Molecular mechanism for preQ₁-II riboswitch function revealed by molecular dynamics

ASAMINEW H. AYTENFISU,¹ JOSEPH A. LIBERMAN,¹ JOSEPH E. WEDEKIND,¹ and DAVID H. MATHEWS^{1,2}

¹Department of Biochemistry & Biophysics and Center for RNA Biology, University of Rochester Medical Center, Rochester, New York 14642, USA

²Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, New York 14642, USA

ABSTRACT

Riboswitches are RNA molecules that regulate gene expression using conformational change, affected by binding of small molecule ligands. A crystal structure of a ligand-bound class II preQ₁ riboswitch has been determined in a previous structural study. To gain insight into the dynamics of this riboswitch in solution, eight total molecular dynamic simulations, four with and four without ligand, were performed using the Amber force field. In the presence of ligand, all four of the simulations demonstrated rearranged base pairs at the 3' end, consistent with expected base-pairing from comparative sequence analysis in a prior bioinformatic analysis; this suggests the pairing in this region was altered by crystallization. Additionally, in the absence of ligand, three of the simulations demonstrated similar changes in base-pairing at the ligand binding site. Significantly, although most of the riboswitch architecture remained intact in the respective trajectories, the P3 stem was destabilized in the ligand-free simulations in a way that exposed the Shine–Dalgarno sequence. This work illustrates how destabilization of two major groove base triples can influence a nearby H-type pseudoknot and provides a mechanism for control of gene expression by a fold that is frequently found in bacterial riboswitches.

Keywords: RNA structure; amber; molecular dynamics; riboswitch

INTRODUCTION

Riboswitches are RNA structures that regulate gene expression in response to the binding of a cognate ligand (Serganov and Nudler 2013). While the thymine pyrophosphate riboswitch is found in all domains of life, most known riboswitches are bacterial, where they typically are located in the 5' leader sequences of mRNA. Riboswitches consist of two functional domains, the aptamer and the expression platform. The aptamer comprises the portion of the riboswitch that is necessary and sufficient for ligand binding, whereas the expression platform contains the gene regulatory sequence. Ligand binding to the aptamer results in a conformational change in both the aptamer and expression platforms, affecting gene regulation. Most commonly, riboswitches regulate transcription or translation. For translational control, which is utilized by the preQ₁-II (class II) riboswitch described herein, ligand binding controls the accessibility of the Shine–Dalgarno (SD) sequence (Serganov and Patel 2007; Peselis and Serganov 2014).

PreQ₁ is an intermediate on the biosynthetic pathway of queuosine (Q), a hypermodified nucleotide in bacteria (Noguchi et al. 1982). Q is found in the wobble position of certain tRNAs (Kimura-Harada et al. 1972), where it confers translational fidelity (Bienz et al. 1981; Meier et al. 1985; Urbona-

vičius et al. 2001). The expression of genes responsible for the biosynthesis of Q or the import of Q precursors is frequently controlled by riboswitches. To date, descriptions of three phylogenetically distinct classes of preQ₁-binding riboswitches, known as the class I (preQ₁-I), class II (preQ₁-II), or class III (preQ₁-III) have been published (Roth et al. 2007; Meyer et al. 2008; McCown et al. 2014). The class II riboswitch is located in the 5' leader sequence of the COG4708 (also known as *queT*) gene, which is hypothesized to be a transporter of Q precursors (Meyer et al. 2008). Upon ligand binding, this riboswitch buries the SD sequence in an H-type (HL_{out}) pseudoknot, and this conformation is hypothesized to attenuate translation by preventing binding by the ribosome. High-resolution structures have been determined for the preQ₁-I, preQ₁-II, and preQ₁-III translational riboswitches in the ligand-bound state (Spitale et al. 2009; Jenkins et al. 2011; Liberman et al. 2013, 2015; Kang et al. 2014). The structures of the apo state, however, have only been determined for the preQ₁-I riboswitch (Jenkins et al. 2011), leaving open the question of how the preQ₁-II riboswitch changes its conformation in the absence of the preQ₁ ligand.

© 2015 Aytenfisu et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Corresponding author: david_mathews@urmc.rochester.edu

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.051367.115>.

Molecular dynamics (MD) has proven valuable in providing insights into many aspects of riboswitch structure and function. More specifically, MD has elucidated chemical features that stabilize ligand binding, the interactions that stabilize conformational endpoints, and features that facilitate conformational switching. Studies have been conducted on the SAM-I riboswitch (Huang et al. 2009; Stoddard et al. 2010; Doshi et al. 2012; Hayes et al. 2012; Huang et al. 2013; Xue et al. 2015), the SAM-II riboswitch (Kelley and Hamelberg 2010; Doshi et al. 2012), the guanine riboswitch (Villa et al. 2009; Sund et al. 2014), the adenine riboswitch (Sharma et al. 2009; Priyakumar and MacKerell 2010; Nozinovic et al. 2014; Sund et al. 2014), the GlmS riboswitch (Banáš et al. 2010; Xin and Hamelberg 2010), and the preQ₁-I and preQ₁-III riboswitches (Petrone et al. 2011; Banáš et al. 2012; Eichhorn et al. 2012; Gong et al. 2014; Liberman et al. 2015). Overall, these studies show the feasibility of using MD to understand riboswitch function and provide an opportunity to look for common themes in ligand binding and conformational dynamics. A review of MD and other computational approaches applied to understanding riboswitch structure and function is available (Sanbonmatsu 2014).

