposed by Batey et al. (2004). However, Stoddard et al. (2008) have demonstrated that this interaction, while important, is not essential for the functioning of purine riboswitches by comparing the performance and association constants of high- and low-affinity ligands.

The proposed model enhances our understanding of the role of the preformed architecture. It is generally acknowledged that while a highly disordered OPEN state usually implies high entropic penalties when switching to an ordered CLOSED state, steric hindrance toward ligand access contraindicates very highly ordered OPEN state structures. In our model, the aptamer domain resolves this issue by bringing about ordering of the dome, which is far away from the binding pocket. Here the dome, being crucial for imparting binding competence, lowers the activation energy for U51 toward carrying the ligand to U74 and thus *speeds up the kinetics*. In terms of thermodynamics also, an ordered dome in the OPEN state would mean that the entropic penalty when going from the unbound state to the bound state would mainly entail the ordering of the binding pocket and the P1 helix.

In addition to this, our model identifies some of the enthalpic penalties associated with binding-induced conformational changes that take place in the roof domain. For example, constraints due to the disposition of the J2-3 loop result in a significant weakening of the G46–C53 pairing in the CLOSED state. Similarly, the P3 base pair A55–U71 is also substantially weakened due to the opposing binding induced movements of the J2-3 loop on the one hand and the backbone stretch in the J3-1 junction on the other hand. These enthalpic penalties are compensated in terms of entropic benefits arising out of enhanced residual flexibility in the roof–dome architecture. The importance of the preformed architecture may thus also be understood in terms of enthalpy–entropy compensation toward stabilizing the closed structure.

Explaining quantitative kinetics data

Experiments, including stopped flow fluorescence spectrometry (Gilbert et al. 2006) and time-resolved NMR spectroscopy (Buck et al. 2007), have provided quantitative estimates of a slow binding kinetics. Gilbert et al. (2006) have also shown that the rate is first order with respect to the ligand for low ligand concentrations and is independent of the ligand for very high concentrations. The model explaining this kinetic data, as proposed by the authors and

as discussed earlier, implies at least two major assumptions, which, though plausible, are not experimentally validated. One is that ligand binding to U74 constitutes the initial docking site, and the other is that this also constitutes the bimolecular step. More importantly, their model does not account for two important features:

1. There is evidence in terms of line broadening in the NMR spectrum of the ligand, which is indicative of a nonrandom initial docking site (Buck et al. 2007). The line broadening and change in chemical shift of residues in the aptamer binding site happens much later.
2. Guanine binds more strongly with uracil than adenine does. So there ought to be a discriminatory molecular recognition step *after* the bimolecular step, even if U74 is involved in the step.

Our model provides plausible explanations for the above features, and can also explain the broad kinetics features. Here, the rate-determining step is the slow unimolecular step involving the ligand loaded aptamer, which is in rapid equilibrium with the mixture of unassociated ligand and aptamer. Thus we have:

$$[\text{Ligand loaded aptamer}] = K_{\text{eq}}[\text{Aptamer}][\text{Ligand}].$$

This would mean that for low ligand concentrations, the rate would be first order with respect to the ligand. In the presence of excess ligand, naturally, the concentration of the ligand loaded aptamer will attain a constant "saturation" value, and hence, the rate would become independent of ligand concentration.

Visualization of the transition dynamics

The two sets of coordinated movement of the backbone during the transition, as described above, have been derived on the basis of comparison of the average structures, in the OPEN and the CLOSED states interpolating between the two average structures using the "morph" server (Krebs and Gerstein 2000; Flores et al. 2006). Though the morph does not address the issue of whether the two sets of coordinated binding-induced motions are concerted or are stepwise, the visualization provides insights into both the stabilization of P1 as well as the weakening of the base pairs mentioned above. The morph movie (Supplemental Information SI4; see also http://molmovdb.org/cgi-bin/movie.cgi, Morph id: m000830-19179) indicates a clockwise wrapping movement of J2-3 and a counterclockwise winding movement of both the strands of the P1 stem. This is likely to be the motion responsible for the gelling together of the bases in positions 51 to 49 with those from 74 to 76, in response to ligand binding. Arguably, the clockwise movement of J2-3 loop, pivoting around the loop closing base pair G46–C53, leads to the weakening of the latter. It may

be noted that the G46–C53 base pair, considered important for anchoring the helices P2 and P3 in the OPEN state, loses its importance, since the CLOSED state anchoring is taken care of by the two floor triplets.

