



Published in final edited form as:

*J Phys Chem B*. 2012 October 25; 116(42): 12721–12734. doi:10.1021/jp309230v.

## Molecular Mechanism of preQ<sub>1</sub> Riboswitch Action: A Molecular Dynamics Study

Pavel Banáš<sup>1,2,‡</sup>, Petr Sklenovský<sup>1,‡</sup>, Joseph E. Wedekind<sup>4</sup>, Jiří Šponer<sup>1,2,3</sup>, and Michal Otyepka<sup>1,\*</sup>

<sup>1</sup>Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University Olomouc, tr. 17. listopadu 12, 771 46 Olomouc, Czech Republic

<sup>2</sup>Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic

<sup>3</sup>CEITEC – Central European Institute of Technology, Masaryk University, Campus Bohunice, Kamenice 5, 625 00 Brno, Czech Republic

<sup>4</sup>Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, 601 Elmwood Ave, Box 712, Rochester, NY 14620, USA

### Abstract

Riboswitches often occur in the 5'-untranslated regions of bacterial mRNA where they regulate gene expression. The preQ<sub>1</sub> riboswitch controls the biosynthesis of a hypermodified nucleoside queuosine in response to binding the queuosine metabolic intermediate. Structures of the ligand-bound and ligand-free states of preQ<sub>1</sub> riboswitch from *Thermoanaerobacter tengcongensis* were determined recently by X-ray crystallography. We used multiple, microsecond-long molecular dynamics simulations (29  $\mu$ s in total) to characterize the structural dynamics of preQ<sub>1</sub> riboswitches in both states. We observed different stability of the stem in the bound and free states, resulting in different accessibility of the ribosome-binding site. These differences are related to different stacking interaction between nucleotides of the stem and the associated loop, which itself adopts different conformations in the bound and free states. We suggest that the loop serves not only to bind preQ<sub>1</sub> but also transmits information about ligand binding from the ligand-binding pocket to the stem, which has implications for mRNA accessibility to the ribosome. We explain functional results obscured by a high salt crystallization medium and help to refine regions of disordered electron density, which demonstrates the predictive power of our approach. Besides investigating the functional dynamics of the riboswitch, we have also utilized this unique small folded RNA system for analysis of performance of the RNA force field on  $\mu$ s time scale. The latest AMBER parmbsc0 $\chi_{OL3}$  RNA force field is capable to provide stable trajectories of the folded molecule on  $\mu$ s timescale. On the other hand, force fields that are not properly balanced lead to significant structural perturbations on the sub- $\mu$ s time scale, which could easily lead inappropriate interpretation of the simulation data.

\*corresponding author: michal.otyepka@upol.cz; FAX: +420 585 634 761; Phone: +420 585 634 756.

‡Both authors contributed equally to this work and should be considered as first authors

**Supporting Information Available:** Supporting information file contains Supporting Methods and Results, Figures S1–S6, and Tables S1–S4. In Supporting Methods are also listed the AMBER prep-files of the four non-standard residues (i.e. preQ0, preQ1, preQ1<sup>+</sup>, and the N3-protonated cytosine). Figures S1–S6 illustrate topology of selected base triples with the N3-protonated cytosine, stability of PQA ligand binding pocket in simulations, stability of A-I/A-U interaction in simulations, P2 stem substates populated over simulations and values of GpU platform signature H-bonds and dihedrals. Tables S1–S4 contain a list of additional simulations, distances between ligand binding, planar bases in ligand-free simulations, a list of average distances of both BPh interactions in individual simulations, and radii of gyration of MD ensembles and X-ray structures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## Keywords

MD simulation; riboswitch; regulatory mechanism; queuosine; pseudoknot; GpU platform

## INTRODUCTION

Riboswitches are RNA structural elements that occur in all domains of life where they function to regulate gene expression during transcription, translation or splicing. The preponderance of riboswitches has been identified in eubacteria, where they typically occur in 5'-leader sequences of mRNAs.<sup>1</sup> Riboswitches usually consist of an aptamer domain that is responsible for ligand binding, and an expression platform that bears a terminator/anti-terminator stem-loop or a ribosome binding site (RBS). These RNA molecules act as “molecular switches”, since they adopt two mutually exclusive conformations that can lead to “gene on” and “gene off” states. So far, more than a dozen of riboswitch classes have been structurally and functionally characterized.<sup>2,3</sup> Riboswitches of each class reveal a unique aptamer domain architecture and sense specific ligands.<sup>4</sup> Some structural properties of ligand recognition, and the overall mechanism by which the expression of genes is controlled, are shared among riboswitch classes. In general, ligand-induced rearrangement of a riboswitch aptamer domain affects conformational behavior of its expression platform. When a terminator stem-loop is formed in the expression platform upon ligand binding – at the expense of the anti-terminator stem-loop – the transcription of a downstream gene is down-regulated. Similarly, when the RBS is sequestered into base pairing upon ligand binding, the subsequent inaccessibility of RBS inhibits translation initiation.<sup>5,6</sup>

Generally, riboswitch aptamer domains adopt complex and highly compact structures upon binding of their cognate ligands.<sup>4,7</sup> This has facilitated crystallographic studies, which has made riboswitches a valued source of three-dimensional structural data. At present, structures of more than 100 ligand-bound riboswitch aptamer domains have been resolved with X-ray crystallography. In contrast, only a few aptamer domains were structurally characterized without their cognate ligands,<sup>3</sup> due to difficulties associated with crystallization of the ligand-free riboswitch aptamers. This suggests higher flexibility of ligand-free aptamers leading to multiple conformational states.

Two classes of phylogenetically unrelated preQ<sub>1</sub> riboswitches have been found. They exist in only three eubacterial phyla (proteobacteria, fusobacteria and firmicutes) making this class of riboswitches less abundant than the others.<sup>1,8–10</sup> The preQ<sub>1</sub> riboswitch regulates biosynthesis of the hypermodified nucleoside queuosine by sensing its metabolic precursor 7-aminomethyl-7-deazaguanine (hereafter labeled preQ<sub>1</sub>; ref. <sup>11</sup>). PreQ<sub>1</sub> riboswitches have been shown to function at the level of transcription or translation, and in both cases, the binding of preQ<sub>1</sub> to the riboswitch aptamer attenuates gene expression leading to down-regulation of queuosine biosynthesis. A hallmark of the class I preQ<sub>1</sub> riboswitch is the small size of its aptamer domain, which comprises only 34 nucleotides (nts) and is thus the smallest of all known riboswitch classes.<sup>11</sup> During the last three years, the preQ<sub>1</sub> structure has been studied intensively by low- as well as high-resolution experimental methods seeking to unravel how the modest 34-nt architecture can achieve highly specific ligand binding. Rieder *et al.* suggested that preQ<sub>1</sub> riboswitch aptamer is organized into a pseudoknot structure upon ligand binding.<sup>12</sup> Nearly simultaneously, three independent high-resolution structural analyses demonstrated that the aptamer folds as an H-type pseudoknot.<sup>13–15</sup> This work revealed that the preQ<sub>1</sub> binding pocket is housed at the junction between the two pseudoknot stems and unmasked the binding pattern of the metabolite. X-ray structures of the only translationally-acting preQ<sub>1</sub> riboswitch from *Thermoanaerobacter tengcongensis* revealed that the aptamer partially overlaps the expression platform and

provided a glimpse of the two bases of the RBS in the ligand-bound and ligand-free states.<sup>13,16</sup>

Solution investigations using NMR in the study of Kang *et al.*<sup>15</sup> and a combined NMR-fluorescence approach by Rieder *et al.*<sup>12</sup> independently revealed that transcriptionally acting preQ<sub>1</sub> riboswitch lacks the pseudoknotted architecture in its ligand-free state and that binding of the ligand triggers folding of the riboswitch aptamer (i.e. ligand-induced folding). In contrast, in the most recent X-ray structure of the translational preQ<sub>1</sub> riboswitch from *Thermoanaerobacter tengcongensis*, the aptamer adopts a compact, near-pseudoknot conformation in its ligand-free state, implying that this riboswitch becomes pre-folded prior to ligand binding.<sup>16</sup> The RBS is sequestered in the X-ray structure of the ligand-free aptamer, which may partly reflect stabilization of P2 stem coming from the coaxial stacking in the X-ray structure. The fact that the apo structure is compact is supported by accompanying small angle scattering data but efforts to fit the apo coordinates to the experimental scattering profile show that the crystal structure was inconsistent with the ligand-free state in solution.

Because the translational, class I preQ<sub>1</sub> riboswitch represents the only preQ<sub>1</sub> riboswitch to be characterized crystallographically in the ligand-bound and free states<sup>16</sup> we choose it to investigate remaining questions about its mechanism of action. First, how does ligand presence or absence in the binding pocket affect the accessibility of the ribosome-binding site? Second, how significantly does crystal packing affect the aptamer structure in both ligand-bound and unbound states? The small size of the preQ<sub>1</sub> riboswitch aptamer makes this system attractive for molecular dynamics (MD) studies. Properly applied simulations can provide unique information on an atomic scale with sub-picosecond time resolution, which can complement existing experimental data.<sup>17</sup> Recent MD studies on various RNA systems helped to identify dynamic features of specific RNA structural motifs and folded RNA molecules.<sup>18–42</sup> It should be stressed that the quality of MD simulations is limited seriously by the quality of the empirical force field employed. Therefore, some RNA simulation results reported earlier in the literature may reflect force field limitations rather than real properties of the molecules being studied. This was demonstrated clearly by the occurrence of senseless “ladder-like” structures in short A-RNA duplexes, which appeared during long simulations with the AMBER force field *ff99*<sup>43,44</sup> on various RNA systems, such as the hairpin ribozyme,<sup>23</sup> GNRA tetraloops,<sup>45</sup> and reverse kink-turns.<sup>19</sup> Similarly, some difficulties have also been identified with the CHARMM27 force field,<sup>46</sup> such as melting of GNRA and UNCG tetraloops and apparent instability of RNA stems.<sup>45,47–49</sup> In addition, explicit solvent model and ion conditions may affect the simulations, albeit less than the solute force field.<sup>19,50,51</sup> These findings underscore the caveats of MD simulations, and suggest that the results should be interpreted with care, keeping in mind potential limitations, as well as how the results account for experimental information.

Herein, we investigated the structural dynamics of the translational, class I preQ<sub>1</sub> riboswitch from *Thermoanaerobacter tengcongensis* (Figure 1) using unrestrained, conventional MD simulations in explicit solvent. The extensive set of simulations (15 simulations in total) probes dynamics of the ligand-bound as well as ligand-free states of the riboswitch on microsecond timescales (the total length of our simulations reaches ~29 μs). We have investigated the principles of structural response of the preQ<sub>1</sub> riboswitch aptamer to ligand presence and absence in the binding pocket and their consequences to accessibility of the part of the ribosome-binding site. We have also used this rather unique RNA system for testing the performance of different solute force fields. We employed three different RNA force fields; two variants (*ff99* and *ff99bsc0χ<sub>OL3</sub>*) of the AMBER (Cornell *et al.*)<sup>43,44</sup> force field, and CHARMM27.<sup>46</sup> Our simulations shed some light on the mechanism by which preQ<sub>1</sub> aptamer, upon ligand binding, sequesters bases of the ribosome binding site. In this

mechanism, loop nucleotides (nts. 11–15) play prominent role. We show that *ff99bsc0* $\chi_{OL3}$  (involving Barcelona *bsc0* reparameterizations for  $\alpha/\gamma$  dihedrals and Olomouc  $\chi_{OL3}$  reparameterization for glycosidic  $\chi$  dihedral) represents a viable force field for long-scale simulations on a wide variety of RNA structural motifs including pseudoknotted RNAs. Our results suggest tangible stabilization compared with earlier force fields. We also demonstrated that even  $\mu$ s-scale simulations are unable to properly capture rearrangements of the active site after the removal of the ligand.