This paper reports an investigation of the conformations adopted by the preQ₁-II riboswitch in the apo and bound states using MD, with a starting model derived from the crystal structure of the ligand-bound preQ₁-II riboswitch determined by Liberman et al. (2013). This analysis was performed to gain a deeper understanding of how this riboswitch behaves in solution, free of the crystal lattice. It is notoriously difficult to obtain crystal structures of riboswitches in the ligand-free state, and this is likely because the apo state is more conformationally heterogeneous and dynamic (Stoddard et al. 2010; Liberman and Wedekind 2012). Therefore, MD simulations provide a tractable method for generating atomic-resolution models of these large RNA molecules. Four MD simulations of 1.3 μ sec each were conducted for both the ligand-bound and apo states, providing a cumulative sampling time of 10.4 μ sec. The ligand-bound simulations underwent a rearrangement of the structure to one predicted by comparative sequence analysis (Meyer et al. 2008), suggesting crystal packing-induced distortions were present in the crystal structure. The apo simulations provided insight into the pathway for rearrangement of the ligand binding regions and the exposure of the SD sequence, although the SD sequence was not fully exposed on the time scale of these simulations. The results support the hypothesis that ligand binding enhances the stability of the pseudoknot, sequestering the SD sequence from interaction with the ribosome.

RESULTS

Starting configuration of the preQ₁-II riboswitch crystal structure

The starting model for the simulations was the crystal structure of the ligand-bound *Lactobacillus rhamnosus* preQ₁-II

riboswitch (Liberman et al. 2013). It contains 77 nt in four stems (P1, P2, P3, and P4), two hairpin loops (L1 and L2), and three junctions (J_{2-3} , J_{2-4} , J_{3-4}) (Fig. 1). The SD sequence is base paired (pseudoknotted) in stem P3. The ligand binding pocket of the preQ₁-II riboswitch lies within a three-way helical junction composed of helices P2, P3, and P4 (Liberman et al. 2013). Within the ligand binding pocket are two major groove base triple interactions, A71-U40•U31 and A72-U39•U32, which form a “floor” under the ligand. The ligand itself participates in another triple interaction, C30•preQ₁•U41. At the “bottom” of P3, the crystal structure demonstrated unpaired nucleotides, U34 and A77 (with A77 unstacked from the helical apex), adjacent to a single hydrogen bond pair of U35 and G76. Here, MD analysis was used to interrogate whether these nucleotides form standard base pairs in the absence of crystal packing (Supplemental Fig. S1; Liberman et al. 2013).

Comparative analysis of global and local stability pinpoints ligand-bound and apo structure flexibility

Four apo and four ligand-bound molecular dynamics simulations were performed for 1.3 μ sec of simulation time each. The stability of the simulations was tested by following RMSD of atomic coordinates compared to the crystal structure as a function of time. For the ligand-bound structure, all helices present in the crystal structure were preserved during the simulation, although some details in the pairing changed (see below). The RMSD for each of the stems in the structure remained below 2 Å for nearly the entire simulation time. Supplemental Table S1 has a summary of stem heavy-atom RMSDs (as minimum, maximum, average, and standard deviation from the mean) compared to the crystal structure for each stem, and the RMSDs are plotted in Supplemental Figures S2–S5.

The binding pocket, comprising C30, U31, U32, U41, and A70 was stable in the ligand-bound simulations, as demonstrated by the mass-weighted, heavy-atom RMSD values, which remained under 2 Å during the simulations (left column of Fig. 2). In contrast, the apo simulations indicated a more flexible binding pocket, with a mass-weighted, heavy-atom RMSD as high as 6.4 Å (right column of Fig. 2).

Atomic fluctuation across the entire production run was calculated and averaged over each nucleotide. The most dynamic nucleotides are in the loop regions, which correlated well with the crystallographic B-factors (Fig. 3). For both the ligand-bound and apo simulations, the nucleotides that constitute the ligand binding pocket were found to be the most rigid part of the riboswitch. Figure 4 shows the atomic fluctuation plotted on the structure for both ligand-bound and apo states as an average over all four simulations. For comparison, the root mean square fluctuation (RMSF) (Henzler-Wildman and Kern 2007) estimated from the crystallographic B-factors was also plotted. The RMSFs from the simulation differ somewhat from the RMSFs estimated

FIGURE 1. Summary of simulations using secondary structure diagrams showing mapped long-range interactions. The *top left* structure is the crystal structure used as the starting conformation for all simulations. All four ligand-bound simulations converged to a common structure shown in the *top right* corner. Three of the apo simulations showed the same altered structure of the binding pocket (apo simulations 2, 3, and 4). This binding pocket includes nucleotides C30, U31, U32, U41, and A70. Apo simulation 1, however, has G76 and U34 stacked at the end of P3. The noncanonical pairs are annotated according to the standard Leontis–Westhof classification (Leontis et al. 2002; Zirbel et al. 2009). The SD sequence is highlighted in yellow. The base triple interactions that are present in ligand-bound simulations are absent in the apo simulations.

from the B-factors because of crystal contacts (Figs. 3, 4). In particular, L1 has strong contacts with J₂₋₄/P4 from an adjacent structure in the crystal lattice (Liberman et al. 2013). As such, L1 has a lower crystallographic B-factors than would be predicted by the significant atomic fluctuations of both ligand-bound and apo simulations. Notably, a difference was observed between ligand-bound and apo simulations in terms of the atomic fluctuations of nucleotides in the binding pocket, or close to the binding pocket. These nucleotides have lower atomic fluctuation for all ligand-bound simulations compared to apo simulations (Fig. 3).

