Running title: Force field dependence of riboswitch dynamics

Force field dependence of riboswitch dynamics

Christian A. Hanke, Holger Gohlke*

Mathematisch-Naturwissenschaftliche Fakultät, Institut für Pharmazeutische und Medizinische Chemie, Heinrich-Heine-Universität Düsseldorf, Germany

3/2/2015

Published as: C.A. Hanke, H. Gohlke, Force Field Dependence of Riboswitch Dynamics. In Shi-Jie Chen, Donald H. Burke-Aguero, editors: Computational Methods for Understanding Riboswitches, Vol. 553, MIE, UK: Academic Press, 2015, pp. 163-191

Keywords: guanine-sensing riboswitch, aptamer domain, molecular dynamics simulations, AMBER, flexibility, structural stability, tertiary structure, magnesium ion

*Address: Universitätsstr. 1, 40225 Düsseldorf, Germany.

Phone: (+49) 211 81 13662; Fax: (+49) 211 81 13847

E-mail: gohlke@uni-duesseldorf.de

1 Abstract

Riboswitches are non-coding regulatory elements that control gene expression in response to the presence of metabolites, which bind to the aptamer domain. Metabolite binding appears to occur through a combination of conformational selection and induced-fit mechanism. This demands to characterize the structural dynamics of the *apo* state of aptamer domains. Molecular dynamics (MD) simulations can in principle give insights at the atomistic level into the dynamics of the aptamer domain. However, it is unclear to what extent contemporary force fields can bias such insights. Here we show that the Amber force field ff99 yields the best agreement with detailed experimental observations on differences in the structural dynamics of wild type and mutant aptamer domains of the guanine-sensing riboswitch (Gsw), including a pronounced influence of Mg^{2+}. In contrast, applying ff99 with parmbsc0 and parmχ_{OL} modifications (denoted ff10) results in strongly damped motions and overly stable tertiary loop-loop interactions. These results are based on 58 MD simulations with an aggregate simulation time > 11 μs, careful modeling of Mg^{2+} ions, and thorough statistical testing. Our results suggest that the moderate stabilization of the χ *anti* region in ff10 can have an unwanted damping effect on functionally relevant structural dynamics of marginally stable RNA systems. This suggestion is supported by crystal structure analyses of Gsw aptamer domains that reveal χ torsions with high-*anti* values in the most mobile regions. We expect that future RNA force field development will benefit from considering marginally stable RNA systems and optimization towards good representations of dynamics in addition to structural characteristics.

2 Introduction

Riboswitches are non-coding genetic regulatory elements mostly found in the 5' untranslated region of bacterial mRNA. Binding of a small molecule metabolite to the aptamer domain results in a large conformational change in the downstream expression platform, which modulates gene expression (Tucker & Breaker, 2005). Structural (Serganov & Nudler, 2013) and related studies have revealed that riboswitches appear to bind their ligands through a combination of conformational selection and induced-fit mechanisms (Serganov *et al.*, 2013). This demands to characterize the structural dynamics and energetics of the *apo* state of aptamer domains in order to understand the full process of gene regulation.

The role of the *apo* state is of particular importance for transcriptional gene regulation by riboswitches as there the decision whether the genes should be transcribed or not has to be made already during transcription (Stoddard, Montange, Hennelly, Rambo, Sanbonmatsu & Batey, 2010). The guanine-sensing riboswitch (Gsw) found in the xpt-pbuX operon of *B. subtilis* exerts transcriptional gene regulation and belongs to the class of purine sensing riboswitches (Batey, 2012; Batey, Gilbert & Montange, 2004). Besides guanine, Gsw binds related ligands such as hypoxanthine (Batey *et al.*, 2004; Gilbert, Stoddard, Wise & Batey, 2006). Its aptamer domain consists of three paired regions P1, P2, and P3, two loops L2 and L3 capping the P2 and P3 region, respectively, and three joining regions J1/2, J2/3, and J3/1 connecting the paired regions (Figure 1A, B). The ligand is tightly bound in the binding pocket formed by the three joining regions (Batey *et al.*, 2004). The aptamer terminal P1 helix connects to the downstream expression platform and is essential for the gene regulation (Batey *et al.*, 2004). Tertiary interactions between the L2 and L3 loops are formed by two base quadruples stacked onto each other (Figure 1C-F). These interactions were found to be present in the unbound state of the aptamer domain (Noeske, Schwalbe & Wöhnert, 2007) and have been suggested to restrict the conformational freedom of the domain, thereby promoting ligand binding under physiological