## METHODS

### Studied systems

We investigated structural dynamics (and indirectly also stability) of the ligand-bound and ligand-free preQ<sub>1</sub> riboswitch. We focused on a translational riboswitch aptamer domain (PQA) from *Thermoanaerobacter tengcongensis* in which two bases of the RBS expression platform are observed in the crystal structure. The starting geometry of the aptamer domain was taken from the X-ray structure (PDB ID: 3GCA) with a resolution of 2.75 Å.<sup>13</sup> In this structure, the preQ<sub>1</sub> precursor (preQ<sub>0</sub>) was bound, which differs from the to preQ<sub>1</sub> by presence of a nitrile group in lieu of 7-methylamine. We chose this structure because the X-ray structure of PQA bound to preQ<sub>1</sub> was not yet available. Thus, the starting structure for simulations utilized an aptamer domain in which preQ<sub>1</sub> was generated by *in silico* modification of preQ<sub>0</sub> to preQ<sub>1</sub> (i.e., the nitrile -CN group of preQ<sub>0</sub> was replaced by -CH<sub>2</sub>-NH<sub>2</sub> (or -CH<sub>2</sub>-NH<sub>3</sub><sup>+</sup>) group of preQ<sub>1</sub>). Two protonation states of preQ<sub>1</sub> were considered bearing an uncharged amino (-CH<sub>2</sub>-NH<sub>2</sub>) or protonated ammonium (-CH<sub>2</sub>-NH<sub>3</sub><sup>+</sup>) group. In hindsight, we believe this approach is valid because the X-ray structure of PQA with preQ<sub>1</sub> ligand was determined recently (PDB ID: 3Q50) with a resolution of 2.75 Å.<sup>16</sup> The X-ray structures with preQ<sub>0</sub> and preQ<sub>1</sub> ligands share almost the same overall folds, architectures of the ligand binding pocket, and ligand binding patterns (the RMSD between the two structures is 0.6 Å). Our starting structure with modeled preQ<sub>1</sub> ligand agrees well with the 3Q50 X-ray structure (involving preQ<sub>1</sub>), which validates its suitability as a model to start the simulations.

Two different PQA starting structures lacking ligand were used in the simulations: (i) the structure based on 3GCA X-ray structure with removed preQ<sub>0</sub> ligand (henceforth “ligand-removed PQA”) and (ii) the structure based on recent ligand-free X-ray structure (PDB ID: 3Q51, resolution of 2.85 Å)<sup>16</sup> termed “ligand-free PQA”. In the 3Q51 X-ray structure, the ligand binding pocket is occupied by the base of A14 while the neighboring A13 is not resolved. The missing nucleotide A13 was thus modeled using the PyMOL program<sup>52</sup> while two different A13 conformations were taken into account (simulations labeled as Q51-free-1 and Q51-free-2 in Table 1). The lack of electron density of one of the tandem adenines and partially disordered electron density of nucleotides U12 and C15 in the 3Q51 X-ray structure suggested the possibility that the adenine resolved in the binding pocket might be A13 instead of A14, although efforts to model A13 were hampered by unrealistic crystal packing. Nonetheless, we tested the possibility that the adenine resolved in the ligand-binding pocket was actually A13 with the neighboring A14 nucleotide extruded to the solvent (data not shown). The simulations starting from this structure were rather inconsistent with the X-ray structure supporting the hypothesis that adenine resolved in the binding pocket was indeed A14.

The arrangement of nucleobases C7, G11 and C30 resembles that of triples C8, G12 and C26 of frame-shifting pseudoknot from Beet Western Yellows Virus (BWYV)<sup>53,54</sup> and C141, C144 and G161 of HDV ribozyme<sup>55,56</sup> (Figure S1, Supporting Information). Stability of the triples from BWYV pseudoknot and HDV ribozyme was studied in earlier MD simulations<sup>53,55</sup> for both canonical and N3-protonated states of the cytosine, which makes

*trans*-Watson-Crick/Hoogsteen (tWH) base pair with guanine (Figure S1, Supporting Information). The simulations of both systems consistently suggested that the base triple architecture is stable only if the respective cytosine is N3-protonated, in agreement with NMR studies.<sup>57</sup> Thus, we treated the C7 cytosine of PQA as N3-protonated.

### AMBER simulation protocol

We performed classical MD simulations using well established simulation protocols.<sup>19,45</sup> Missing hydrogen atoms were added by the LEaP module of the AMBER package.<sup>58</sup> The riboswitch aptamer was neutralized, depending on the ligand net-charge and/or the C7 protonation state (some simulations with canonical C7 were performed to explicitly test the influence of the C7 protonation state on structural stability of the C7, G11, C30 base triple, shown in Table S1 in the Supporting Information), with 30–32 Na<sup>+</sup> counter ions (radius 1.868 Å and well depth 0.00277 kcal/mol),<sup>59</sup> and immersed in a rectangular TIP3P<sup>60</sup> water box with a 10-Å thick layer of water molecules. For extensive justification of this ion/water model condition and comparison of different ion treatments in RNA simulations see refs.<sup>19,61</sup>. Parameters for all non-standard residues, preQ<sub>0</sub>, preQ<sub>1</sub>, protonated preQ<sub>1</sub> and protonated cytosine, were developed according to the Cornell *et al.* procedure.<sup>43,62</sup> (See Supporting Information for more details and the parameter files). The RNA-solvent system was minimized prior to the simulation as follows. Minimization of the solute hydrogen atoms was followed by minimization of counter ions and water molecules. Subsequently, the aptamer was frozen and solvent molecules with counter ions were allowed to move during a 10-ps long MD run in order to relax the density of the system. After that, the nucleobases were allowed to relax in several minimization runs with decreasing force constants applied to the backbone phosphate atoms. After full relaxation, the system was slowly heated to 298.15 K over 100 ps using 2-fs time steps and *NpT* conditions using a weak-coupling scheme with a coupling time of 1 ps.<sup>63</sup> The production phases of MD simulations were carried out under periodic boundary conditions (PBC) in the *NpT* ensemble (298.15 K, 1 atm) with 2-fs time steps. The particle-mesh Ewald (PME) method<sup>64,65</sup> was used to calculate electrostatic interactions with a cubic-spline interpolation and ~1 Å grid spacing; a 10.0-Å cut-off was applied for Lennard-Jones interactions with automatic rebuilding of the buffered pair-list when atoms moved more than 0.5 Å. The SHAKE algorithm was applied to fix all bonds containing hydrogen atoms. The PMEMD module of AMBER 11.0 (ref. <sup>66</sup>) was used for simulations. Simulations were run with the Cornell *et al.* *f99* force field<sup>43,44</sup> and its modified *f99bsc0*χ<sub>OL3</sub> variant, including Barcelona *bsc0*<sup>67</sup> and Olomouc χ<sub>OL3</sub><sup>68</sup> corrections to the α/γ and χ torsions, respectively. The χ<sub>OL3</sub> reparameterization was originally marked as χ<sub>OL</sub>.<sup>19,45,68</sup> When this correction was incorporated in the AMBER code, χ<sub>OL3</sub> notation was used in AmberTools 1.5 manual. As such, we decided to unify the notation to avoid any further confusion and henceforth the χ<sub>OL3</sub> name will be used. The cumulative simulation time of all AMBER simulations equals to 29 microseconds (16.2 and 12.8 microseconds for ligand-bound and ligand-free states, respectively). See Table 1 and Table S1 in the Supporting Information for overview of MD simulations.

### CHARMM simulation protocol

MD simulations of the ligand-removed PQA (with canonical as well as protonated C7) were also performed with the CHARMM27 force field<sup>46</sup> using the NAMD<sup>69</sup> package (ver. 2.6) in the context of the following protocol. To avoid any differences in starting geometries, the geometry of neutralized and solvated riboswitch were identical to that prepared for AMBER simulations, i.e., even the same starting positions of ions and water molecules. The CHARMM<sup>70</sup> software package (ver. 34b2) was used to prepare CHARMM27 topologies and coordinates. The water molecules and counter ions were minimized in 2,500 steps and shaken by a short *NpT* dynamics (100 ps) at 300 K and 1 atm. The RNA solute was minimized prior to simulation in 3000 steps, then slowly heated to 300 K over 100 ps using

1-fs time steps and  $NpT$  conditions using Langevin dynamics.<sup>71,72</sup> The simulation was conducted under the periodic boundary conditions in the  $NpT$  ensemble (300 K, 1 atm) with 1-fs time steps, because the 2-fs integration step produced considerably less stable trajectories for A-RNA stems.<sup>45</sup> The PME method was applied to calculate electrostatic interactions (PME tolerance of  $10^{-6}$ ) and the 12.0 Å cut-off with an 8.0 Å switching distance was applied for Lennard-Jones interactions.

### Force field effect and data analysis

Our simulation data indicate that the pseudoknotted architecture of PQA is very sensitive to the force field employed. Stable trajectories for both the ligand-free as well as ligand-bound aptamer were obtained only with the  $f99\text{bsc0}\chi_{\text{OL3}}$  force field (these trajectories are summarized in Table 1). In contrast, neither  $f99$  nor CHARMM27 provide sufficiently stable trajectories on long time scale (Table S1, Supporting Information). The aptamer pseudoknotted fold was distorted substantially compared to the X-ray structure. With  $f99$  it occurred on sub- $\mu\text{s}$  time scale and with CHARMM27 at  $\sim 100$  ns. In  $f99$  simulations, the aptamer rearranged into the “ladder-like” architecture, which is characterized by a shift of glycosidic torsions of almost all nucleobases from *anti* to the high-*anti* region, reduction of P1 stem twist from  $\sim 32^\circ$  to  $\sim 15^\circ$ , complete distortion of the ligand binding pocket, disruption of a majority of native non-WC base pairs and formation of multiple non-native base pairs or stacking interactions (Figure 2). The “ladder-like” rearrangement is associated with the parameters of the glycosidic  $\chi$  torsion in the  $f99$  force field as explained in our previous studies.<sup>19,23,45,68</sup> The appearance of the sign of “ladder-like” misfolding indicates that the bias toward “ladder-like” geometries is rather strong even in such compactly folded structure as the preQ<sub>1</sub> riboswitch and implies that  $f99$  cannot be recommended for RNA simulations (note that the parmb<sub>sc0</sub> correction does not prevent formation of ladders). In CHARMM27 simulations, PQA adopts a geometry, with distorted P2 stem, quartet and triplets of the junction and a portion of the P1 stem (C1=G20 and U2-A19 base pairs). In addition, all tertiary contacts between the L3 loop and the P1 stem minor groove are lost (Figures 1 and 2).

Analyses of trajectories were performed by Ptraj (from AmberTools package) and X3DNA<sup>73</sup> programs. The B-factors were calculated as squared atomic positional fluctuations multiplied by factor  $8\pi^2/3$ . The VMD<sup>74</sup> and PyMOL programs were used for visualizations. Stacking energies for a given pair of bases were calculated using the NAMD Energy Plugin in VMD. The cut-off for non-bonded components was set to 99 Å while the calculations were performed with the dielectric constant of 1 (i.e. *in vacuo*).

### Non-standard residues preQ<sub>1</sub>, preQ<sub>0</sub> and protonated cytosine

Parameters for all non-standard residues were developed according to the Cornell *et al.* procedure.<sup>43,62</sup> The parameterization strategy is thoroughly described in the Supporting Information. In addition, prep files of all four non-standard residues are provided in the Supporting Information.

## RESULTS

### Starting structures

**Architecture of the preQ<sub>0</sub>-aptamer complex**—PQA harboring preQ<sub>0</sub> in its binding pocket folds into an H-type pseudoknot.<sup>16</sup> The pseudoknot consists of two stems (P1 and P2) that are separated by three loops (L1, L2, and L3; Figure 1). The P1 stem is a canonical A-RNA duplex (comprising C1-G5 and C16-G20 strand segments) (Figure 1). The P2 stem consists of only two base pairs, the canonical C9=G33 pair and the *trans*-Hoogsteen/Sugar-

edge ( $\theta$ HS) A32/A10 base pair. A32 and G33 belong to the RBS so that when P2 stem is formed, part of the RBS is sequestered into the PQA core structure (Figure 1).

The adenine A23 of the L3 loop orients its sugar edge towards the canonical U2-A19 P1 stem base pair (Figure 1), thus forming an A-minor I interaction<sup>75</sup> with an A-U base pair receptor instead of the more common C=G one (hereafter called the A-I/A-U interaction). L3 is further fixed by a tertiary base-phosphate (BPh) interaction<sup>76</sup> formed between G3 of the P1 stem and A24 of L3. The PQA structure with bound preQ<sub>0</sub> reveals another BPh interaction between G8 of L1 and G11 of L2 (Figure 1). The stability of both BPh interactions in  $\theta$ 99bsc0 $\chi$ <sub>OL3</sub> simulations is analyzed in Table S2 in the Supporting Information.

The preQ<sub>0</sub> ligand is almost completely buried within the aptamer structure. Only the –CN group and O1 atom (formally the ligand's Hoogsteen edge) are fully or in part exposed to the solvent, respectively (see Figure 1 for atom numbering). The ligand binding pocket comprises two base triples and one quartet (Figure 1B). The G5, C16, A27 triple and the C7, G11, A14, C30 quartet represent the “floor” and “ceiling” of the ligand binding pocket, respectively. Namely, the G5=C16 base pair and G11 base interact, via stacking, with preQ<sub>0</sub>. The second triple of bases (the ligand binding plane) is responsible for H-bonding of preQ<sub>0</sub> to the pocket. It includes U6, C15, and A29, of which C15 orients its Watson-Crick face to the “Watson-Crick” face of preQ<sub>0</sub>. Additionally, preQ<sub>0</sub> is stabilized by one H-bond with U6 and two H-bonds with A29 (Figure 1B).