MD supports the formation of an additional, expected pair in the P3 terminal stem

The mode of gene regulation for translational riboswitches entails complete or partial sequestration of the SD sequence upon ligand binding (Breaker 2012). Although the entire SD sequence is observed to participate in P3 formation (Fig. 1), the crystal structure lacked an adjacent base pair between A77 and U34 at the terminus of P3, which had been predicted by comparative analysis (Meyer et al. 2008). Additionally, there is significant disorder seen in the G76•U35 base pair.

FIGURE 2. PreQ₁-II riboswitch binding pocket mass-weighted, heavy-atom RMSD to starting structure as function of time. The *left* column shows four ligand-bound simulations and the *right* column shows four apo simulations. The binding pocket is defined as nucleotides C30, U31, U32, U41, and A70.

This discrepancy was likely the result of crystal packing, which prompted further analysis by MD. All four ligand-bound simulations consistently showed the emergence of Watson–Crick pairing for A77–U34 and wobble pairing for G76•U35. In contrast, the crystal structure showed a *trans*-Hoogsteen/sugar-edge U35•G76 pair (Figs. 1, 5). The switch to a wobble interaction of G76•U35 occurred in <400 nsec from the start in all four ligand-bound simulations, and the wobble interaction remained for the rest of the simulation time. In contrast, the formation of the pair between A77 and U34 was nearly immediate (in the first few nanoseconds of simulations 1–3 and by 200 nsec in simulation 4). In all ligand-bound simulations, there was subsequent transient opening of the A77–U34 pair. The emergence of these pairs can be followed by the atomic distances between N3 of U34 and N1 of A77 (average distance of 5.42 Å in the pair) and the distances between N3 of U35 and O6 of G76 (average distance of 5.10 Å in the wobble pair; Fig. 6).

Conformational change of the ligand binding pocket in apo simulations

Removal of the ligand for apo simulations led to altered pairing in the binding pocket. The triple interactions that were found in the crystal structure (A71–U40•U31 and A72–U39•U32) rearranged to form three new base pairs. The first change was the formation of an A29•A43–U28 triple interaction with an A43–U28 Watson–Crick base pair and a *trans*-Hoogsteen/sugar-edge A29•A43 interaction (Fig. 7). Stacked on this triplex was a new Watson–Crick C30–G42 base pair. This pair was followed by a Watson–Crick/Watson–Crick noncanonical pair of U31•U41. The resulting changes closed the preQ₁ pocket to ligand binding, and all three of the new pairings were present in three of the four apo simulations.

These rearrangements occurred within the initial 300–500 nsec of the simulation (right side of Figs. 2, 8).

Conformational changes during the simulation were monitored by distance and solvent accessible surface area (SASA) (Figs. 8, 9). The distance between N6 of A29 and N3 of A43 was measured to show the emergence of the *trans*-Hoogsteen/sugar-edge A29•A43 interaction during the simulations. The distance between N6 of A29 and N3 of A43 of ~6 Å corresponded to a Watson–Crick G42•A29 noncanonical interaction, as observed in the crystal structure in the binding pocket ceiling; the distance of ~2.8 Å corresponded to the *trans*-Hoogsteen/sugar-edge A29•A43 interaction that emerged in the apo simulations. The distance stayed close to the starting conformation at 6 Å in the ligand-bound simulations. The distances between N3 of C30 and N1 of G42 and between N3 of U31 and N3 of U41 were also measured (Fig. 8) to detect the emergence of the G–C (Watson–Crick; at 2.9 Å) and U•U base-pairing (at 3.4 Å), respectively, in the apo simulations. These distances also remained stable, i.e., at roughly the distances observed in the crystal structure, for the ligand-bound simulations, where these base pairs were not observed.

Monitoring disruption of the pairing of the Shine–Dalgarno sequence in the apo simulations

While the simulation of the ligand-bound riboswitch revealed additional base pairs forming in P3, the apo simulation showed the dissolution of base pairs in P3. To measure the extent to which pairs were disrupted in apo and ligand-bound simulations, distances were monitored for four base pairs in P3. The four distances were between N3 of U34

FIGURE 3. RMSF averaged over each nucleotide. Apo simulations are plotted with blue lines and ligand-bound simulations are plotted in red lines. The fluctuation data (black line) estimated from the temperature factors of the crystal structure using Equation 1 is also included for comparison. Ligand binding regions are boxed in green rectangles. The x-axis is color-coded according to the secondary structure regions shown in Figure 1.

FIGURE 4. RMSF of the ligand-bound simulations (*top*) and apo simulations (*middle*) from MD trajectories. Each RMSF was plotted on a common structure (from 830 nsec in ligand-bound simulation 3) to facilitate comparison. The fluctuation data (*bottom*) estimated from the temperature factor of the crystal structure using Equation 1 is also included for comparison. The scaling is the same as Figure 3, the RMSF 1D plot.

and N1 of A77, between N3 of U35 and O6 of G76, between N3 of U36 and N1 of A75, and between N3 of U39 and N1 of A72 (Fig. 6). These four distances showed how the absence of ligand destabilized the P3 terminal base pair within a region that includes the SD sequence (summarized in Fig. 1). In all apo simulations, pairs U34-A77 and U35-G76 are not observed. Simulations 1 and 4 demonstrated a loss of the pair U36-A75. In simulation 4, the U39-A72 pair, which is Watson-Crick in the crystal structure, rearranges to a *cis*-Watson-Crick/sugar-edge pair.