conditions (Stoddard, Gilbert & Batey, 2008). In a mutant without these tertiary interactions, the aptamer's ability to bind hypoxanthine was abolished (Batey *et al.*, 2004). In contrast, the G37A/C61U double mutation, reducing the number of hydrogen bonds within the upper base quadruple from seven in the wild type to five, led to an aptamer domain where formation of loop-loop interactions in the *apo* state and binding of hypoxanthine show a pronounced Mg^{2+} dependence (Buck, Noeske, Wöhnert & Schwalbe, 2010; Noeske, Buck, Fürtig, Nasiri, Schwalbe & Wöhnert, 2007). To what extent the loop-loop interactions contribute to the stabilization of the P1 helix, in addition to the ligand (Batey *et al.*, 2004), has not been described so far.

[Insert Figure 1 here]

With the long-term goal to gain insights at the atomistic level by molecular dynamics (MD) simulations into the function of the Gsw, we set out here to identify simulation conditions that allow appropriate modeling of the experimentally demonstrated marginal stability of the unbound aptamer domain, in particular when the G37A/C61U double mutation is present, and the pronounced influence of Mg^{2+} on the structural dynamics. With this we address one of the challenges that MD simulations of RNA face, the potential bias due to the applied force field (Sponer, Banas, Jurecka, Zgarbova, Kuhrova, Havrila *et al.*, 2014).

While contemporary force fields for RNA well describe canonical helices, they may show different successes on more complex RNAs (see refs. (Cheatham III & Case, 2013; Sponer *et al.*, 2014; Sponer, Cang & Cheatham III, 2012) for recent reviews). Here we test the Amber (Case, Cheatham III, Darden, Gohlke, Luo, Merz *et al.*, 2005) force field ff99 (Wang, Cieplak & Kollman, 2000) itself and together with the refinements parmbsc0 (Perez, Marchan, Svozil, Sponer, Cheatham III, Laughton *et al.*, 2007) and parmbsc0 + parmχOL$_3$ (Zgarbova, Otyepka, Sponer, Mladek, Banas, Cheatham III *et al.*, 2011). Note that the latter combination, ff99 +

parmbsc0 + parmχOL$_3$, has been termed ff10 in the context of RNA simulations in Amber (as of Amber12, this force field is also referred to as ff12). In parmbsc0, the α/γ dihedrals have been reparameterized in order to prevent the collapse of B-DNA due to accumulation of irreversible non-native γ-*trans* dihedral states (Perez *et al.*, 2007; Sponer *et al.*, 2014). As γ-*trans* degradation has never substantially affected canonical A-RNA simulations (Reblova, Lankas, Razga, Krasovska, Koca & Sponer, 2006) it was initially not clear if parmbsc0 brings any substantial changes for RNA simulations (Besseova, Otyepka, Reblova & Sponer, 2009). However, long-timescale MD simulations of RNA revealed sudden irreversible transitions to ladder-like structures (Mlynsky, Banas, Hollas, Reblova, Walter, Sponer *et al.*, 2010), which required a reparametrization of the χ dihedral to suppress *anti* to high-*anti* χ shifts in RNA (Sponer *et al.*, 2014), leading to the parmχOL$_3$ refinement (Zgarbova *et al.*, 2011). At present, it is recommended to use parmχOL$_3$ together with parmbsc0 (Banas, Hollas, Zgarbova, Jurecka, Orozco, Cheatham III *et al.*, 2010).