**Architecture of the ligand-free aptamer**—The most significant structural differences between ligand-free (3Q51) and ligand-bound (3GCA and 3Q50) X-ray structures occur in A14 and C15 nucleotides of L2. A14 is stacked between bases A13 and C15 in the ligand-bound PQA structure, but it occupies the ligand binding pocket in the ligand-free PQA structure (Figure 3). Essentially, A14 replaces the preQ<sub>0</sub>/preQ<sub>1</sub> ligands since it interacts with the same set of bases with exception of C15. However, the A14 base is rotated around the normal axis (perpendicular to the A14/ligand ring) by  $\sim 90^\circ$  compared to the ligand so that A29 and U6 are H-bonded with A14 WC-edge, whereas they bind the “sugar edge” of the preQ<sub>0</sub>/preQ<sub>1</sub> ligand in ligand-bound state (Figure 3C). C15 is not part of the ligand binding pocket in the ligand-free PQA structure, as it is unstacked and protruded into solvent (Figure 3). By contrast, the P2 stem adopts highly similar conformations in both aptamer states (Figure 3B). A main difference is that the  $\theta$ HS A32/A10 base pair is slightly less compact in the ligand-bound aptamer due to larger propeller twist ( $-28^\circ$ ) (Figure 3B).

### Crystal contacts seen in the *Thermoanaerobacter tengcongensis* PQA X-ray structures

Because crystal packing can affect the X-ray structure, it is important to analyze the crystal contacts. In both X-ray structures used in this study as starting geometries (3GCA and 3Q51), the most apparent crystal contacts are two coaxial stacks. The first stack occurs at the terminal C1=G20 base pair of the P1 stem, which packs against the same base pair from a dyad-related molecule in the lattice. A second co-axial stack exists between the C9=G33 base pair of the P2 stem and its dyad-related symmetry mate (Figures 4A and 4D). These coaxial stacking interactions cap the PQA structure at both 5'- and 3'-ends. Furthermore, the X-ray structures involve four additional crystal contacts. These include a stacking interaction between U22 and A24 (either between U22 and A24 of the symmetry copy or *vice versa*) (Figures 4B and 4E) and an interaction (stack in 3Q51 and hydrogen bond interaction in 3GCA structure) between U12 and A32 (either between U12 of a given molecule and A32 of its symmetry copy or *vice versa*) (Figures 4C and 4F). The recently published X-ray structure of PQA with bound preQ<sub>1</sub> reveals the same crystal contacts.

## Stability of the simulations and selection of appropriate solute force field

Three different force fields were initially used for the simulations of preQ<sub>1</sub> riboswitch, i.e. two Cornell *et al.* (AMBER) type force fields *ff99* and *ff99bsc0χ<sub>OL3</sub>*, and CHARMM27 force field. We found that only the most recent version of AMBER force field *ff99bsc0χ<sub>OL3</sub>* is able to provide stable simulations of the preQ<sub>1</sub> riboswitch on microsecond time scale. Formation of artificial “ladder-like” structures on a time scale of hundreds of nanoseconds as well as partial unfolding on ~100 ns time scale were observed with *ff99* or CHARMM27 force fields, respectively (Supporting Information). It is worth noting that such behavior is independent of the starting structure, the protonation state of C7 cytosine or the presence of ligand. We suggest that the structural perturbations are attributed unambiguously to the respective solute force fields. The results are consistent with recent observations for other RNA systems.<sup>19,23,45,47,48,50,68,77</sup> Therefore, all further analyses present in this study are based on the simulations performed using the *ff99bsc0χ<sub>OL3</sub>* variant of the AMBER force field. Note that the CHARMM force field has been recently upgraded to the CHARMM36 version, which provides more stable RNA trajectories.<sup>78</sup> The present riboswitch is stable after ~300 ns with CHARMM36. As the force field testing is not the main scope of this work, more thorough comparison of force fields will be published separately.

## Simulation behavior of the preQ<sub>1</sub> riboswitch aptamer

**Global conformational behavior of PQA**—We did not observe any significant global structural rearrangement of PQA in any simulation carried out using the *ff99bsc0χ<sub>OL3</sub>* force field (Table 1). Analyses of global structural characteristics showed that the aptamer domain maintains its initial architecture along the trajectories of microsecond-long simulations. The RMSD values are converged below 3.2 Å and the radii of gyration R<sub>g</sub>s lie in a narrow interval from 13.1–13.5 Å, which is close to crystallographic values of 12.9 Å (ligand-bound PQA) and 13.2 Å (ligand-free PQA) (see Table 1); notably, the latter values were also in agreement with measurements made for the crystallization construct in solution by small angle X-ray scattering.<sup>16</sup>

The P1 stem remains folded as a stable A-form RNA duplex in all simulations (regardless of the presence or absence of the ligand). The analysis of B-factors indicates that values of the P1 stem agree well with values of the L1 loop, suggesting these are relatively rigid parts of PQA (Figure 5). Only some fluctuations in roll and inclination, which are interrelated parameters,<sup>50,77</sup> are observed in some simulations (Table 2). The ligand binding pocket is also very rigid in all simulations with exception of the simulation in which the ligand was removed from the ligand-bound structure. Removal of preQ<sub>0</sub> markedly increased the flexibility of cytosine C15 (Figure 5B).

## Binding patterns of PreQ<sub>0</sub> and PreQ<sub>1</sub>; comparison of simulation and

**experimental results**—PreQ<sub>0</sub> as well as both variants of preQ<sub>1</sub> (i.e. neutral and protonated forms) remain tightly bound to the PQA pocket during the whole simulations. The same behavior was observed also for adenine A14 occupying the ligand binding pocket in the ligand-free aptamer. The ligands as well as the A14 base permanently retain their X-ray stacking patterns within the binding pocket (Figure S2, Supporting Information). Similarly, the hydrogen bonding network formed between the binding pocket and preQ<sub>0</sub> ligand or adenine remains unmodified (i.e. in X-ray-like arrangement, Table 3 and Table S3, Supporting Information). In contrast, preQ<sub>1</sub> in both neutral and protonated forms establishes additional H-bonds between amino/ammonium group as a proton donor and G5(O6) carbonyl as a proton acceptor. In addition, the ammonium group of protonated preQ<sub>1</sub> ligand forms second additional H-bond between ammonium group as a proton donor and G11(O2<sup>-</sup>) hydroxyl group as a proton acceptor (see Table 3 and Table S3, Supporting Information). It is worth noting that the amino/ammonium group of preQ<sub>1</sub> is also involved in an intra-



molecular H-bond contact to its own O1 carbonyl of ligand (see Figure 1B for preQ<sub>1</sub> atom numbering). Taken together, our simulation data provide evidence that the preQ<sub>1</sub> ligand in the aptamer's pocket is stabilized by one or two extra H-bonds in neutral or protonated form, respectively, in comparison to the preQ<sub>0</sub> ligand. These extra H-bonds might explain the 17-fold higher affinity of preQ<sub>1</sub> compared to preQ<sub>0</sub> as observed by surface plasmon resonance (SPR).<sup>16</sup> Interestingly, the orientation of the methylamino group of the preQ<sub>1</sub> ligand differs between MD simulations and the recent X-ray structure of preQ<sub>1</sub> bound. Although functional analysis predicted an H-bond between O6 of G5 and the methylamine of preQ<sub>1</sub>, such interaction is not seen in the X-ray structure. We suggest that it has been obfuscated in the crystal structure by contact between the methylamine and a sulfate ion in the mother liquor.<sup>16</sup> By contrast, our MD analyses indicate that the H-bond contact of the G5(O6) to preQ<sub>1</sub> methylamine is geometrically reasonable and that this favorable interaction accounts for the observed experimental SPR data.<sup>16</sup>

**Dynamics of L3, P1 and L1 segments**—The L3 loop reveals a non-uniform flexibility pattern in our simulations. The 3'-tail of L3 proximal to the P2 stem (A26–A31) is very rigid since it is an integral part of the rigid ligand binding pocket (see above). In contrast, the 5'-tail of the L3 loop (U21–C25), which runs parallel to the P1 stem minor groove, is one of the most flexible PQA segments in our simulations (see B-factors in Figure 5). The fluctuations occurring at the 5'-tail of L3 are associated with the dynamics of the A-minor type I interaction between A23 and A19–U2 base pair (A-I/A-U) and fluctuations of cHS U21/G20 pair (see Supporting Information for more details about structural dynamics of A-I/A-U interaction). Overall, we found that structural dynamics of P1 stem, L3 loop and the binding pocket involving the L1 loop is independent of ligand presence or absence. These structural parts of PQA fluctuate near the initial X-ray geometry and remain stable in the simulations.

#### **Dynamics of L2 and P2 stem and their possible role in the riboswitch function**

—The behavior of the P2 stem and L2 loop depends on the presence of the ligand. In general, the differences of L2 loop conformations between ligand-free and ligand-bound systems as observed in X-ray structures affect the stability (indirectly inferred from the observed structural dynamics) of the P2 stem in our simulations. While the L2 loop conformation of the ligand-bound system remains intact and fluctuates near its X-ray geometry, some conformational plasticity of the L2 topology was found in both ligand-free PQA simulations based on the 3Q51 structure (see Methods). We observed a spontaneous formation of GpU platform by G11 and U12 nucleotides (Figure 6). The GpU platform is a recurrent RNA structural submotif found e.g. in sarcin-ricin loop domain.<sup>79–81</sup>

The GpU platform is formed within tens of nanoseconds by conformation adjustments of the sugar-phosphate backbone around the U12 nucleotide and sequestration of U12 into the intramolecular interactions within the L2 loop (Figure 6). In contrast, the U12 as modeled in the X-ray structure is bulged to the solvent and forms a crystal contact with A32 of the neighboring molecule in the crystal lattice (see above). As the formation of GpU platform was observed to be a spontaneous and relatively fast process in both ligand-free simulations (Q51-free-1 and Q51-free-2), we hypothesize that the ligand-free structure of PQA in solvent most likely involves such GpU platform conformations and its absence in the X-ray structure might be an artifact of crystal contacts.

The structural dynamics and stability of the P2 stem are affected by two different structural components: (i) the base fraying of the terminal G33=C9 base pair, and (ii) the disruption of the A10/A32 base pair. Dynamics of the former were found in both ligand-free as well as ligand-bound simulations and comprise disruption of the G=C base pair accompanied by unstacking of G33, and possibly rebuilding of the G=C interactions after a *syn/anti* flip of G33 (see substates P2(ii), P2(vii), P2(viii), and P2(ix) in Figure S5 in the Supporting

Information). Such base pair fraying is not unusual for terminal base pairs.<sup>82</sup> On the other hand, the stability of the terminal base pairs is increased significantly in the presence of a 3'-overhang.<sup>82</sup> The A32 and G33 nucleotides of PQA are part of the RBS and thus the full-length native PQA would inherently include the 3'-overhang of P2 stem represented by the genuine extension of the RBS. Thus, the G33=C9 base pair fraying observed in our MD simulations may be due to the incomplete sequence of PQA used for the crystallization and in our MD simulations lacking the rest of RBS in 3'-overhang of the P2 stem.

In contrast, the disruption of A10/A32 base pair (see substates P2(iii), P2(iv), P2(v), P2(vi), and P2(x) in Figure S5 in the Supporting Information) was observed almost exclusively in ligand-free simulations. Namely, we evidenced degradation of P2 stem structure in one of two independent ligand-free simulations (Q51-free-2), while we observed only transient and fully reversible fluctuations of A10/A32 base pair in other simulations and we never observed irreversible P2 stem degradation in any of the three ligand-bound simulations (GCA-Q0, GCA-Q1, and GCA-Q1+) (Figure 7). The fluctuations of A10/A32 base pair as well as irreversible degradation of P2 stem in ligand-free simulation occurred on microsecond time-scale. It means that even our several-microsecond-long simulations, representing state-of-the-art time-scales, are likely not fully converged. Despite the evident limits of sampling, we suggest that the simulations capture relevant difference in the structural dynamics of ligand-free and ligand bound systems. Detailed analyses of the simulations suggest that the observed difference in stability of the P2 stem between ligand-free and ligand-bound states could be rationalized most likely by differences in intramolecular interactions that stabilize the P2 stem. On the basis of the simulation data we suggest that the stability of the A10/A32 base pair is indeed modulated by the extent of stacking interaction between A32 and the bases of L2 loop. In the ligand-bound system, the A32 is firmly stacked to A13 of the L2 loop, which is in turn base paired with A31. The mean van der Waals interaction energy between A32 and A13 is  $\sim -6$  kcal/mol (calculated using the Lennard-Jones term of the force field). We observed transient disruptions of A31/A13 base pair followed by partial disruptions of A10/A32 base pair of the P2 stem. In such cases, A32 remains stacked on A13 and thus the stabilities of the A10/A32 and A31/A13 base pairs are interrelated (Figure 7). This transient disruption is, however, short-living and fully reversible. In most of the simulation time, both A10/A32 and A13/A31 base pairs are stably paired. On the other hand, in the ligand-free PQA, the A13 is no longer stacked to A32 as its adjacent nucleotide A14 is bound in the ligand binding pocket. Instead, the A32 is stacked on U12, which is involved in the aforementioned GpU platform. However, the stacking energy represented by mean van der Waals interaction energy between A32 and U12 is significantly lower and equals to  $\sim -3$  kcal/mol. In contrast to ligand-bound simulations, the A10/A32 base pair of the P2 stem is significantly more flexible, which finally leads to disruption of the whole P2 stem after hundreds of nanoseconds.