DISCUSSION

Molecular dynamics simulations were performed for the preQ₁-II riboswitch from *L. rhamnosus* in ligand-bound and apo states. The goal of this investigation was to provide insight into how the ligand-bound riboswitch behaves in solution, without the confines of a crystal lattice. An additional goal was to understand the conformation adopted by the apo riboswitch, which is expected to have an exposed SD sequence that is

competent to base pair with the 3' end of the 16S rRNA within the small ribosomal subunit as a preface to translation initiation. Such pairing is only possible in the “gene on” conformational state of the *queT* mRNA transcript.

First, the crystal structure did not show pairing between U34 and A77, and the observed U35•G76 pairing was not a canonical wobble interaction (Liberman et al. 2013), despite the fact that the secondary structure derived from comparative analysis suggested such pairing would occur (Meyer et al. 2008). Additionally, the NMR structure of the *Streptococcus pneumoniae* preQ₁-II riboswitch reveals canonical C-G Watson-Crick base-pairing at the location equivalent to U35•G76 in accord with these MD results; in contrast, the U34-A77 position is replaced by a buckled A•G pair that is not oriented properly for hydrogen bonding in the lowest energy NMR structures (Kang et al. 2014). In all four of the ligand-bound MD simulations described here, a rearrangement of U35•G76 to a canonical wobble pair followed by an A77-U34 canonical Watson-Crick pairing occurred by 400 nsec (Fig. 6). As a result, the SD sequence becomes more sequestered with canonical pairs. These observations are consistent with the mode of translational attenuation reported for comparable translational riboswitches containing an H-type pseudoknot that sequesters the SD sequence (Serganov and Patel 2007), such as SAM-II (Gilbert et al. 2008) in which the entire ribosome-binding site is buried, and the preQ₁-I riboswitch in which the first two or three bases of the SD sequence are buried in helices equivalent to P3 (Jenkins et al. 2011). The SAH riboswitch, an LL-type pseudoknot, also sequesters its entire SD sequence for translational control (Batey 2011).

The differences in the four apo simulations (Fig. 1) clearly indicate that the simulations are not long enough for the structures to equilibrate. This is not surprising, as dinucleotide stacking and unstacking both have rate constants on the order of 10^7 s⁻¹ (Turner 2000). A number of these types of events need to occur for the rearrangement of the structure to the apo form.

FIGURE 5. Representative structures of the P3 terminus. (A) The crystal structure. (B) The consensus structure of the ligand-bound simulations, selected from simulation number 3 at 730 nsec.

FIGURE 6. Distance between atoms in P3 nucleotides. The *left* panel corresponds to ligand-bound simulations and the *right* panel corresponds to the apo simulations. Four distances are color annotated and these distances each correspond to a base pair, as shown in the key. In many simulations, the plot of the N3 of U36 to N1 of A75 distance (green) is obscured by the N3 of U39 to N1 of A72 distance (purple), and is therefore not visible.

Although the simulation time is insufficient to observe the complete relaxation of the structure to a final conformation, insights can be obtained from the simulations from the ligand binding pocket changes in the apo structure. Consistent trends were observed that suggest the expression platform is more solvent accessible in the apo state. The two major groove base triples, U32•A72-U39 and U31•A71-U40, that stack directly on the ligand were lost, thus opening a pocket close to the SD sequence. Additionally, loss of the ligand in the preQ₁-II riboswitch is expected to have the immediate effect of disrupting the coaxial helical stack of P2 and P3, further destabilizing the pseudoknot that controls SD accessibility as suggested (Jenkins et al. 2011; Liberman et al. 2013). The simulations reported here confirmed this hypothesis, while also showing this pocket region rearranges to a similar conformation in three of four simulations that is different from the ligand-bound structure. This conformation included a new Watson–Crick pair (G42 to C30), a Watson–Crick/Watson–Crick noncanonical pair between U31 and U41, and a triple interaction between A29, A43, and U28 at the ligand binding pocket. The SASA was plotted for the binding pocket with the ligand-bound simulations as a control to show the conformational switch (Fig. 9). Lower SASA corresponded to nucleotides that were buried in interactions, and SASA above the ligand-bound control indicated exposure to solvent (Petrone

et al. 2011). The nucleotides from the base pairs, U31-U41 and C30-G42, were less accessible to solvent in apo simulations than ligand-bound simulation, whereas nucleotide A29 showed more exposure to solvent. These nucleotides are parts of the binding pocket and, overall for apo simulation, the binding pocket became less solvent accessible. This indicated the binding pocket closed in the apo simulations. The unique apo simulation (simulation 1) exhibited a cross-strand stack between A77 and U36 at the 3' terminus that formed at 350 nsec and persisted for nearly 1 μ sec of the simulation (Fig. 7). A distance of 4.36 Å average between the center of mass of the pyrimidine ring of A77 and U36 was observed in apo simulation 1. This distance indicated that the structure entered a local free energy minimum that MD cannot escape during a simulation of this duration. The observation of at least one “open pocket” conformation during MD is consistent with the notion that some fraction of apo preQ₁-II riboswitches must be receptive to ligand binding for gene regulation. Conversely, the “closed” pockets observed here would need to change conformation for effector recognition (Liberman and Wedekind 2012).