Another level of complexity arises in terms of describing the pronounced influence of Mg^{2+} ions on the structural dynamics of the aptamer domain of the Gsw (Buck *et al.*, 2010). Mg^{2+} ions play an important role for the stability of nucleic acid structures (Draper, Grilley & Soto, 2005; Pyle, 2002; Woodson, 2005). However, Mg^{2+} ions are challenging to treat in MD simulations due to the slow exchange kinetics of their first shell ligands (Ohtaki & Radnai, 1993), the low diffusion coefficient (Mills & Lobo, 1989), and difficulties in modeling polarization and charge-transfer effects with fixed-charge additive force fields (Auffinger, 2012; Sponer *et al.*, 2014). Here, we test two parameter sets for Mg^{2+} ions available for the use with Amber force fields: the long-used parameters developed by Aqvist (Aqvist, 1992) and a set of parameters developed recently with the aim to improve agreement with experimentally determined kinetic properties of Mg^{2+} ions, especially with respect to the exchange of the first solvation shell (Allner, Nilsson & Villa, 2012). Furthermore, we probe the influence of the

method of Mg^{2+} ion placement on the simulation results. In total, we report on MD simulations with an aggregate sampling > 11 μs (Table 1).

3 Methods

3.1 Setup of the simulations

The starting structures for MD simulations of the aptamer domain of the Gsw in the *apo* state were obtained from the respective ligand bound crystal structures after removal of the ligand, ions, and crystal water molecules. The sequence of the wild type structure (PDB ID: 4FE5 (Stoddard, Widmann, Trausch, Marcano-Velazquez, Knight & Batey, 2013)) was adapted to match the sequence of the G37A/C61U mutant (PDB ID 3RKF (Buck, Wacker, Warkentin, Wöhnert, Wirmer-Bartoschek & Schwalbe, 2011)). As a result, the wild type and mutant structures only differ in the nucleotides 37 and 61. The modified wild type will be referred to as Gsw^{apt}, and the G37A/C61U mutant will be referred to as Gsw^{loop}, according to Schwalbe *et al.*(Buck et al., 2011).

For both structures, three simulation systems with different Mg^{2+} concentrations (0, 12 and 20 Mg^{2+} ions per RNA molecule) were set up according to investigations on the Mg^{2+} dependence of Gsw^{loop} properties by Buck *et al.* (Buck *et al.*, 2010). For the setup, the *leap* program from the AmberTools suite of programs was used (Case, Darden, Cheatham III, Simmerling, Wang, Duke *et al.*, 2012).

For investigating the influence of the initial Mg^{2+} ion placement, one set of simulations was set up using the default placement procedure for ions of *leap*. The other simulations were set up using a modified protocol to place the Mg^{2+} ions with a first hydration shell of six water molecules: Initially, a dummy atom was placed by *leap* instead of a Mg^{2+} ion, which had a 2+ charge, but a larger radius of 4 Å. This radius is similar to the one of a hexahydrated Mg^{2+} ion (Robinson, Gao, Sanishvili, Joachimiak & Wang, 2000). Afterwards, the dummy atom was replaced by a Mg^{2+} ion with the correct radius and six water molecules surrounding it. Using a larger dummy atom initially in this two-step procedure results in a position of the Mg^{2+} ion that also allows accommodating the first hydration shell. Using a Mg^{2+} ion with the correct radius

in the initial step instead would result in the ion be placed too closely to the RNA to accommodate the waters.