The counterclockwise coiling up of the strands of the P1 helix, during its transition to the CLOSED state, leads to the shortening of the helix and clearly explains its enhanced stability. However, while the coaxially placed P3 inherits this movement only for its 3′ strand, the 5′ strand of P3 inherits its winding movement from J2-3, which moves clockwise. Thus, the two strands of P3 tend to wind in opposite directions, and this understandably puts a strain on the helix. This strain cannot get distributed to the upper base pairs because the helix is capped by the loop L3, which is stably tied up with L2 through two base quartets. We, therefore, observe a large distortion in the base pair A55–U71, going from the OPEN state to the CLOSED state.

CONCLUSIONS

We have carried out MD simulation studies of the aptamer domain of the *add* A-riboswitch, both in the presence and in the absence of the adenine ligand. Based on characterization and analysis of conformations, noncovalent interactions and solvent accessibilities of the individual nucleotides in the simulation trajectory, we have added functionally significant molecular level details to the existing experimentally developed picture of the structure and dynamics of the aptamer in the ligand-free OPEN state. We have also compared these parameters with those in the CLOSED state and have proposed a detailed model for the mechanism of ligand binding, which is consistent with currently available experimental data. Our model explains the mechanism of communication of the ligand binding event in the aptamer to the expression platform, in terms of the tightening of the P1 helix, in molecular details. It also helps to identify two crucial nucleotides, A24 and U51, which apart from U74 possibly play important roles in ligand binding and ligand discrimination. Our results open up newer possibilities for hypothesis driven investigations, both experimental as well as computational, into the functioning of purine riboswitches.

MATERIALS AND METHODS

MD simulations

All MD simulations were carried out using software NAMD2.6 (Phillips et al. 2005) with CHARMM27 force fields (Foloppe and MacKerell 2000; MacKerell and Banavali 2000). The *add* A-riboswitch coordinates were taken from pdb: 1Y26 (Serganov et al. 2004). Since the crystal structure is broken near the P–O5′ bond of G37, with a misplaced O5′ atom, we have refined the structure by minimization using CHARMM27 force field. This minimized structure, referred to as "crystal structure" in the

article, was used for preparing the respective initial structures for the production run for the "ligand-bound" CLOSED as well as "ligand-free" OPEN state simulations.

For explicit solvent calculations, the nucleic acid systems represented by the "crystal structure" coordinates, with and without the ligand adenine, respectively, were solvated in a cubic box with a 10 Å layer of TIP3P (Jorgensen et al. 1983) water and neutralized with addition of appropriate Na^+ ions. Though suggested to be problematic in terms of the effects of their hydration shell water molecules (Rhodes et al. 2006), Mg^{2+} ions present in 1Y26 were retained in the simulations. This was necessary, since, apart from considering their importance in stabilizing RNA folds, the crystal structure showed two of the Mg^{2+} ions to be interacting with backbone phosphate oxygen atoms.

For both CLOSED as well as OPEN state simulations, corresponding respectively to the ligand-bound and ligand-free states, minimization was carried out in two steps: first, for 50,000 cycles with all nucleic acid atoms fixed, followed by 50,000 cycles with no fixed atoms. The temperature was gradually raised from 0 K to 300 K with an increment of 6 K/psec, followed by backbone-restrained equilibration of 100 psec and unrestrained equilibration for 1.5 nsec. These thermally equilibrated structures were then used for 15 nsec production MD runs. Integration time step of 2 fsec was used, and the SHAKE algorithm was used to constrain bond lengths involving hydrogens. Periodic boundary conditions were employed to eliminate surface effects, and Particle Mesh Ewald (PME) was used for full electrostatic computations with tolerance of 10e-06 and grid-spacing of 1 Å. Langevin dynamics was used to perform NPT simulations at controlled temperature of 300 K at 1 atm pressure.

Trajectory and structural analyses

Trajectory analyses were done using built-in programs of VMD 1.8.6 (Humphrey et al. 1996). These include calculation of RMSD of all backbone atoms for each of the two states, respectively during the simulations, with reference to their corresponding preequilibration geometries or other geometries as mentioned in the Results section. We also calculated the RMSF of the C1′ atoms, in the trajectories of both the states, with reference to their respective initial positions at the beginning of the production run. The backbone torsion angles (δ, χ, ζ) of all the nucleotides were calculated using NUPARM (Bansal et al. 1995). Base pairs were identified and analyzed using BPFIND (Das et al. 2006) and 3DNA (Lu and Olson 2003). The variation in interaction energies involving individual base pairs and/or complete structural blocks were estimated, wherever required, with CHARMM27 force fields using software NAMD.