Taken together, we observed that the L2 loop is able to establish two different local arrangements depending on presence or absence of the ligand. These two arrangements of the L2 loop provide different stacking pedestals for the P2 stem, thereby influencing its stability, which in turn should lead to an interrelation between ligand binding and the accessibility of the RBS. The L2 loop thus could relay information about presence or absence of ligand to the P2 stem, thus controlling the sequestration of the RBS to regulate translation; (the proposed molecular mechanism of preQ<sub>1</sub> action based on the simulation data is shown in Figure 8).

The role of the L2 loop in communication between the ligand binding pocket and the P2 stem is also supported by the ligand-removed PQA simulation (labeled as GCA-Q0del in Table 1). In this simulation, the starting structure is based on the 3GCA crystal structure, but the preQ<sub>0</sub> ligand is removed from the binding pocket. We observed significantly increased

flexibility of C15, which is originally bound to the ligand by its WC-edge. This flexibility is propagated by the stacking interactions through A14 and A13 adenines of the L2 loop up to A32 and finally leads to disruption of the A32/A10 base pair (see the P2(vi) substate representing this P2 stem conformation). It should be noted that in this simulation neither the A14 nor the A13 adenines bind to the ligand binding pocket in a manner similar to that seen in the ligand-free X-ray PQA structure. Evidently, the ligand-removed simulation, despite the microsecond time scale, captures only initial part of the ligand binding pocket/L2 loop reorganization process. The simulation is definitely not converged. In other words, complete conformational rearrangement of the ligand-bound structure into a ligand-free structure after removing the ligand most likely happens on a time scale that is beyond the applicability of contemporary classical MD simulations. This suggestion is also consistent with absence of any significant fluctuations of A14 in the ligand-free simulations, where A14 remains firmly locked inside the ligand-binding pocket. Any mechanism describing the complete pathway from ligand-free state to preQ<sub>1</sub> bound state would require to sample at least minor populated “open” conformational state competent to interact with the ligand.<sup>16</sup>

## DISCUSSION AND CONCLUSIONS

We employed MD simulations to investigate the flexibility and structural dynamics of a translational preQ<sub>1</sub> riboswitch aptamer domain from *Thermoanaerobacter tengcongensis*. The molecular dynamics, when used insightfully with enough attention paid to force field limitations, can provide unique information that complements experimental data.<sup>17,83</sup> Here, we used two X-ray structures representing different functional states of the preQ<sub>1</sub> riboswitch as starting structures for our simulations -- namely the ligand-bound (PDB ID 3GCA; ref. <sup>13</sup>) and the ligand-free (PDB ID 3Q51; ref. <sup>16</sup>) states. The NMR experiments carried out on related transcriptionally acting preQ<sub>1</sub> riboswitch aptamers identified that local conformational motions occurred on pico-to nano-seconds time scales, while larger conformational changes were detected on micro-to (mili-)seconds time scales.<sup>15,84–86</sup> The microseconds time scale used in this study approaches the lower limit of large-scale motions, however, it is still too short to fully sample the global conformational changes. Even having in hand much longer simulations, which are really difficult to obtain with current computer power, a direct comparison with NMR data would not be straightforward because of two different aptamer domains used in NMR and MD studies.

We found that in case of the ligand-bound PQA the aptamer stably fluctuates near the X-ray geometry. In other words, the equilibrium structure (on the present time-scale) as observed in MD simulations, does not differ significantly from the X-ray geometry. The only differences between structures as observed by X-ray and MD simulations in the ligand-bound state (but also in the ligand-free PQA structure) are the stability of the A-minor type I interaction and binding pattern of preQ<sub>1</sub> ligand. The A-minor interaction motif is observed in the X-ray structure between adenine A23 and A19-U2 base pair of P1 stem, while it is rather flexible and loses its signature geometry during MD simulations. Crystal packing may contribute to the difference in stability since the X-ray structure includes crystal contacts between U22 (the adjacent nucleotide to A23 that forms the A-minor interaction) and A24, each donated from neighboring molecules. Nonetheless, we also cannot rule out that the difference in stability might be attributed to slightly underestimated H-bonding interactions in empirical force fields.<sup>87</sup> In case of the binding pattern of preQ<sub>1</sub> ligand, MD simulations reveal interaction between O6 of G5 and ligand's methylamino group, while such interaction is not observed in X-ray due to alternative interaction of ligand's methylamino group to sulfate ion of mother liquor.

In contrast, we suggest that the structure of ligand-free state of PQA might be more affected by crystal contacts. Namely, we found that the U12, which forms a crystal contact with A32

of neighboring molecule in the crystal lattice spontaneously forms a GpU platform with G11 in simulations and thus helps to form a well-defined conformation of the L2 loop even in the ligand-free state. Nevertheless, the bioinformatics analysis does not reveal any sequence conservation of this GpU platform. Also the suggested mechanism of the riboswitch action does not assume any specific biological role of this GpU, indicating no evolution pressure to conserve the GpU platform. Its formation thus may be rather a coincidence of using a specific variant of the riboswitch. Still, we assume that the loss of the above-noted crystal contact would lead to L2 remodeling even in case of different sequences.

Another crystal contact most likely affecting the structure of ligand-free PQA is a coaxial stacking of the terminal G33=C9 base pair of the P2 stem with its dyad-related symmetry mate, which increases the stability of P2 stem in the crystal. When this contact is removed, the P2 stem of ligand-free PQA reveals significant fluctuations. The last observed crystal contact is located between the terminal C1=G20 base pair of P1 stem and its dyad-related symmetry mate. This interaction does not seem to have any significant effect on the PQA structure in either the ligand-bound or ligand-free states.

On the basis of the difference in observed structural stability of the P2 stem between ligand-free and ligand-bound simulations we hypothesize that a part of the RBS undergoes sequestration in response to ligand binding, as summarized in Figure 8. We observed that the L2 loop establishes structurally well-defined but distinct conformations in ligand-bound and ligand-free states. In the ligand-bound state, the L2 loop is structured, so that A32 of the P2 stem is stacked with A13, which belongs to the L2 loop. In contrast, in the ligand-free state, where A14 is bound within the ligand binding pocket, A13 can no longer stack with A32, which instead stacks with the GpU platform formed by G11 and U12 nucleotides of the L2 loop. Interaction energy calculations show that the stacking interaction between A32 and the GpU platform of the L2 loop in ligand-free state ( $\sim -3$  kcal/mol, van der Waals part of interaction energy) is significantly weaker than the stacking interaction between A32 and A13 in the ligand-bound state ( $\sim -6$  kcal/mol). This difference in the stacking interaction between P2 and L2 affects the stability of the A32/A10 base pair and consequently also the whole P2 stem (Figure 8). In other words, the different conformation of the L2 loop, which is induced by the presence or absence of ligand in the binding pocket, affects the stability of the P2 stem via a stacking interaction. Consequently, since one strand of the P2 stem is part of the RBS, the stability of the P2 stem affects the structure available for translation initiation. The L2 loop thus can mediate the information flow about the ligand's presence between the binding pocket and the P2 stem, which has the capability to sequester part of the RBS.

Finally, we should note that only small structural changes were observed in the ligand-removed PQA simulation, where we used the 3GCA ligand-bound PQA X-ray structure but manually removed the ligand from the starting structure of our simulation. Although this type of simulation, i.e. ligand-removed simulations, is common in the literature,<sup>27-30,88</sup> we suggest some caution is needed in the interpretation. Note that we have used microsecond simulations, longer than employed for similar systems in the past.<sup>27-30,88</sup> Still, the ligand-removed simulation based on a ligand-bound structure apparently does not provide converged information about structural dynamics of the genuine ligand-free state and in fact may be very far from being converged. Also the fact that we do not see any sign of transition from the A14 “closed” to an “open” state receptive to preQ<sub>1</sub> binding when starting from the genuine ligand-free structure confirms the MD trajectories are far from being converged. This suggests that conformational plasticity of the aptamer in response to ligand binding or its absence from the pocket can only be realized on much longer time scale that affordable in current state-of-the-art microsecond atomistic simulations. In addition, when the ligand-removed simulation is not accompanied at least by ligand-bound simulation, the possible

force field artifacts in the simulations might be easily misinterpreted as a native response of the system to removal of the ligand. In the present study we have for the first time used the latest  $\chi_{OL3}$  AMBER RNA force field version for riboswitch simulations. This force field brings substantial improvement in stabilizing RNA simulations compared to force fields that have been used for riboswitches in the past.

Overall, our results suggest that MD can provide useful complementary information that explains experimental results not represented by the crystal structures, such as the methylamine group interacting with O6 of G5, and the formation of a GpU platform that represents part of the observed RNA structural repertoire. By scrutinizing all the data we were able to suggest a plausible qualitative mechanism of the riboswitch action, as seen in Figure 8. It is, however, important to note that the stability of the riboswitch will likely be significantly different when confronted by the translation initiation machinery, which should shift the equilibrium to unfolded states. Additional experimental information is needed to understand this phenomenon, which can then be used to guide additional MD calculations aimed at providing an atomistic-level understanding of the gene regulation process.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by the project "CEITEC -Central European Institute of Technology" CZ.1.05/1.1.00/02.0068 (J.S.), Operational Program Research and Development for Innovations - European Regional Development Fund (project CZ.1.05/2.1.00/03.0058), and the Operational Program Education for Competitiveness - European Social Fund (CZ.1.07/2.3.00/20.0017) of the Ministry of Education, Youth and Sports of the Czech Republic (M.O., P.B., P.S.), and by grants 203/09/H046 (M.O., P.S., J.S.), 203/09/1476 (J.S.), P208/12/1878 (J.S., M.O.), P305/12/G03 (J.S.), P208/12/G016 (M.O., P.S.) and P301/11/P558 (P.B.) by the Grant Agency of the Czech Republic. J.E.W. is supported by a U.S. Public Health Service grant from NIH/NIGMS GM063162-09A1.