Aside from the ligand binding regions, the apo simulations also demonstrated differences compared to the ligand-bound simulations in terms of the 3' stem nucleotides. In the 1.3 μ sec simulations, two base pairs (U34-A77 and G76-U35) were disrupted for all apo simulations, although the final disposition of the nucleotides was variable. According to NMR experiments (Kang et al. 2014), the SD sequence is not fully exposed in the apo state, and the simulation results here were consistent with this. In the presence of preQ₁, the SD sequence is fully sequestered in the NMR model structure. Evidence about the apo structure is provided by the intensity of hydrogen bonding in the P3 regions, which is shown to be reduced in the absence of preQ₁ (Kang et al. 2014). All ligand-bound simulations here demonstrated the SD sequence was fully sequestered, whereas all apo simulations here

FIGURE 7. The structure of the binding pocket with ligand-bound (*left* panel) and in apo simulations (*right* panel). The apo structure (*right* panel) was the consensus for three of four of the apo simulations. Nucleotides A29 and A43 formed a *trans*-Hoogsteen/sugar-edge A29•A43 pair. At the *top*, a noncanonical, Watson–Crick face U31•U41 base pair (Leontis et al. 2002) and, in the *middle*, a C30-G42 Watson–Crick base pair formed. This structure was taken from simulation number 3 at 560 nsec. For comparison, the ligand-bound binding pocket is in the *left* panel with the ligand drawn in green from simulation number 3 at 730 nsec.

FIGURE 8. Distances between base atoms in the binding pocket. The distance between N6 of A29 and N3 of A43 (red dots) of ~ 6 Å corresponds to a Watson–Crick G42•A29 noncanonical interaction, as observed in the crystal structure in the binding pocket ceiling; the distance of ~ 2.8 Å corresponded to a *trans*-Hoogsteen/sugar-edge A29•A43 interaction that emerged in the apo simulations. Simulations 2, 3, and 4 of the apo structure showed similar trends, but larger fluctuations were observed in simulation 1. Three distances are color-coded as shown in Figure 6 legend. These distances were selected to monitor new base pairs that were observed in the apo simulations.

demonstrated a partially exposed SD sequence. In-line probing, SHAPE and single molecule FRET experiments also show that P3 is more exposed to solvent in the absence of preQ₁ (Meyer et al. 2008; Liberman et al. 2013; Soulière et al. 2013).

Although the SD sequence does not become fully unpaired in the simulations here, it is conceivable that full exposure of the SD sequence is not required for ribosome binding and translation. Studies on hybridization kinetics found that lengthening of toeholds (nucleotides in a target sequence that are available for pairing) up to 5 nt can dramatically accelerate the kinetics of hybridization (Srinivas et al. 2013; Šulc et al. 2015). For each unpaired nucleotide of additional toehold up to 5 nt, the hybridization rate increases by approximately an order of magnitude (Šulc et al. 2015). The changes in apo structure observed in these simulations would explain one to two orders of magnitude of faster ribosome-binding kinetics. Further equilibration of the apo structure might reveal more exposure of the SD sequence, which would explain even faster kinetics for expression of the COG4708 gene in the absence of preQ₁.

Overall, this work demonstrates that simulation can complement structural studies to provide novel insights into RNA structure. The simulations provided additional information about the solution structure and dynamics that are not available in the crystal structure because of crystal contacts. The simulations, starting from the crystal structure, also provide information not provided by the solution structure, which is limited by the number of spatial restraints. Therefore, the integration of crystallography, NMR, single molecule tech-

niques, chemical mapping data, and simulation provides a more complete understanding of the preQ₁-II riboswitch. The observations and method utilized here should be broadly applicable to other RNA regulatory motifs.

MATERIALS AND METHODS

The preQ₁-II riboswitch atomic coordinates were taken from the X-ray structure at an advanced stage of crystallographic refinement (subsequently PDB accession number 4JF2) (Liberman et al. 2013). The atomic model was used as the starting structure for both apo and ligand-bound simulations. For all simulations, the Amber 11 or Amber 12 software packages were used with the amber ff10 force field, which is unchanged for nucleic acids in ff12SB (Cornell et al. 1995; Cheatham et al. 1999; Pérez et al. 2007; Joung and Cheatham 2008; Zgarbová et al. 2011). In the undeposited structure, five Mg²⁺ ions were modeled, but one of the Mg²⁺ ions was changed to solvent prior to the final PDB deposition because of its high B-factors and lack of a complete coordination sphere. Nonetheless, the assigned waters and all five original Mg²⁺ ions were preserved in the model for simulation. Fifteen Cs⁺ ions were replaced with Na⁺. The parameters for monovalent salts were those by Joung and Cheatham (2008) and the parameters for Mg²⁺ were Amber-adapted Åqvist parameters (Åqvist 1990).

The ligand was removed to obtain an initial apo structure. The apo structure was solvated with the TIP3P water model (Jorgensen et al. 1983) in an isometric box with a 10 Å margin in all directions. It was necessary to add an additional 51 Na⁺ ions to neutralize the system. Finally, 65 Na⁺ and 65 Cl[−] ions were added in order to include 0.1 M NaCl salt, based on the number of waters. The whole system consisted of 122,958 atoms in a box of dimensions 113 Å × 113 Å × 113 Å.