Using either placement method, 12 or 20 Mg^{2+} ions with parameters of Aqvist (Aqvist, 1992) were added to Gsw^{apt} and Gsw^{loop}. To neutralize the simulation systems, also the ones without Mg^{2+} ions, Na^{+} ions (Na^{+} parameters for ff99 and ff99 + parmbsc0: (Aqvist, 1990), and for ff10: (Joung & Cheatham, 2008)) were added by *leap*. The systems were then placed in an octahedral box of TIP3P water (Jorgensen, Chandrasekhar, Madura, Impey & Klein, 1983) such that the distance between the edge of the water box and the closest RNA atom was at least 11 Å.

For each of the six simulation systems (Gsw^{apt} and Gsw^{loop} with 0, 12, or 20 Mg^{2+} ions per RNA molecule), three independent MD simulations were performed. These simulations were performed for each of the three force fields described earlier: ff99 (Wang *et al.*, 2000), ff99 + parmbsc0 (Perez *et al.*, 2007), and ff10 (Zgarbova *et al.*, 2011) (Table 1).

Before the start of the MD simulation, each system was minimized by 200 steps of steepest descent minimization, followed by 50 steps of conjugate gradient minimization. The particle mesh Ewald method (Darden, York & Pedersen, 1993) was used to treat long range electrostatic interactions, and bond lengths involving bonds to hydrogen atoms were constrained using the SHAKE algorithm (Ryckaert, Ciccotti & Berendsen, 1977). The time step for all MD simulations was 2 fs, with a direct-space nonbonded cutoff of 9 Å. Applying harmonic force restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms (and if present, also to the Mg^{2+} ions and their first hydration shell waters), we carried out canonical ensemble (NVT)-MD simulations for 50 ps, during which the system was heated from 100 to 300 K. Subsequent isothermal isobaric ensemble (NPT)-MD simulations were used for 50 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on positions of RNA, Mg^{2+} ions, and their first hydration shell waters were gradually reduced to 1 kcal mol^{-1} $Å^{-2}$ during 250 ps of NVT-MD simulations. This was followed by 50 ps of NVT-MD simulations without

applying positional restraints. From the following 200 ns of NVT-MD simulations at 300 K performed with the GPU version of *pmemd* (Salomon-Ferrer, Gotz, Poole, Le Grand & Walker, 2013) from the Amber 11 suite of programs (Case *et al.*, 2005; Case, Darden, Cheatham III, Simmerling, Wang, Duke *et al.*, 2010) conformations were extracted every 20 ps.

In order to investigate the influence of the Mg^{2+} parameters on the structural dynamics of the aptamer domain, we also performed MD simulations of Gsw^{apt} in the presence of 20 Mg^{2+} ions now using the recently developed Mg^{2+} parameters (Allner *et al.*, 2012). The simulated system was otherwise identical in terms of the initial coordinates and the force field parameters for the RNA (ff99) and the water model (TIP3P). For the system three independent MD simulations of 100 ns length each were performed.

3.2 Trajectory analysis

The MD trajectories were analyzed using the *cpptraj* tool (Roe & Cheatham III, 2013) from the AmberTools suite of programs (Case *et al.*, 2012). The initial 50 ns of each trajectory were not considered for analysis.

The radius of gyration (R_g), a measure for the compactness of a structure, was calculated for all atoms except the P1 region. Prior to the calculation of root mean square fluctuations (RMSF) over all atoms, global translational and rotational differences between the structures along the trajectory were removed by least-squares fitting onto the average structure. In order to reduce the influence of very mobile regions on the picture of internal motions (Gohlke, Kuhn & Case, 2004), conformations were root mean square-fitted on those 80% of the nucleotides with the lowest fluctuations. These 54 nucleotides, which henceforth will be referred to as "core nucleotides", were chosen from an initial calculation of the mean RMSF over the three simulations of Gsw^{apt} in the absence of Mg^{2+} ions. The root mean square deviations (RMSD) of atomic positions were calculated as a measure for the structural similarity with respect to the initial structure. The RMSD was calculated after aligning the conformations onto the starting

structure using the core nucleotides; the RMSD was then calculated for all atoms of these core nucleotides (Table 2).