## REFERENCES

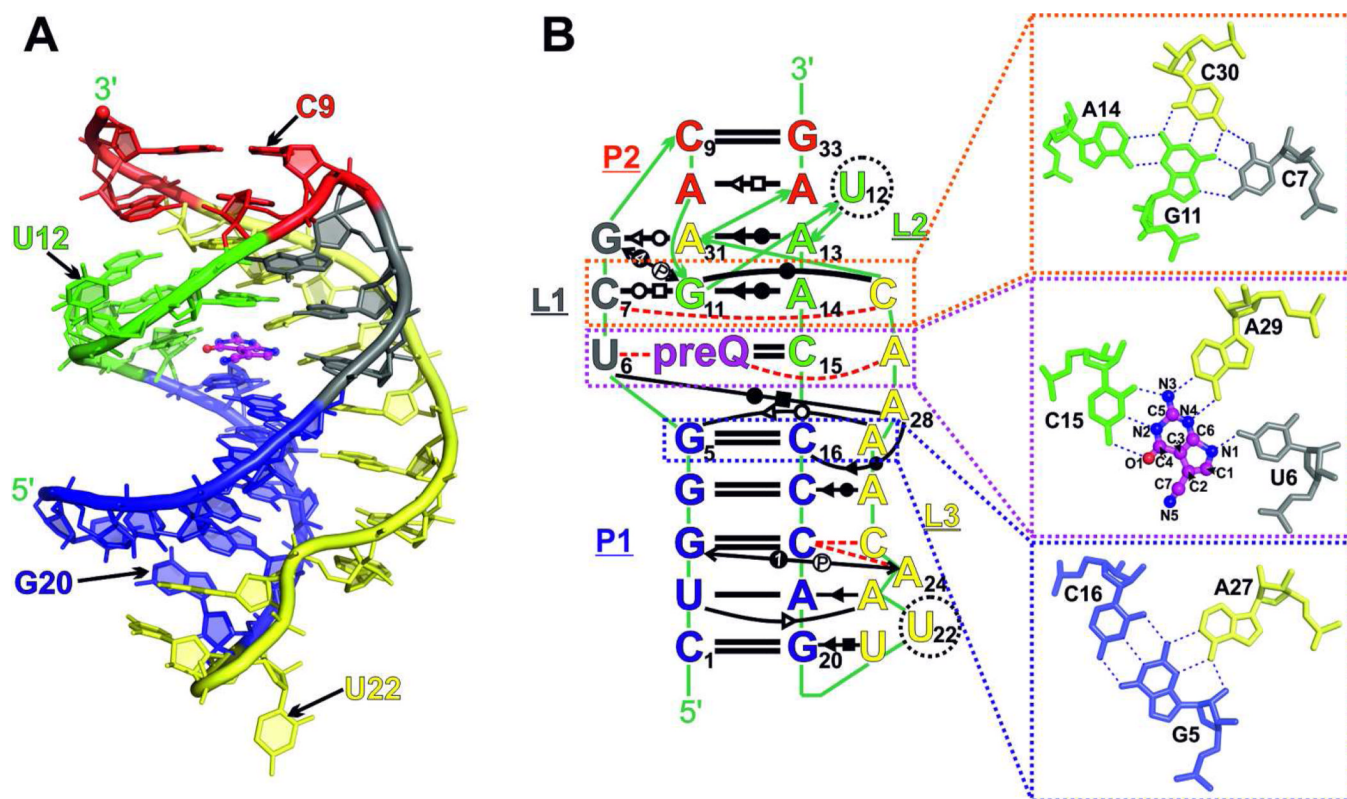
1. Barrick JE, Breaker RR. *Genome Biol.* 2007; 8:1.
2. Serganov A. *Rna Biol.* 2010; 7:98. [PubMed: 20061809]
3. Liberman JA, Wedekind JE. *WIREs RNA.* 2012; 3:369. [PubMed: 21957061]
4. Zhang J, Lau MW, Ferre-D'Amare AR. *Biochemistry-U.S.* 2010; 49:9123.
5. Blouin S, Mulhbacher J, Penedo JC, Lafontaine DA. *Chembiochem.* 2009; 10:400. [PubMed: 19101979]
6. Liberman JA, Wedekind JE. *Curr Opin Struct Biol.* 2011; 21:327. [PubMed: 21530235]
7. Montange RK, Batey RT. *Annu Rev Biophys.* 2008; 37:117. [PubMed: 18573075]
8. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. *Nucleic Acids Res.* 2005; 33:D121. [PubMed: 15608160]
9. Weinberg Z, Barrick JE, Yao Z, Roth A, Kim JN, Gore J, Wang JX, Lee ER, Block KF, Sudarsan N, Neph S, Tompa M, Ruzzo WL, Breaker RR. *Nucleic Acids Res.* 2007; 35:4809. [PubMed: 17621584]
10. Meyer MM, Roth A, Chervin SM, Garcia GA, Breaker RR. *RNA.* 2008; 14:685. [PubMed: 18305186]
11. Roth A, Winkler WC, Regulski EE, Lee BWK, Lim J, Jona I, Barrick JE, Ritwik A, Kim JN, Welz R, Iwata-Reuyl D, Breaker RR. *Nat Struct Mol Biol.* 2007; 14:308. [PubMed: 17384645]
12. Rieder U, Lang K, Kreutz C, Polacek N, Micura R. *Chembiochem.* 2009; 10:1141. [PubMed: 19382115]
13. Spitale RC, Torelli AT, Krucinska J, Bandarian V, Wedekind JE. *J Biol Chem.* 2009; 284:11012. [PubMed: 19261617]

14. Klein DJ, Edwards TE, Ferre-D'Amare AR. *Nat Struct Mol Biol.* 2009; 16:343. [PubMed: 19234468]
15. Kang M, Peterson R, Feigon J. *Mol Cell.* 2009; 33:784. [PubMed: 19285444]
16. Jenkins JL, Krucinska J, McCarty RM, Bandarian V, Wedekind JE. *J Biol Chem.* 2011; 286:24626. [PubMed: 21592962]
17. Ditzler MA, Otyepka M, Sponer J, Walter NG. *Accounts Chem Res.* 2010; 43:40.
18. Razga F, Zacharias M, Reblova K, Koca J, Sponer J. *Structure.* 2006; 14:825. [PubMed: 16698544]
19. Sklenovsky P, Florova P, Banas P, Reblova K, Lankas F, Otyepka M, Sponer J. *J Chem Theory Comput.* 2011; 7:2963.
20. Reblova K, Spackova N, Stefl R, Csaszar K, Koca J, Leontis NB, Sponer J. *Biophys J.* 2003; 84:3564. [PubMed: 12770867]
21. Reblova K, Strelcova Z, Kulhanek P, Besscova I, Mathews DH, Van Nostrand K, Yildirim I, Turner DH, Sponer J. *J Chem Theory Comput.* 2010; 6:910. [PubMed: 21132104]
22. Spackova N, Sponer J. *Nucleic Acids Research.* 2006; 34:697. [PubMed: 16456030]
23. Mlynsky V, Banas P, Hollas D, Reblova K, Walter NG, Sponer J, Otyepka M. *J Phys Chem B.* 2010; 114:6642. [PubMed: 20420375]
24. Banas P, Walter NG, Sponer J, Otyepka M. *J Phys Chem B.* 2010; 114:8701. [PubMed: 20536206]
25. Reblova K, Lankas F, Razga F, Krasovska MV, Koca J, Sponer J. *Biopolymers.* 2006; 82:504. [PubMed: 16538608]
26. Besseova I, Reblova K, Leontis NB, Sponer J. *Nucleic Acids Res.* 2010; 38:6247. [PubMed: 20507916]
27. Villa A, Wohnert J, Stock G. *Nucleic Acids Res.* 2009; 37:4774. [PubMed: 19515936]
28. Sharma M, Bulusu G, Mitra A. *RNA.* 2009; 15:1673. [PubMed: 19625387]
29. Kelley JM, Hamelberg D. *Nucleic Acids Res.* 2010; 38:1392. [PubMed: 19969538]
30. Petrone PM, Dewhurst J, Tommasi R, Whitehead L, Pomerantz AK. *J Mol Graphics Modell.* 2011; 30:179.
31. Nguyen PH, Derreumaux P, Stock G. *J Phys Chem B.* 2009; 113:9340. [PubMed: 19569726]
32. Priyakumar UD, MacKerell AD. *J Mol Biol.* 2010; 396:1422. [PubMed: 20026131]
33. Priyakumar UD. *J Phys Chem B.* 2010; 114:9920. [PubMed: 20614931]
34. Huang W, Kim J, Jha S, Aboul-Ela F. *Nucleic Acids Res.* 2009; 37:6528. [PubMed: 19720737]
35. Doshi U, Kelley JM, Hamelberg D. *RNA.* 2012; 18:300. [PubMed: 22194311]
36. Quarta G, Sin K, Schlick T. *Plos Comput Biol.* 2012; 8:e1002368.
37. Eichhorn CD, Feng J, Suddala KC, Walter NG, Brooks CL, Al-Hashimi HM. *Nucleic Acids Res.* 2012; 40:1345. [PubMed: 22009676]
38. Feng J, Walter NG, Brooks CL. *J Am Chem Soc.* 2011; 133:4196. [PubMed: 21375305]
39. Veeraraghavan N, Ganguly A, Golden BL, Bevilacqua PC, Hammes-Schiffer S. *J Phys Chem B.* 2011; 115:8346. [PubMed: 21644800]
40. Veeraraghavan N, Ganguly A, Chen JH, Bevilacqua PC, Hammes-Schiffer S, Golden BL. *Biochemistry-US.* 2011; 50:2672.
41. Veeraraghavan N, Bevilacqua PC, Hammes-Schiffer S. *J Mol Biol.* 2010; 402:278. [PubMed: 20643139]
42. Lee TS, Giambasu GM, Harris ME, York DM. *J Phys Chem Lett.* 2011; 2:2538. [PubMed: 22200005]
43. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. *J Am Chem Soc.* 1995; 117:5179.
44. Wang JM, Cieplak P, Kollman PA. *J Comput Chem.* 2000; 21:1049.
45. Banas P, Hollas D, Zgarbova M, Jurecka P, Orozco M, Cheatham TE, Sponer J, Otyepka M. *J Chem Theory Comput.* 2010; 6:3836.
46. MacKerell AD, Banavali NK. *J Comput Chem.* 2000; 21:105.
47. Deng NJ, Cieplak P. *Biophys J.* 2010; 98:627. [PubMed: 20159159]