To generate the ligand-bound system, ligand charge and force field parameters were prepared. First hydrogens were added to the ligand with the “reduce” program found in AmberTools (Supplemental Fig. S7; Case et al. 2012). Then geometry was optimized with the RED RESP server (Bayly et al. 1993; Dupradeau et al.

FIGURE 9. Solvent accessible surface area (SASA) per nucleotide, averaged over the last 1.2 μ sec of production simulation time. For this analysis, the ligand-bound simulations were controls in order to identify conformational switching in the apo simulations. Nucleotide A29 became more exposed and nucleotides C30, G42, U31, and U41 formed new base pairs, as shown by the lower SASA compared to the ligand-bound simulations. The error bars indicate the standard error of the mean for the four simulations. The results of the individual simulations are provided as Supplemental Figure S6.

2010; Vanqualef et al. 2011) with an HF/6-31G* quantum calculation. Atomic point charges were also obtained from the RED RESP server with the same level of quantum theory as the geometry optimization (Supplemental Table S2). Other force field parameters were obtained using the “antechamber” program in Amber (Supplemental Table S3; Case et al. 2005). The ligand-bound system was also solvated in an isometric box with a 10 Å margin. Fifty Na⁺ ions were added to neutralize the system, and then 65 Na⁺ and 65 Cl[−] were added to include 0.1 M NaCl (Joung and Cheatham 2008, 2009). The whole system comprises 121,145 atoms in a 112 Å × 112 Å × 112 Å box.

Simulations

Initially the systems were subjected to a two-stage energy minimization. In the first stage, the RNA molecule was held fixed in space using a positional restraint with a force constant of 500 kcal/(mol × Å²), and the solvent and counterions were energy-minimized for 5000 steps. In the second stage, both solvent and RNA were energy-minimized for 5000 steps. Each minimization had 2500 steps of steepest descent, followed by 2500 steps of conjugate gradient minimization. The final minimized systems were subjected to 100 psec of heating from 0 K to 300 K in the NVT ensemble by holding the RNA in space with a positional restraint with force constant of 10 kcal/(mol × Å²). Finally, 100 psec of NPT ensemble simulation was performed to equilibrate at 300 K. The Langevin thermostat with a collision frequency of 1000/psec^{−1} was used for temperature control for both NVT and NPT simulations. The particle-mesh Ewald (PME) (Toukmaji et al. 2000; Sagui et al. 2004) method was used to calculate long-range electrostatics, and a 10 Å cutoff was used for the direct space sum. SHAKE (Miyamoto and Kollman 1992; Ryckaert et al. 1977) was used to restrain all bonds involving hydrogen. The PMEMD and PMEMD.CUDA programs were used for the simulations (Case et al. 2005, 2012; Götz et al. 2012; Le Grand et al. 2013; Salomon-Ferrer et al. 2013).

Analyses

Analyses of trajectories were performed by Ptraj and Cpptraj from AmberTools (Case et al. 2012). The VMD (Humphrey et al. 1996) and PyMOL (DeLano 2002) programs were used for visualization and image preparation. Solvent accessible surface area (SASA) was calculated using the program “surf” implemented in AmberTools under Cpptraj (Case et al. 2012) with the default settings. Using Ptraj (Case et al. 2012), the RNA flexibility was calculated on a residue basis using RMSF (Henzler-Wildman and Kern 2007) of the position of heavy atoms. The RMSF calculation was done after trajectories were fitted to the average structure. Atomic fluctuation for the crystal was estimated from the temperature factor according to

$$\text{RMSF} = \sqrt{\frac{3\beta}{8\pi^2}}, \quad (1)$$

where β is the temperature factor or B-factor of the crystal and RMSF (Henzler-Wildman and Kern 2007) is the root mean of isotropic atomic fluctuations.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants R01 GM076485 to D.H.M. and R01 GM063162 to J.E.W. Computer time was provided by the University of Rochester Center for Integrated Research Computing.

Received February 10, 2015; accepted July 2, 2015.

REFERENCES

- Åqvist J. 1990. Ion-water interaction potentials derived from free energy perturbation simulations. *J Phys Chem* **94**: 8021–8024.
- Banáš P, Walter NG, Šponer J, Otyepka M. 2010. Protonation states of the key active site residues and structural dynamics of the *glmS* riboswitch as revealed by molecular dynamics. *J Phys Chem B* **114**: 8701–8712.
- Banáš P, Sklenovský P, Wedekind JE, Šponer J, Otyepka M. 2012. Molecular mechanism of preQ₁ riboswitch action: a molecular dynamics study. *J Phys Chem B* **116**: 12721–12734.
- Batey RT. 2011. Recognition of S-adenosylmethionine by riboswitches. *Wiley Interdiscip Rev RNA* **2**: 299–311.
- Bayly CI, Cieplak P, Cornell W, Kollman PA. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* **97**: 10269–10280.
- Bienz M, Kubli E. 1981. Wild-type transfer-RNA GTyr reads the TMV RNA stop codon, but Q-base modified transfer-RNA GTyr does not. *Nature* **294**: 188–190.
- Breaker RR. 2012. Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol* **4**: a003566.
- Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. 2005. The Amber biomolecular simulation programs. *J Comput Chem* **26**: 1668–1688.
- Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, et al. 2012. *AmberTools 1.3*. <http://ambermd.org/>.
- Cheatham TE III, Cieplak P, Kollman PA. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J Biomol Struct Dyn* **16**: 845–862.
- Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* **117**: 5179–5197.
- DeLano WL. 2002. *The PyMOL molecular graphics system*. DeLano Scientific, San Carlos, CA. <http://www.pymol.org>.
- Doshi U, Kelley JM, Hamelberg D. 2012. Atomic-level insights into metabolite recognition and specificity of the SAM-II riboswitch. *RNA* **18**: 300–307.
- Dupradeau FY, Pigache A, Zaffran T, Savineau C, Lelong R, Grivel N, Lelong D, Rosanski W, Cieplak P. 2010. The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys Chem Chem Phys* **12**: 7821–7839.
- Eichhorn CD, Feng J, Suddala KC, Walter NG, Brooks CL III, Al-Hashimi HM. 2012. Unraveling the structural complexity in a single-stranded RNA tail: implications for efficient ligand binding in the prequeuosine riboswitch. *Nucleic Acids Res* **40**: 1345–1355.
- 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.
- Gong Z, Zhao Y, Chen C, Duan Y, Xiao Y. 2014. Insights into ligand binding to preQ₁ riboswitch aptamer from molecular dynamics simulations. *PLoS ONE* **9**: e92247.
- Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. 2012. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J Chem Theory Comput* **8**: 1542–1555.