The occupancy of hydrogen bonds between bases in the base quadruples of the L2/L3 region was determined using default geometrical parameters (distance: 3.5 Å; angle: 120°). The statistical uncertainty within one trajectory was determined from block averaging over blocks of 10 ns length. Finally, the values reported in Table 3 were averaged over all hydrogen bonds between a pair of bases.

In order to characterize the behavior of Mg^{2+} ions, we calculated diffusion coefficients from the mean square displacements of all Mg^{2+} ions as moving averages with varying time windows. Mg^{2+} ions with a reduced number of water molecules in the first hydration shell were identified as those that have at least one RNA atom within 3.5 Å.

Circular variances of dihedral angles were first calculated on a per-nucleotide basis and afterwards averaged over all nucleotides. The circular variance for a torsion angle θ was calculated following MacArthur and Thornton (MacArthur & Thornton, 1993) as

$$Var(\theta) = 1 - \bar{R} \quad (1)$$

where $\bar{R} = \frac{R}{n}$, n is the number of values, and R is given as

$$R^2 = (\sum_{i=1}^{n} \cos\theta_i)^2 + (\sum_{i=1}^{n} \sin\theta_i)^2 \quad (2)$$

Reported mean values (± standard error in the mean (SEM)) were calculated over the independent simulations that were performed for each system. The overall SEM was calculated from the SEMs of the single trajectories according to the laws of error propagation. For testing the equality of two mean values a Welch two-sample *t*-test (Snedecor & Cochran, 1989; Welch, 1947) was applied, which does not require that the two samples have equal variances, as implemented in the program R .

4 Results and Discussion

4.1 Placement of Mg^{2+} ions

Mg^{2+} ions have a strong influence on the structure and stability of RNA molecules (Draper *et al.*, 2005; Woodson, 2005). However, only few MD simulations of RNA systems have been carried out using Mg^{2+} ions (Auffinger, 2012; Saini, Homeyer, Fulle & Gohlke, 2013). This is a consequence of the challenges that persist in modeling these cations (see above), which result in poor sampling properties and the difficulty with current MD techniques to simulate losing of a ligand (e.g., a water or a nucleotide atom) from the inner coordination sphere as a prerequisite to form a new inner-sphere complex (e.g., with a(nother) nucleotide atom) (Auffinger, 2012). This necessitates to pay particular attention to the initial placement of Mg^{2+} ions when setting up a simulation system, particularly if no direct experimental information on the location of Mg^{2+} is available as in the case of the aptamer domain of Gsw.

Applying the default placement procedure of the *leap* program (Schafmeister, Ross & Romanovski, 1995) of the Amber software (Case *et al.*, 2005) ions are added to favorable positions near the solute according to gas-phase electrostatics. For RNA, this is usually close to the negatively charged phosphate groups of the backbone, as shown for Gsw^{apt} in Figure 2A for 20 Mg^{2+} ions after the thermalization step involving 500 ps of MD simulations. Already then all ions possess an inner-sphere hydration shell of at most five water molecules, i.e., they have at least one inner-sphere contact with an oxygen atom of the RNA backbone, likely due to the lack of a complete hydration shell around the ions after the placement. As a consequence, none of the ions moves away from its initial position during the first 100 ns of the MD simulations (Figure 2B; note that a broader ion cloud particularly in the P1 region results from ion movements coupled to RNA motions).

A particularly favorable position exists if two phosphate groups are close together in the starting structure, as given for G27 in the P2 region and G68 in the P3 region (Figure 2C). The Mg^{2+} ion placed there practically acts as a distance restraint between the two groups even over simulation times of 200 ns due to the slow exchange kinetics of inner-sphere ligands: The ion-phosphate oxygen interactions never break during that time. Such stable ion-phosphate oxygen interactions cause compaction of the surrounding structure (Figure 2D). In addition, we observe that the connected nucleotides move in a correlated manner, which extends to sequentially neighboring nucleotides (data not shown). Even if placed as a hexahydrated Mg^{2+} bound initially as an outer-sphere complex at such a position, the Mg^{2+} ion can lose a water molecule from the inner sphere and directly chelate to the phosphate groups (Figure 3A). Such an exchange is fostered by the force field bias towards inner-shell binding of Mg^{2+} ions in RNA simulations (Auffinger, 2012; Reblova *et al.*, 2006; Reblova, Spackova, Stefl, Csaszar, Koca, Leontis *et al.*, 2003).