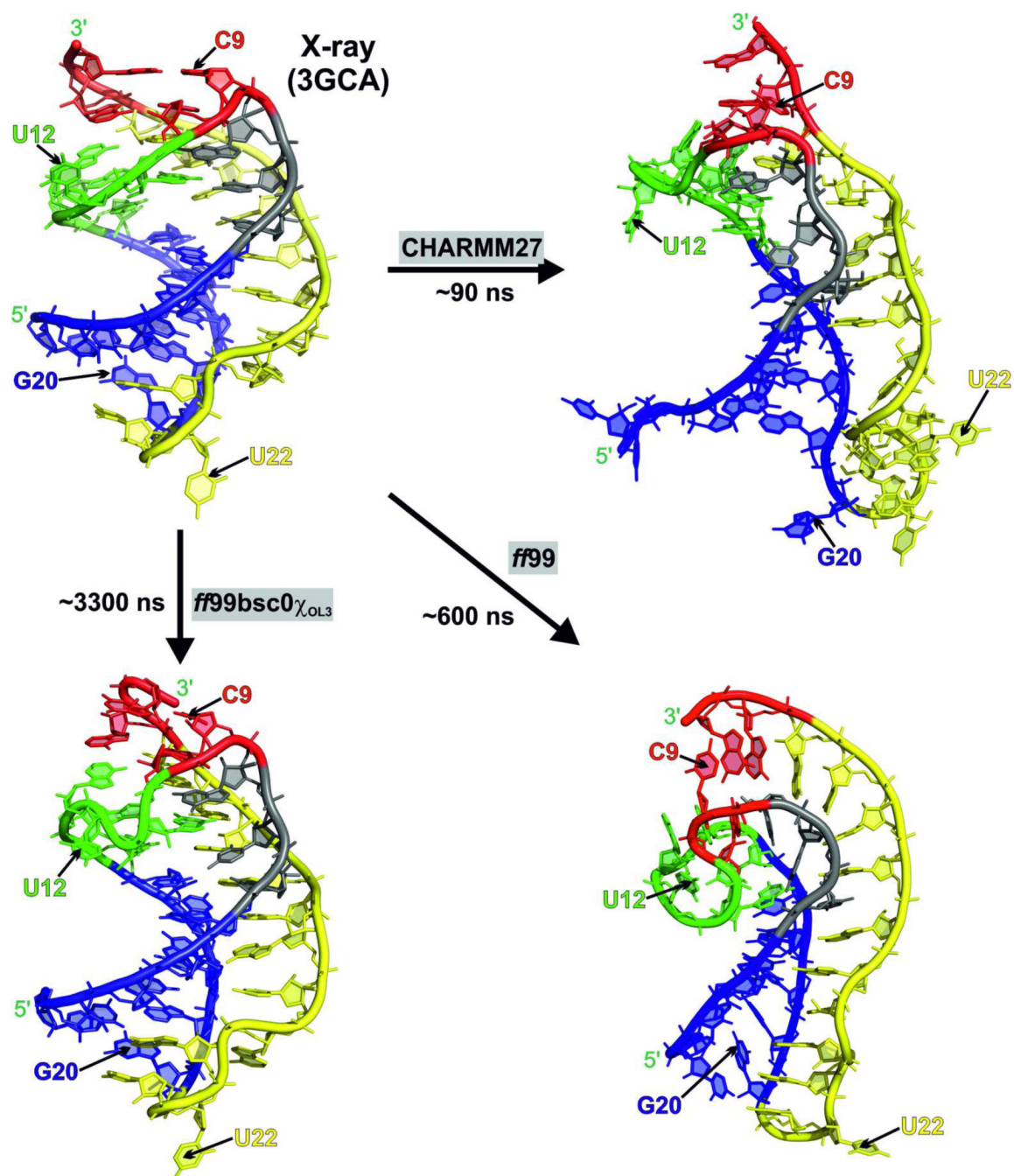
48. Faustino I, Perez A, Orozco M. *Biophys J*. 2010; 99:1876. [PubMed: 20858433]
49. Reblova K, Fadrna E, Sarzynska J, Kulinski T, Kulhanek P, Ennifar E, Koca J, Sponer J. *Biophys J*. 2007; 93:3932. [PubMed: 17704156]
50. Besseova I, Otyepka M, Reblova K, Sponer J. *Phys Chem Chem Phys*. 2009; 11:10701. [PubMed: 20145814]
51. Florova P, Sklenovsky P, Banas P, Otyepka M. *J Chem Theory Comput*. 2010; 6:3569.
52. DeLano, WL. *The PyMOL Molecular Graphics System*. Palo Alto, CA, USA: DeLano Scientific LLC; 2008.
53. Csaszar K, Spackova N, Stefl R, Sponer J, Leontis NB. *J Mol Biol*. 2001; 313:1073. [PubMed: 11700064]
54. Su L, Chen LQ, Egli M, Berger JM, Rich A. *Nat Struct Biol*. 1999; 6:285. [PubMed: 10074948]
55. Krasovska MV, Sefcikova J, Spackova N, Sponer J, Walter NG. *J Mol Biol*. 2005; 351:731. [PubMed: 16045932]
56. Ferre-D'Amare AR, Zhou KH, Doudna JA. *Nature*. 1998; 395:567. [PubMed: 9783582]
57. Nixon PL, Rangan A, Kim YG, Rich A, Hoffman DW, Hennig M, Giedroc DP. *J Mol Biol*. 2002; 322:621. [PubMed: 12225754]
58. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Debolt S, Ferguson D, Seibel G, Kollman P. *Computer Physics Communications*. 1995; 91:1.
59. Aqvist J. *J Phys Chem-US*. 1990; 94:8021.
60. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. *J Chem Phys*. 1983; 79:926.
61. Reblova K, Sponer JE, Spackova N, Besseova I, Sponer J. *J Phys Chem B*. 2011; 115:13897. [PubMed: 21999672]
62. Cornell WD, Cieplak P, Bayly CI, Kollman PA. *J Am Chem Soc*. 1993; 115:9620.
63. Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR. *J Chem Phys*. 1984; 81:3684.
64. Darden T, York D, Pedersen L. *J Chem Phys*. 1993; 98:10089.
65. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. *J Chem Phys*. 1995; 103:8577.
66. Case, DA.; Darden, TA.; Cheatham, TE., 3rd; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Walker, RC.; Zhang, W.; Merz, KM.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, KF.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, SR.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M-J.; Cui, G.; Roe, DR.; Mathews, DH.; Seetin, MG.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, PA. *AMBER 11*. San Francisco: University of California; 2010.
67. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M. *Biophys J*. 2007; 92:3817. [PubMed: 17351000]
68. Zgarbova M, Otyepka M, Sponer J, Mladek A, Banas P, Cheatham TE, Jurecka P. *Journal of Chemical Theory and Computation*. 2011; 7:2886. [PubMed: 21921995]
69. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. *J Comput Chem*. 2005; 26:1781. [PubMed: 16222654]
70. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. *J Comput Chem*. 1983; 4:187.
71. Martyna GJ, Tobias DJ, Klein ML. *J Chem Phys*. 1994; 101:4177.
72. Feller SE, Zhang YH, Pastor RW, Brooks BR. *J Chem Phys*. 1995; 103:4613.
73. Lu XJ, Olson WK. *Nat Protoc*. 2008; 3:1213. [PubMed: 18600227]
74. Humphrey W, Dalke A, Schulten K. *J Mol Graphics*. 1996; 14:33.
75. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. *P Natl Acad Sci USA*. 2001; 98:4899.
76. Zirbel CL, Sponer JE, Sponer J, Stombaugh J, Leontis NB. *Nucleic Acids Research*. 2009; 37:4898. [PubMed: 19528080]
77. Besseova I, Banas P, Kührova P, Kosinova P, Otyepka M, Sponer J. *J Phys Chem B*. 2012 *Accepted*.

78. Denning EJ, Priyakumar UD, Nilsson L, Mackerell AD. *J Comput Chem.* 2011; 32:1929. [PubMed: 21469161]
79. Olieric V, Rieder U, Lang K, Serganov A, Schulze-Briese C, Micura R, Dumas P, Ennifar E. *RNA.* 2009; 15:707. [PubMed: 19228585]
80. Lu XJ, Olson WK, Bussemaker HJ. *Nucleic Acids Res.* 2010; 38:4868. [PubMed: 20223772]
81. Mladek A, Sponer JE, Kulhanek P, Lu XJ, Olson WK, Sponer J. *J Chem Theory Comput.* 2012; 8:335. [PubMed: 22712001]
82. O'Toole AS, Miller S, Haines N, Zink MC, Serra MJ. *Nucleic Acids Res.* 2006; 34:3338. [PubMed: 16820533]
83. Banas P, Jurecka P, Walter NG, Sponer J, Otyepka M. *Methods.* 2009; 49:202. [PubMed: 19398008]
84. Zhang Q, Kang MJ, Peterson RD, Feigon J. *J Am Chem Soc.* 2011; 133:5190. [PubMed: 21410253]
85. Rieder U, Kreutz C, Micura R. *P Natl Acad Sci USA.* 2010; 107:10804.
86. Santner T, Rieder U, Kreutz C, Micura R. *J Am Chem Soc.* 2012; 134:11928. [PubMed: 22775200]
87. Banas P, Mladek A, Otyepka M, Zgarbova M, Jurecka P, Svozil D, Lankas F, Sponer J. *J Chem Theory Comput.* 2012; 8:2448.
88. Gong Z, Zhao YJ, Chen CJ, Xiao Y. *J Biomol Struct Dyn.* 2011; 29:403. [PubMed: 21875158]
89. Leontis NB, Stombaugh J, Westhof E. *Nucleic Acids Res.* 2002; 30:3497. [PubMed: 12177293]

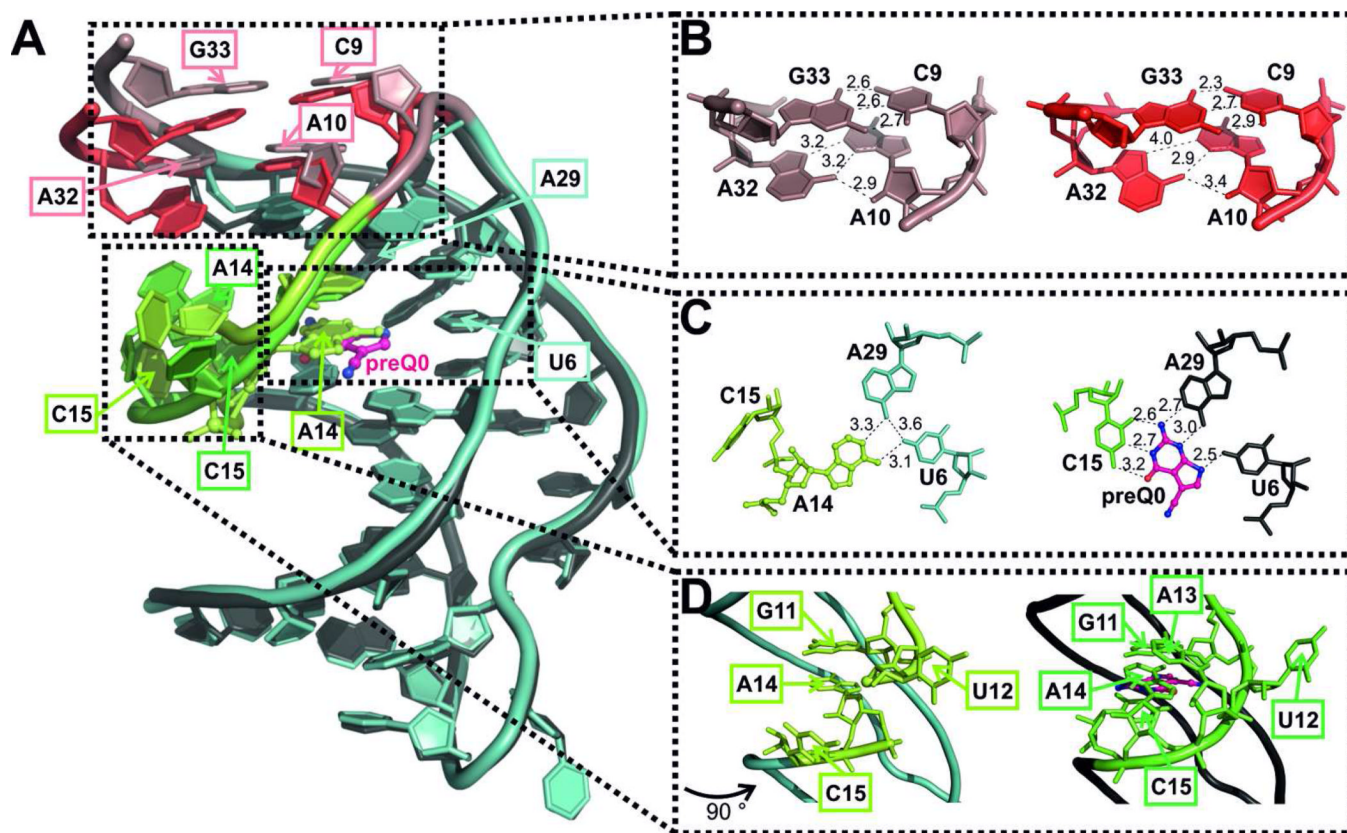




**Figure 1.** Three-dimensional (A) and secondary (B) structures of preQ<sub>1</sub> riboswitch aptamer bound to preQ<sub>0</sub> ligand. The secondary structure is annotated by standard classification<sup>76,89</sup> according to the X-ray structure (PDB ID: 3GCA). Blue, green and orange boxes show details of the ligand-binding floor, plane and ceiling, respectively.

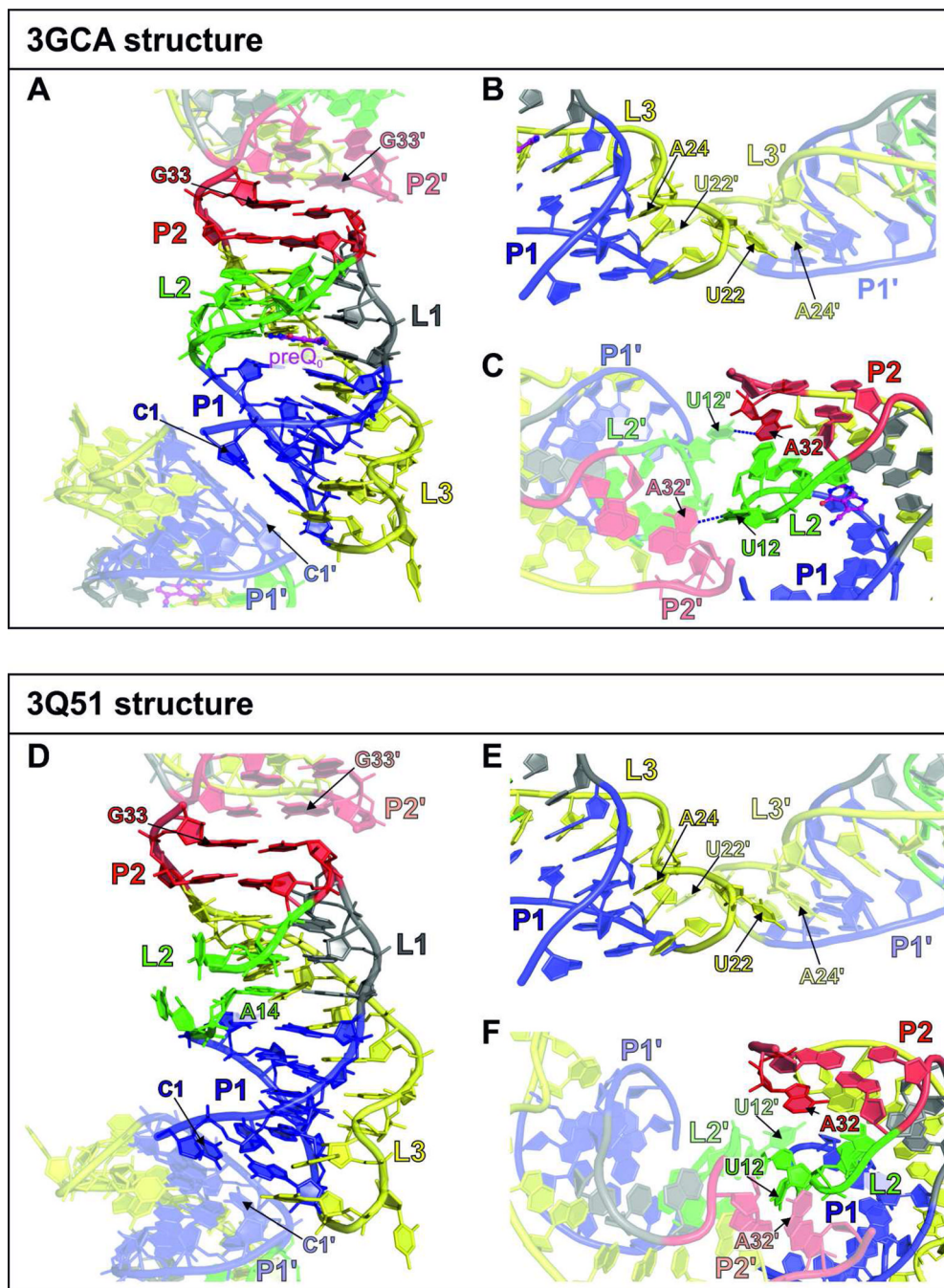


**Figure 2.** Development of the ligand-removed preQ<sub>1</sub> riboswitch aptamer simulations with different force fields, as indicated. Only *ff99bsc0 $\chi_{OL3}$*  provides stable trajectory.