Analysis of average structures

From the trajectories in the ligand-free as well as in the ligand-bound states, representing the flexibilities of the CLOSED and the OPEN states, respectively, we used VMD to calculate the coordinates for their respective "all atom" average structures as representatives for the two states. The structures were minimized for 200 cycles with fixed backbone atoms, to remove clashes, using NAMD2.6. The two refined structures were compared by superposition and RMSD calculations, and were individually analyzed for the SASAs using the Lee–Richard algorithm (Lee and Richards

1971) implemented in CCP4 Program Suite (CCP4, 1994). Interpolation between the crystal structure and the two average structures was carried out using the morph server (Krebs and Gerstein 2000). Figures involving structures were generated using VMD.

Dynamic cross-correlation analysis

Studying the correlated motion between different regions of a biomolecule can provide insights into its dynamic behavior. Such correlated motions can be analyzed with a DCC map. The extent of correlated motions between two nucleotides is calculated as the magnitude of the correlation coefficient between their respective atoms. The cross-correlation coefficient, C_{ij}, for each pair of atoms i and j is calculated as

$$C_{ij} = \frac{<\Delta r_i> \cdot <\Delta r_j>}{\sqrt{<\Delta r_i^2><\Delta r_j^2>}},$$

where Δr_i is the displacement from mean position of the ith atom and $< >$ symbol represents the time average (Harte et al. 1990). Positive value of correlation coefficient indicates that the given pair of atoms is moving together in the same phase and same direction, whereas negative value indicates movement of atoms in opposite directions. The cross-correlation coefficient is calculated for the centroids of nucleobases, and the DCC map is calculated as time average over the simulation. DCCM analysis is done using in-house software.

Supplemental data

Fluctuations in torsion angles (δ, χ, and ζ), DCC coefficient values, plots of interaction energies for relevant base pairs and base clusters, and a morph movie showing transition from OPEN to CLOSED state average structures.

SUPPLEMENTAL MATERIAL

Supplemental material can be found at http://www.rnajournal.org.

NOTE ADDED IN PROOF

A 1-μsec MD simulation of the guanine riboswitch was reported online (Villa et al. 2009) about a week before this manuscript was accepted. Interestingly, these authors also conclude that residue U51 in the aptamer domain functions as a general docking platform for purine bases, whereas the interactions between C74 and the ligand are crucial for ligand selectivity. Our current article, on the adenine riboswitch, independently builds up on this basic

assertion and further elaborates upon molecular level details of the switching mechanism.

REFERENCES

Al-Hashimi HM, Walter NG. 2008. RNA dynamics: It is about time. *Curr Opin Struct Biol* **18:** 321–329.

Bansal M, Bhattacharyya D, Ravi B. 1995. NUPARM and NUCGEN: Software for analysis and generation of sequence dependent nucleic acid structures. *Comput Appl Biosci* **11:** 281–287.

Batey RT, Gilbert SD, Montange RK. 2004. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432:** 411–415.

Bhattacharyya D, Koripella SC, Mitra A, Rajendran VB, Sinha B. 2007. Theoretical analysis of noncanonical base pairing interactions in RNA molecules. *J Biosci* **32:** 809–825.

Buck J, Fürtig B, Noeske J, Wöhnert J, Schwalbe H. 2007. Time-resolved NMR methods resolving ligand-induced RNA folding at atomic resolution. *Proc Natl Acad Sci* **104:** 15699–15704.

Buskirk AR, Landrigan A, Liu DR. 2004. Engineering a ligand-dependent RNA transcriptional activator. *Chem Biol* **11:** 1157–1163.

CCP4. 1994. Collaborative computational project, number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50:** 760–763.

Cromie MJ, Shi Y, Latifi T, Groisman EA. 2006. An RNA sensor for intracellular. Mg^{2+}. *Cell* **125:** 71–84.

Dann CE, Wakeman CA, Sieling CL, Baker SC, Irnov I, Winkler WC. 2007. Structure and mechanism of a metal-sensing regulatory RNA. *Cell* **130:** 878–892.

Das J, Mukherjee S, Mitra A, Bhattacharyya D. 2006. Noncanonical base pairs and higher order structures in nucleic acids: Crystal structure database analysis. *J Biomol Struct Dyn* **24:** 149–161.