As an alternative to the default placement, we developed a procedure in order to place Mg^{2+} ions with a first hydration shell by *leap*. Initially, we added dummy ions to the RNA with a charge equal to Mg^{2+} but a radius of 4 Å, which is close to the radius of a hexahydrated Mg^{2+} ion (Robinson *et al.*, 2000). Subsequently, these dummy ions were replaced by Mg^{2+} ions with a first hydration shell of six water molecules, and the whole system was solvated in a box of TIP3P water. Harmonic restraints were then applied to the RNA, the Mg^{2+} ions, and the first hydration shell waters during the thermalization in order to counteract ion-RNA backbone interactions. As a result, it took at least 5 ns during the production runs after removal of all restraints before the first Mg^{2+} ion lost an inner-sphere water; further water losses of other ions were observed throughout the trajectory such that after 200 ns between 60 to 80% of the Mg^{2+} ions were still hexahydrated (Figure 3A). If restraints on the ions and first hydration shell waters were omitted in the thermalization phase, occasionally Mg^{2+} ions lost one water molecule already there and formed an inner-sphere complex with the RNA (data not shown).

The hexahydrated Mg^{2+} ions show a dramatically different mobility during the MD simulations compared to the ions placed by the default procedure (Figure 2B), with ions now preferentially and thoroughly sampling the space between two backbones, but also exploring almost all of the simulation box (Figure 2E). This should allow the ions sufficient time to equilibrate prior to the formation of outer-sphere complexes with the RNA, considering fluctuation times of the bulk Mg^{2+} density of 5 to 8 ns, depending on the Mg^{2+} abundance, found by Hayes *et al.* in MD simulations of the SAM-I riboswitch (Hayes, Noel, Mohanty, Whitford, Hennelly, Onuchic *et al.*, 2012).

In line with the high mobility of hexahydrated Mg^{2+} ions, which show tumbling motions even when bound to RNA (Auffinger, Bielecki & Westhof, 2003), a much less restricted structural dynamics of Gsw^{apt} is observed now in the region between G27 and G68: the distance between the two nucleotides is up to 1.5-fold larger than when restrained by inner-sphere binding of a Mg^{2+} ion and shows fluctuations up to 3 Å (Figure 2 D), resulting in non-correlated movements of the two nucleotides (data not shown).

Thus, for all subsequent simulations, we used our optimized procedure for Mg^{2+} placement.

[Insert Figure 2 here]

4.2 Parameters of Mg^{2+} ions

Describing Mg^{2+} ions by the nonbonded model, i.e. representing their interactions by Coulombic and Lennard-Jones terms, allows switching the coordination number and ligand exchange at the metal center. This and the simplicity of the model are reasons why this model is still widely used for representing Mg^{2+} ions in biomolecular simulations, despite it oversimplifying interactions between the ions and surrounding atoms and the existence of more sophisticated models (Li, Roberts, Chakravorty & Merz, 2013). One only needs to determine