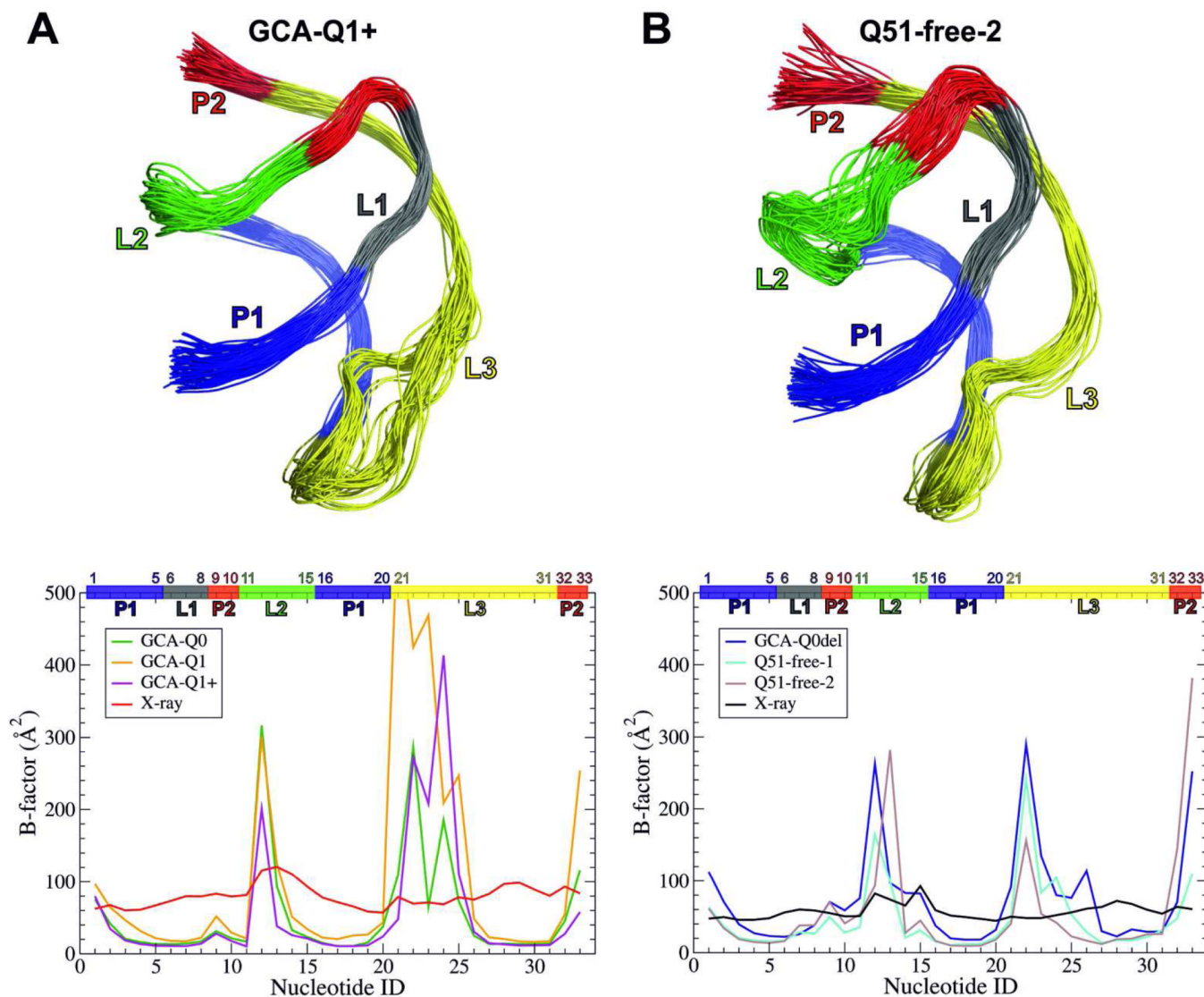


**Figure 3.**

(A) Three-dimensional structure of the ligand-bound (gray) and ligand-free (teal) preQ<sub>1</sub> riboswitch aptamer states. P2 stems of the ligand-bound and ligand-free structures are colored with red and salmon, respectively. L2 loops of the ligand-bound and ligand-free structure are colored in green and lime, respectively. The aptamers are superimposed over their P1 stems. (B) Structure of P2 stem from ligand-free (left) and ligand-bound (right) aptamer states with H-bond distances between the individual base pairs shown. (C) Depiction of the ligand-binding plane from the ligand-free aptamer with A14 occluding the binding pocket (left) and ligand-bound aptamer with the preQ<sub>0</sub> ligand bound (right). (D) Arrangement of L2 bases in the ligand-free (left) and ligand-bound (right) aptamer X-ray geometries.

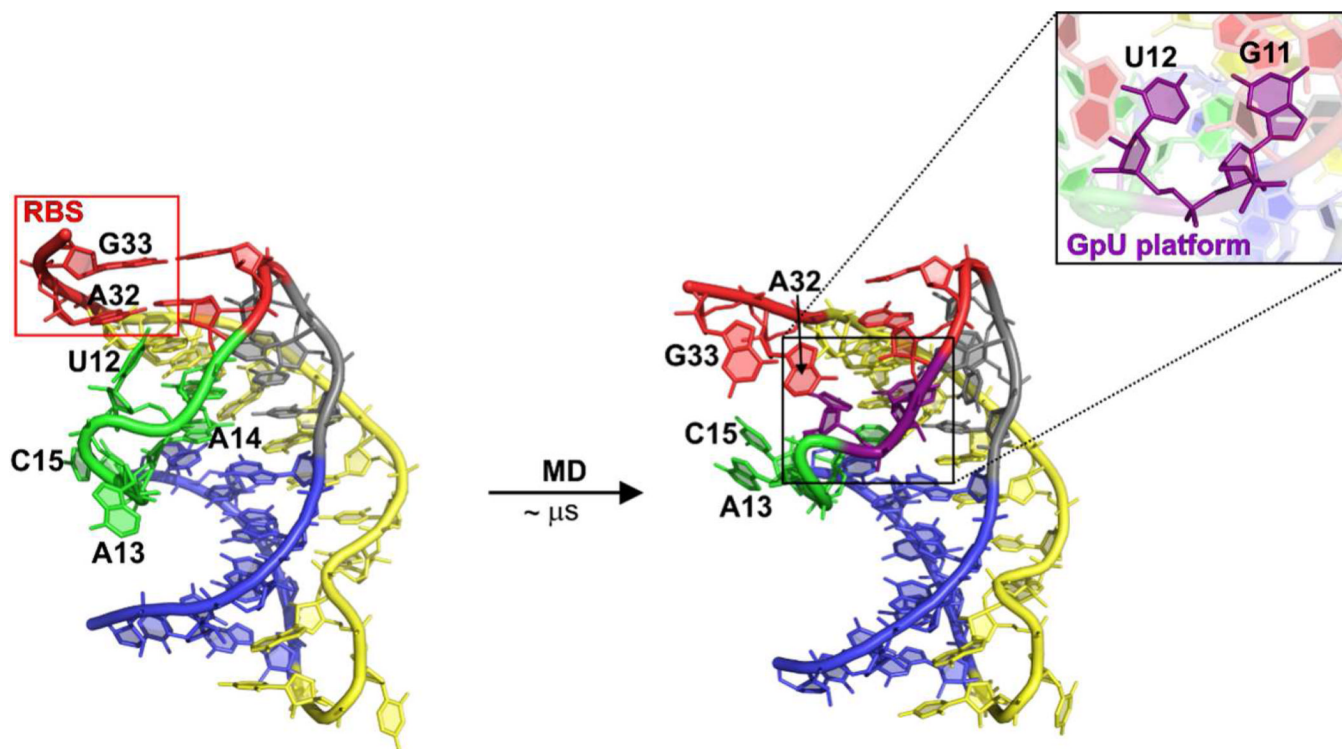


**Figure 4.** Crystal contacts visible in ligand-bound (upper part) and ligand-free (lower part) X-ray structures. Panels A and D show the coaxial stacks between the terminal base pairs of P1 stem from neighboring molecules in crystal lattice and between the terminal P2 stem base pair and its neighboring periodic image. Panels B and E depict contacts between U22 and A24 residues, while panels C and F between U12 and A32 residues.

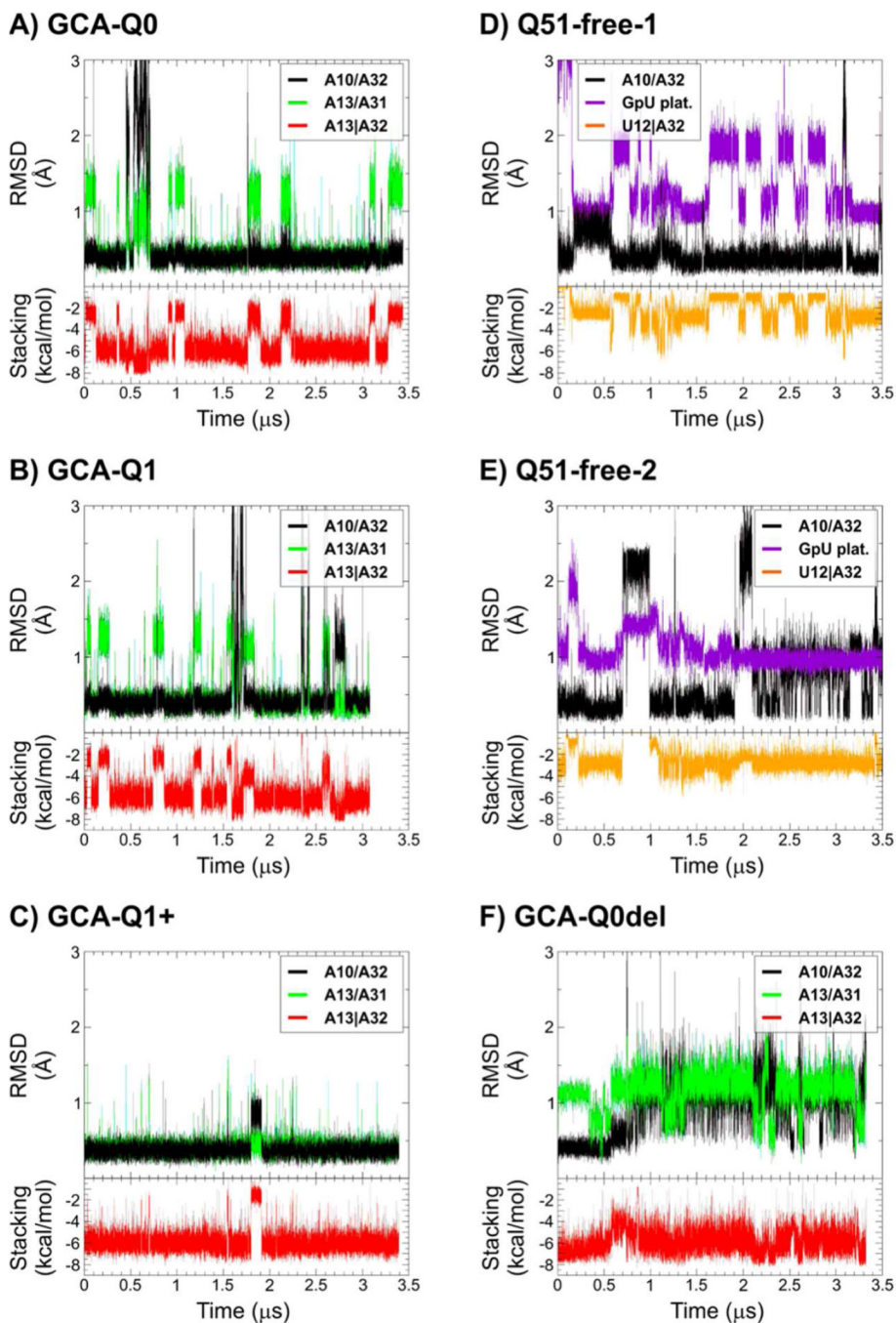


**Figure 5.**

The upper part shows the superposition of the snapshots of the preQ<sub>1</sub> riboswitch aptamer taken from A) GCA-Q1+ (ligand-bound) and B) Q51-free-2 (ligand-free) simulations, showing the rigid and flexible regions of the aptamer in each functional state. The lower part documents corresponding thermal B-factors of each aptamer's residue. In red and black are colored B-factors derived from the X-ray structures of ligand-bound (PDB ID 3GCA) and ligand-free (PDB ID 3Q51) PQA, respectively. The simulations are indicated top left. The coloring of the stripes at the top of each panel matches the colors of the stems and loops shown in Figure 1.