- Hayes RL, Noel JK, Mohanty U, Whitford PC, Hennelly SP, Onuchic JN, Sanbonmatsu KY. 2012. Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. *J Am Chem Soc* **134**: 12043–12053.
- Henzler-Wildman K, Kern D. 2007. Dynamic personalities of proteins. *Nature* **450**: 964–972.
- Huang W, Kim J, Jha S, Aboul-ela F. 2009. A mechanism for S-adenosyl methionine assisted formation of a riboswitch conformation: a small molecule with a strong arm. *Nucleic Acids Res* **37**: 6528–6539.
- Huang W, Kim J, Jha S, Aboul-ela F. 2013. The impact of a ligand binding on strand migration in the SAM-I riboswitch. *PLoS Comput Biol* **9**: e1003069.
- Humphrey W, Dalke A, Schulten K. 1996. VMD—visual molecular dynamics. *J Molec Graphics* **14**: 33–38.
- Jenkins JL, Krucinska J, McCarty RM, Bandarian V, Wedekind JE. 2011. Comparison of a preQ₁ riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J Biol Chem* **286**: 24626–24637.
- 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.
- Joung IS, Cheatham TE III. 2008. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J Phys Chem B* **112**: 9020–9041.
- Joung IS, Cheatham TE III. 2009. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J Phys Chem B* **113**: 13279–13290.
- Kang M, Eichhorn CD, Feigon J. 2014. Structural determinants for ligand capture by a class II preQ₁ riboswitch. *Proc Natl Acad Sci* **111**: E663–E671.
- Kelley JM, Hamelberg D. 2010. Atomistic basis for the on-off signaling mechanism in SAM-II riboswitch. *Nucleic Acids Res* **38**: 1392–1400.
- Kimura-Harada F, Harada F, Nishimura S. 1972. The presence of N-[9-(*c*-D-ribofuranosyl)purin-6-ylcarbamoyl] threonine in isoleucine, threonine and asparagine tRNAs from *Escherichia coli*. *FEBS Lett* **21**: 71–74.
- Le Grand S, Götz AW, Walker RC. 2013. SPFP: speed without compromise—a mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* **184**: 374–380.
- Leontis NB, Stombaugh J, Westhof E. 2002. The non-Watson–Crick base pairs and their associated isosteric matrices. *Nucleic Acids Res* **30**: 3497–3531.
- Lieberman JA, Wedekind JE. 2012. Riboswitch structure in the ligand-free state. *Wiley Interdiscip Rev RNA* **3**: 369–384.
- Lieberman JA, Salim M, Krucinska J, Wedekind JE. 2013. Structure of a class II preQ₁ riboswitch reveals ligand recognition by a new fold. *Nat Chem Biol* **9**: 353–355.
- Lieberman JA, Suddala KC, Aytenfisu A, Chan D, Belashov IA, Salim M, Mathews DH, Spitale RC, Walter NG, Wedekind JE. 2015. Structural analysis of a class III preQ₁ riboswitch reveals an aptamer distant from a ribosome-binding site regulated by fast dynamics. *Proc Natl Acad Sci* **112**: E3485–E3494.
- McCown PJ, Liang JJ, Weinberg Z, Breaker RR. 2014. Structural, functional, and taxonomic diversity of three preQ₁ riboswitch classes. *Chem Biol* **21**: 880–889.
- Meier F, Suter B, Grosjean H, Keith G, Kubli E. 1985. Queuosine modification of the wobble base in tRNA^{His} influences in vivo decoding properties. *EMBO J* **4**: 823–827.
- Meyer MM, Roth A, Chervin SM, Garcia GA, Breaker RR. 2008. Confirmation of a second natural preQ₁ aptamer class in *Streptococcaceae* bacteria. *RNA* **14**: 685–695.
- Miyamoto S, Kollman PA. 1992. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem* **13**: 952–962.
- Noguchi S, Nishimura Y, Hirota Y, Nishimura S. 1982. Isolation and characterization of an *Escherichia coli* mutant lacking tRNA-guanine transglycosylase. *Function and biosynthesis of queuosine in tRNA*. *J Biol Chem* **257**: 6544–6550.
- Nozinovic S, Reining A, Kim Y-B, Noeske J, Schlepckow K, Wöhnert J, Schwalbe H. 2014. The importance of helix P1 stability for structural pre-organization and ligand binding affinity of the adenine riboswitch aptamer domain. *RNA Biol* **11**: 655–666.
- Pérez A, Marchán I, Svozil D, Sponer J, Cheatham TE III, Laughton CA, Orozco M. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys J* **92**: 3817–3829.
- Peselis A, Serganov A. 2014. Structure and function of pseudoknots involved in gene expression control: structure and function of pseudoknots. *Wiley Interdiscip Rev RNA* **5**: 803–822.
- Petrone PM, Dewhurst J, Tommasi R, Whitehead L, Pomerantz AK. 2011. Atomic-scale characterization of conformational changes in the preQ₁ riboswitch aptamer upon ligand binding. *J Mol Graphics Modell* **30**: 179–185.
- Priyakumar UD, MacKerell AD Jr. 2010. Role of the adenine ligand on the stabilization of the secondary and tertiary interactions in the adenine riboswitch. *J Mol Biol* **396**: 1422–1438.
- 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.
- Ryckaert J-P, Ciccotti G, Berendsen HJC. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **23**: 327–341.
- Sagui C, Pedersen LG, Darden TA. 2004. Towards an accurate representation of electrostatics in classical force fields: efficient implementation of multipolar interactions in biomolecular simulations. *J Chem Phys* **120**: 73–87.
- Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. 2013. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent Particle Mesh Ewald. *J Chem Theory Comput* **9**: 3878–3888.
- Sanbonmatsu KY. 2014. Dynamics of riboswitches: molecular simulations. *BBA Gene Regul Mech* **1839**: 1046–1050.
- Serganov A, Nudler E. 2013. A decade of riboswitches. *Cell* **152**: 17–24.
- Serganov A, Patel DJ. 2007. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet* **8**: 776–790.
- Sharma M, Bulusu G, Mitra A. 2009. MD simulations of ligand-bound and ligand-free aptamer: molecular level insights into the binding and switching mechanism of the *add* A-riboswitch. *RNA* **15**: 1673–1692.
- Soulière MF, Altman RB, Schwarz V, Haller A, Blanchard SC, Micura R. 2013. Tuning a riboswitch response through structural extension of a pseudoknot. *Proc Natl Acad Sci* **110**: E3256–E3264.
- Spitale RC, Torelli AT, Krucinska J, Bandarian V, Wedekind JE. 2009. The structural basis for recognition of the preQ₀ metabolite by an unusually small riboswitch aptamer domain. *J Biol Chem* **284**: 11012–11016.
- Srinivas N, Ouldrige TE, Šulc P, Schaeffer JM, Yurke B, Louis AA, Doye JPK, Winfree E. 2013. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res* **41**: 10641–10658.
- Stoddard CD, Montange RK, Hennelly SP, Rambo RP, Sanbonmatsu KY, Batey RT. 2010. Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure* **18**: 787–797.
- Šulc P, Ouldrige TE, Romano F, Doye JPK, Louis AA. 2015. Modelling toehold-mediated RNA strand displacement. *Biophys J* **108**: 1238–1247.
- Sund J, Lind C, Åqvist J. 2014. Binding site preorganization and ligand discrimination in the purine riboswitch. *J Phys Chem B* **119**: 773–782.
- Toukmaji A, Sagui C, Board J, Darden T. 2000. Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J Chem Phys* **113**: 10913–10927.