Lennard-Jones parameters for the nonbonded model. Aqvist developed Lennard-Jones parameters for Mg^{2+} more than 20 years ago (Aqvist, 1992) using the hydration free energy as objective value; these parameters have been adopted for use in Amber force fields. However, these parameters lead to an ion-water exchange rate from the first solvation shell two orders of magnitude slower than the experimental rate (Allner *et al.*, 2012). Hence, Allner *et al.*(Allner *et al.*, 2012) reparameterized the repulsive Lennard-Jones term starting from the CHARMM27 parameters for Mg^{2+} (Foloppe & MacKerell Jr., 2000; MacKerell Jr. & Banavali, 2000) with the objective to reproduce this rate. Recently, Li et al. (Li *et al.*, 2013) designed Lennard-Jones parameters for Mg^{2+} ions specifically for use in PME MD simulations, with the objective to reproduce the hydration free energy and the ion-oxygen distance in the first hydration shell ("CM set"). As absolute hydration free energies for Mg^{2+} computed with the CM set deviate more strongly from the experimental value than those computed with the Amber force field-adopted Aqvist parameters (Aqvist, 1992), we only tested the latter and the ones by Allner *et al.* (Allner *et al.*, 2012) with respect to their influence on the structural dynamics of Gsw^{apt}.

Atomic root mean-square fluctuations (RMSF) averaged at the nucleotide level (Figure 3B) show that, in general, the MD simulations with 20 Mg^{2+} ions using either one of the two parameter sets are more similar to each other than to the MD simulations without Mg^{2+} ions; the latter simulations result in the highest RMSF, indicating the lack of a stabilizing effect on the RNA in the absence of Mg^{2+} ions. This result is at variance with findings by Reblova *et al.* (Reblova *et al.*, 2006), which reported that a similar concentration of Mg^{2+} ions exerted surprisingly little effect on the dynamics of a long duplex RNA, but in agreement with findings by Reblova *et al.* of stabilizing effects of Mg^{2+} in MD simulations of the 5S rRNA Loop E (Reblova *et al.*, 2003). Comparing the two sets of MD simulations with Mg^{2+} ions, the one with parameters by Allner *et al.* shows slightly higher RMSF than the one with Aqvist parameters (Figure 3B) and, hence, a less stabilizing effect by the Mg^{2+} ions. This may reflect findings by Allner *et al.* on *add* A-riboswitch MD simulations, where their new parameters resulted in, on

average, less ions bound to the RNA via outersphere contacts than when using the CHARMM27 parameters, which result in similar ion-water characteristics as the Aqvist parameters (Allner *et al.*, 2012). From a global perspective, the stabilizing influence of the Mg^{2+} ions is also reflected in the radius of gyration (R_g) as a measure of the compactness of Gsw^{apt}: R_g = 16.5 ± 0.2 Å with the Mg^{2+} parameters from Allner *et al.* and 16.2 ± 0.1 Å with the Mg^{2+} parameters from Aqvist; for comparison, without Mg^{2+} ions, R_g = 17.1 ± 0.3 Å.

Higher exchange rates have been found to be coupled to faster ion diffusion likely because the solvent structure is disrupted during the exchange (Moller, Rey, Masia & Hynes, 2005). In line with this, our calculations of diffusion coefficients (D) over the 20 Mg^{2+} ions in the Gsw^{apt} system show 50 to 100% higher D values with the parameters of Allner *et al.* than with the ones of Aqvist for different time windows used in the moving average calculations (Figure 3C). The computed D values for very short time windows (2 ns; 0.86 * 10^{-5} cm^2 s^{-1} for the parameters of Allner *et al.*; 0.65 0.86 * 10^{-5} cm^2 s^{-1} for the parameters of Aqvist) are in very good agreement with experimental values reported for $MgSO_4$ (0.85 * 10^{-5} cm^2 s^{-1} at infinite dilution (Harned & Hudson, 1951)) and the Mg^{2+} ion (0.71 * 10^{-5} cm^2 s^{-1} (Mills *et al.*, 1989)). This suggests that the majority of ions move essentially freely within this time window over the course of the MD simulations of 200 ns. For time windows larger than 50 ns, D drops by ~1/4 in the case of Aqvist parameters and ~1/3 in the case of parameters of Allner *et al.* with respect to the initial values (Figure 3C). This is a result of a reduced ion mobility over these time windows due to the formation of inner-sphere contacts with Gsw^{apt} in the case of Aqvist parameters (between 20 to 40% of all ions; Figure 3A) and outer-sphere contacts formed by ions of both parameterizations (Figure 3A).