Edwards AL, Batey RT. 2009. A structural basis for the recognition of 2′-deoxyguanosine by the purine riboswitch. *J Mol Biol* **385:** 938–948.

Epshtein V, Mironov AS, Nudler E. 2003. The riboswitch mediated control of sulfur metabolism in bacteria. *Proc Natl Acad Sci* **100:** 5052–5056.

Eskandari S, Prychyna O, Leung J, Avdic D, O'Neill MA. 2007. Ligand-directed dynamics of adenine riboswitch conformers. *J Am Chem Soc* **129:** 11308–11309.

Foloppe N, MacKerell AD Jr. 2000. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem* **21:** 86–104.

Flores S, Echols N, Milburn D, Hespenheide B, Keating K, Lu J, Wells S, Yu EZ, Thorpe M, Gerstein M. 2006. The Database of Macromolecular Motions: New features added at the decade mark. *Nucleic Acids Res* **34:** D296–D301.

Fuchs RT, Grundy FJ, Henkin TM. 2006. The S_{MK} box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat Struct Mol Biol* **13:** 226–233.

Gilbert SD, Batey RT. 2006. Riboswitches: Fold and function. *Chem Biol* **13:** 805–807.

Gilbert SD, Stoddard CD, Wise SJ, Batey RT. 2006. Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *J Mol Biol* **359:** 754–768.

Gilbert SD, Love CE, Edwards AL, Batey RT. 2007. Mutational analysis of the purine riboswitch aptamer domain. *Biochemistry* **46:** 13297–13309.

Gilbert SD, Rambo RP, Van Tyne D, Batey RT. 2008. Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nat Struct Mol Biol* **15:** 177–182.

Greenleaf WJ, Frieda KL, Foster DAN, Woodside MT, Block SM. 2008. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319:** 630–633.

Grundy FJ, Lehman SC, Henkin TM. 2003. The L box regulon: Lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc Natl Acad Sci* **100:** 12057–12062.

Harte WE Jr, Swaminathan S, Mansuri MM, Martin JC, Rosenberg IE, Beveridge DL. 1990. Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proc Natl Acad Sci* **87:** 8864–8868.

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

Humphrey W, Dalke A, Schulten K. 1996. VMD: Visual molecular dynamics. *J Mol Graph* **14:** 33–38.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79:** 926–935.

Krebs WG, Gerstein M. 2000. The morph server: A standardized system for analyzing and visualizing macromolecular motions in a database framework. *Nucleic Acids Res* **28:** 1665–1675.

Lea CR, Piccirilli JA. 2007. 'Turning on' riboswitches to their antibacterial potential. *Nat Chem Biol* **3:** 16–17.

Lee B, Richards FM. 1971. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* **55:** 379–400.

Lemay JF, Lafontaine DA. 2007. Core requirements of the adenine riboswitch aptamer for ligand binding. *RNA* **13:** 339–350.

Lemay JF, Penedo JC, Tremblay R, Lilley DMJ, Lafontaine DA. 2006. Folding of the adenine riboswitch. *Chem Biol* **13:** 857–868.

Leontis NB, Westhof E. 2001. Geometric nomenclature and classification of RNA base pairs. *RNA* **7:** 499–512.

Leontis NB, Stombaugh J, Westhof E. 2002. The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res* **30:** 3497–3531.

Lin JC, Thirumalai D. 2008. Relative stability of helices determines the folding landscape of adenine riboswitch aptamers. *J Am Chem Soc* **130:** 14080–14081.

Lu XJ, Olson WK. 2003. 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **31:** 5108–5121.

MacKerell AD Jr, Banavali N. 2000. All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J Comput Chem* **21:** 105–120.

Mandal M, Breaker RR. 2004a. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* **5:** 451–463.

Mandal M, Breaker RR. 2004b. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat Struct Mol Biol* **11:** 29–35.

Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* **113:** 577–586.

Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL, Breaker RR. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* **306:** 275–279.

McDaniel BA, Grundy FJ, Artsimovitch I, Henkin TM. 2003. Transcription termination control of the S box system: Direct measurement of S-adenosylmethionine by the leader RNA. *Proc Natl Acad Sci* **100:** 3083–3088.

Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, Perumov DA, Nudler E. 2002. Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell* **111:** 747–756.

Montange RK, Batey RT. 2006. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **441:** 1172–1175.

Montange RK, Batey RT. 2008. Riboswitches: Emerging themes in RNA structure and function. *Ann. Rev Biophys* **37:** 117–133.