- Turner DH. 2000. Conformational changes. In *Nucleic acids* (ed. Bloomfield V, Crothers D, Tinoco I), pp. 259–334. University Science Books, Sausalito, CA.
- Urbonavičius J, Qian Q, Durand JM, Hagervall TG, Björk GR. 2001. Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J* **20**: 4863–4873.
- Vanquielef E, Simon S, Marquant G, Garcia E, Klimerak G, Delepine JC, Cieplak P, Dupradeau FY. 2011. R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res* **39**: W511–W517.
- Villa A, Wöhnert 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* **37**: 4774–4786.
- Xin Y, Hamelberg D. 2010. Deciphering the role of glucosamine-6-phosphate in the riboswitch action of *glmS* ribozyme. *RNA* **16**: 2455–2463.
- Xue X, Yongjun W, Zhihong L. 2015. Folding of SAM-II riboswitch explored by replica-exchange molecular dynamics simulation. *J Theor Biol* **365**: 265–269.
- Zgarbová M, Otyepka M, Šponer J, Mládek A, Banáš P, Cheatham TE III, Jurečka P. 2011. Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* **7**: 2886–2902.
- Zirbel CL, Šponer JE, Šponer J, Stombaugh J, Leontis NB. 2009. Classification and energetics of the base-phosphate interactions in RNA. *Nucleic Acids Res* **37**: 4898–4918.

RNA

A PUBLICATION OF THE RNA SOCIETY

Molecular mechanism for preQ₁-II riboswitch function revealed by molecular dynamics

Asaminew H. Aytenfisu, Joseph A. Liberman, Joseph E. Wedekind, et al.

RNA 2015 21: 1898-1907 originally published online September 14, 2015

Access the most recent version at doi:[10.1261/rna.051367.115](https://doi.org/10.1261/rna.051367.115)

Supplemental Material

<http://rnajournal.cshlp.org/content/suppl/2015/09/09/rna.051367.115.DC1>

References

This article cites 64 articles, 12 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/21/11/1898.full.html#ref-list-1>

Creative Commons License

This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *RNA* go to:

<http://rnajournal.cshlp.org/subscriptions>