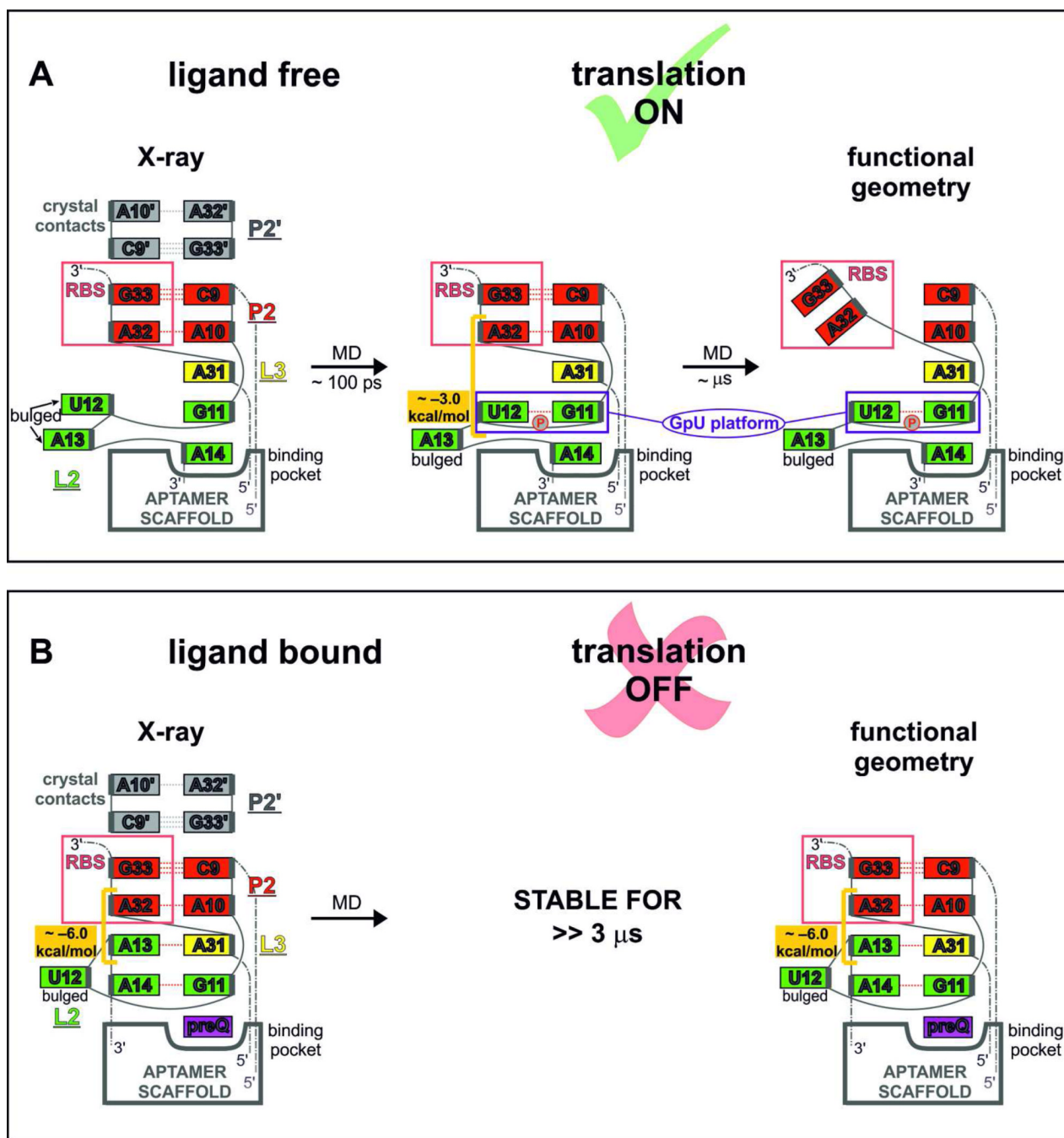


**Figure 6.** Starting X-ray (left) and functional “ON” state (right) geometry of the ligand-free PQA taken from Q51-free-2 simulation. The GpU platform, which is formed by bases U12 and G11, is highlighted in the context of the whole “ON” state PQA as well as in the close-up using the violet sticks. Note that the ribosome binding site (RBS), including bases A32 and G33, becomes considerably less structured in the “ON” state ligand-free PQA structure.



**Figure 7.**

Time developments of RMSD of the *t*HS A10/A32 (black curve) base pair and either the cWS A13/A31 base pair (green curve) or the GpU platform (violet curve). The GpU platform RMSD was calculated with respect to sarcin-ricin loop domain nucleotides G2655 and U2656 (PDB ID: 3DVZ),<sup>79</sup> which also adopt GpU platform topology in this 23S rRNA motif. Below each plot is displayed development of stacking interaction energy (Van der Waals part only) either between the bases A13 and A32 (red curve) or bases U12 and A32 (orange curve).



**Figure 8.** Suggested molecular-level model of preQ<sub>1</sub> riboswitch regulatory mechanism based on the simulation data. Van der Waals stacking energies between a given pair of bases are shown in the orange boxes.



Table 1

Overview of MD simulations<sup>a</sup>

Simulation label	Ligand	C7 prot. State	Time (μs)	Force-Field	Initial Structure	RMSD <sup>b</sup> (Å)	R <sub>g</sub> <sup>c</sup> (Å)
GCA-Q0del	removed	C7H <sup>+</sup>	3.3	<i>ff99bsc0</i> χ <sub>OL3</sub>	3GCA	2.3 ± 0.4	13.2 ± 0.3
GCA-Q0	preQ <sub>0</sub>	C7H <sup>+</sup>	3.4	<i>ff99bsc0</i> χ <sub>OL3</sub>	3GCA	1.7 ± 0.3	13.3 ± 0.2
GCA-Q1	preQ <sub>1</sub>	C7H <sup>+</sup>	3.1	<i>ff99bsc0</i> χ <sub>OL3</sub>	3GCA	2.1 ± 0.7	13.4 ± 0.3
GCA-Q1+	preQ <sub>1</sub> <sup>+</sup>	C7H <sup>+</sup>	3.4	<i>ff99bsc0</i> χ <sub>OL3</sub>	3GCA	2.0 ± 0.4	13.3 ± 0.2
Q51-free-1	no	C7H <sup>+</sup>	3.5	<i>ff99bsc0</i> χ <sub>OL3</sub>	3Q51	2.2 ± 0.3	13.3 ± 0.2
Q51-free-2	no	C7H <sup>+</sup>	3.5	<i>ff99bsc0</i> χ <sub>OL3</sub>	3Q51	2.4 ± 0.3	13.2 ± 0.2

<sup>a</sup>Some additional MD simulations are listed in Table S1 in the Supporting Information.<sup>b</sup>Mean RMSD (mass-weighted) of nucleotides 1–33 (all atoms) calculated with respect to initial (X-ray) structure and over entire trajectory of a given simulation.<sup>c</sup>Mean radius of gyration of nucleotides 1–33 (all atoms) calculated over entire trajectory of a given simulation. The X-ray R<sub>g</sub>s of the ligand-bound (PDB ID 3GCA) and ligand-free (PDB ID 3Q51) PQA are equal to 12.9 Å and 13.2 Å, respectively.

Table 2

Helical parameters of the P1 stem calculated using X3DNA<sup>73</sup> from the selected X-ray and MD-average geometries

Structure		Helical parameter <sup>a</sup>									
Type	Label/PDB ID	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (°)	Roll (°)	Twist (°)	Inclination (°)	Tip (°)		
X-ray	3GCA	0.3	-1.8	3.1	2.4	7.7	32.1	13.8	-4.1		
X-ray	3Q51	0.2	-1.7	3.3	1.7	7.7	33.5	12.8	-2.4		
MD-average	GCA-Q0del	0.1	-1.8	3.4	1.7	6.8	31.5	12.3	-3.0		
MD-average	GCA-Q0	0.1	-1.8	3.3	1.6	5.4	31.4	9.9	-2.8		
MD-average	GCA-Q1	0.1	-1.8	3.3	1.5	6.2	31.3	11.4	-2.7		
MD-average	GCA-Q1+	0.1	-1.9	3.3	1.4	5.4	31.4	9.9	-2.5		
MD-average	Q51-free-1	0	-1.8	3.3	0.9	6.7	31.1	12.3	-1.8		
MD-average	Q51-free-2	0.1	-1.9	3.3	1.0	4.8	31.2	9.0	-1.9		

<sup>a</sup>Parameters are averaged over the entire P1 stem (i.e. four central steps) and the second half of a given MD trajectory.

Table 3

Stability of H-bonds formed between the preQ<sub>0</sub>, preQ<sub>1</sub> and preQ<sub>1</sub><sup>+</sup> ligands and the riboswitch aptamer

Simulation name	Watson-Crick edge of preQ (heavy atom distances in Å) <sup>a</sup>				Sugar edge of preQ (heavy atom distances in Å)				Hoogsteen edge of preQ (heavy atom distances in Å)			
	preQ(O1)-C15(N4) <sup>d</sup>	preQ(N2)-C15(N3)	preQ(N3)-C15(O2)	preQ(N3)-A29(N1)	preQ(N4)-A29(N6)	preQ(N1)-U6(O4)	preQ(N5)-preQ(O1)	preQ(N5)-G5(O6)	preQ(N5)-G11(O2)	preQ(N5)-G11(O2)	preQ(N5)-G11(O2)	preQ(N5)-G11(O2)
GCA-Q0	3.0 ± 0.3	2.9 ± 0.2	2.9 ± 0.1	3.1 ± 0.1	2.9 ± 0.1	2.9 ± 0.1	-	-	-	-	-	-
GCA-Q1	2.9 ± 0.3	2.9 ± 0.2	2.9 ± 0.1	3.1 ± 0.2	2.9 ± 0.1	2.9 ± 0.2	3.1 ± 0.5	3.2 ± 0.5	4.3 ± 1.1	4.3 ± 1.1	4.3 ± 1.1	4.3 ± 1.1
GCA-Q1+	2.9 ± 0.1	2.9 ± 0.1	2.8 ± 0.1	3.0 ± 0.1	3.0 ± 0.1	2.9 ± 0.2	2.9 ± 0.2	2.9 ± 0.2	2.9 ± 0.2	2.9 ± 0.2	2.9 ± 0.2	3.2 ± 0.5
X-ray (preQ <sub>0</sub> lig.) <sup>b</sup>	3.2	2.7	2.6	2.7	3.0	2.5	-	-	-	-	-	-
X-ray (preQ <sub>1</sub> lig.) <sup>c</sup>	2.7	2.9	3.2	2.7	3.1	3.1	4.5	4.0	3.2 <sup>e</sup>	3.2 <sup>e</sup>	3.2 <sup>e</sup>	3.2 <sup>e</sup>

<sup>a</sup>The mean distances ± S.D. (standard deviations) were calculated over the entire trajectories.

<sup>b</sup>The respective distances were derived from the X-ray structure under the PDB ID 3GCA.

<sup>c</sup>The respective distances were derived from the X-ray structure under the PDB ID 3Q50.

<sup>d</sup>Ligand atom names are given in Figure 1B. For the sake of simplicity, the heavy atoms of -CN and -CH<sub>2</sub>NH<sub>2</sub> groups share the same names.

<sup>e</sup>Albeit the heavy atoms lie at H-bond distance, the angle between H-bond donor and acceptor is far beyond the cut-off.