Mulhbacher J, Lafontaine DA. 2007. Ligand recognition determinants of guanine riboswitches. *Nucleic Acids Res* **35:** 5568–5580.

Murray LJ, Arendall WB 3rd, Richardson DC, Richardson JS. 2003. RNA backbone is rotameric. *Proc Natl Acad Sci* **100:** 13904–13909.

Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR. 2002. Genetic control by a metabolite binding mRNA. *Chem Biol* **9:** 1043–1049.

Noeske J, Richter C, Grundl MA, Nasiri HR, Schwalbe H, Wöhnert J. 2005. An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs. *Proc Natl Acad Sci* **102:** 1372–1377.

Noeske J, Buck J, Fürtig B, Nasiri HR, Schwalbe H, Wöhnert J. 2007. Interplay of 'induced fit' and preorganization in the ligand induced folding of the aptamer domain of the guanine binding riboswitch. *Nucleic Acids Res* **35:** 572–583.

Nudler E, Mironov AS. 2004. The riboswitch control of bacterial metabolism. *Trends Biochem Sci* **29:** 11–17.

Oliva R, Tramontano A, Cavallo L. 2007. Mg^{2+} binding and archaeosine modification stabilize the G15–C48 Levitt base pair in tRNAs. *RNA* **13:** 1427–1436.

Ottink OM, Rampersad SM, Tessari M, Zaman GJ, Heus HA, Wijmenga SS. 2007. Ligand-induced folding of the guanine-sensing riboswitch is controlled by a combined predetermined-induced fit mechanism. *RNA* **13:** 2202–2212.

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. 2005. Scalable molecular dynamics with NAMD. *J Comput Chem* **26:** 1781–1802.

Rhodes MM, Réblová K, Sponer J, Walter NG. 2006. Trapped water molecules are essential to structural dynamics and function of a ribozyme. *Proc Natl Acad Sci* **103:** 13380–13385.

Rieder R, Lang K, Graber D, Micura R. 2007. Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control. *ChemBioChem* **8:** 896–902.

Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2003. Regulation of lysine biosynthesis and transport genes in bacteria: Yet another RNA riboswitch? *Nucleic Acids Res* **31:** 6748–6757.

Roth A, Winkler WC, Regulski EE, Lee BW, Lim J, Jona I, Barrick JE, Ritwik A, Kim JN, Welz R, et al. 2007. A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nat Struct Mol Biol* **14:** 308–317.

Schneider B, Morávek Z, Berman HM. 2004. RNA conformational classes. *Nucleic Acids Res* **32:** 1666–1677.

Serganov A, Yuan YR, Pikovskaya O, Polonskaia A, Malinina L, Phan AT, Hobartner C, Micura R, Breaker RR, Patel DJ. 2004. Structural basis of discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem Biol* **11:** 1729–1741.

Serganov A, Polonskaia A, Phan AT, Breaker RR, Patel DJ. 2006. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441:** 1167–1171.

Sharma P, Sharma S, Chawla M, Mitra A. 2009. Modeling the noncovalent interactions at the metabolite binding site in purine riboswitches. *J Mol Model* **15:** 633–649.

Soukup JK, Soukup GA. 2004. Riboswitches exert genetic control through metabolite-induced conformational change. *Curr Opin Struct Biol* **14:** 344–349.

Stoddard CD, Gilbert SD, Batey RT. 2008. Ligand-dependent folding of the three-way junction in the purine riboswitch. *RNA* **14:** 675–684.

Sudarsan N, Nahvi A, Roth A, Collins JA, Breaker RR. 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes & Dev* **17:** 2688–2697.

Thore S, Leibundgut M, Ban N. 2006. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312:** 1208–1211.

Villa A, Wohnert J, Stock G. 2009. Molecular dynamics simulation study of the binding of purine bases to the aptamer domain of the guanine sensing riboswitch. *Nucleic Acids Res* (in press). doi: 10.1093/nar/gkp486.

Wickiser JK, Winkler WC, Breaker RR, Crothers DM. 2005a. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell* **18:** 49–60.

Wickiser JK, Cheah MT, Breaker RR, Crothers DM. 2005b. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **44:** 13404–13414.

Winkler WC, Nahvi A, Breaker RR. 2002a. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419:** 952–956.

Winkler WC, Cohen-Chalamish S, Breaker RR. 2002b. An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci* **99:** 15908–15913.

Winkler WC, Nahvi A, Sudarsan N, Barrick JE, Breaker RR. 2003. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol* **10:** 701–707.

Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428:** 281–286.