While no high resolution information on Mg^{2+} binding to the Gsw is available (Batey *et al.*, 2004; Buck *et al.*, 2011), it was unexpected that in neither of the three simulations with parameters of Allner *et al.* an inner-sphere contact was formed (Figure 3A), taking into account that such contacts do occur in related purine-binding riboswitches (PDB ID: 1Y26 (Serganov,

Yuan, Pikovskaya, Polonskaia, Malinina, Phan *et al.*, 2004), 3LA5 (Dixon, Duncan, Geerlings, Dunstan, McCarthy, Leys *et al.*, 2010)). The larger drop in D in the case of parameters of Allner *et al.* may reflect that these parameters result in a Mg^{2+} hydration free energy that is ~15 kcal mol^{-1} less favorable (-417. 3 kcal mol^{-1} (Allner *et al.*, 2012)) than that computed with Aqvist parameters (-432.6 kcal mol^{-1} (Li *et al.*, 2013)), employing the PME method in both cases. We speculate that this may result in a stronger coupling between the ions, observed for outer-sphere Mg^{2+} ions recently (Hayes *et al.*, 2012), which may not have become visible during MD simulations of only 10 ns length by Allner *et al.*

From the above results we find it difficult to discern that one of the two Mg^{2+} parameterizations is advantageous over the other with respect to MD simulations of the aptamer domain. We thus decided to use the Mg^{2+} parameters by Aqvist *et al.* for all subsequent simulations because of the longer experience with these parameters.

[Insert Figure 3 here]

4.3 Force field influence

4.3.1 Formation of ladder-like structures

The irreversible formation of ladder-like RNA structures in long MD simulations when using ff99 and ff99 + prmbsc0 has been described for a series of systems (Banas *et al.*, 2010; Mlynsky *et al.*, 2010; Sklenovsky, Florova, Banas, Reblova, Lankas, Otyepka *et al.*, 2011), with different simulation times until the structural collapse was observed. In our case of 18 MD simulations of Gsw^{apt} and Gsw^{loop} with ff99 of 200 ns length, we observed the irreversible formation of a ladder-like structure of the aptamer domain only once at the end of a simulation in the terminal P1 region. When extending some of these simulations to 500 ns (data not shown), all aptamer domains were stable but one. This is similar to MD simulations of the $preQ_1$

riboswitch where a collapse occurred after ~1 μs (Banas, Sklenovsky, Wedekind, Sponer & Otyepka, 2012). In the 18 MD simulations with ff99 + parmbsc0 of 200 ns length, two showed the sudden transition to a ladder-like structure after ~100 ns in the terminal P1 region. All trajectories with ladder-like structures were excluded from further analysis. Using ff10, which includes the parmχOL$_3$ refinement, neither of the 18 MD simulations showed a transition to a ladder-like structure.

4.3.2 Structural deviations of the aptamer domain

Data on the structural deviations of Gsw^{apt} and Gsw^{loop} with respect to the starting structures in terms of mean RMSD calculated over three independent simulations for each combination of force field (ff99, ff99 + parmbsc0, ff10) and number of Mg^{2+} ions (0, 12, 20) are shown in Table 2. The structural deviations are in general moderate (≤ 3 Å for all but four out of the 18 cases, where the RMSD is up to 4 Å), which agrees with findings of MD simulations on related systems (Priyakumar & MacKerell Jr., 2010; Sharma, Bulusu & Mitra, 2009; Villa, Wöhnert & Stock, 2009) and experimental data according to which Gsw^{apt} with and without Mg^{2+}, and Gsw^{loop} with Mg^{2+}, have a stable tertiary structure (Buck *et al.*, 2010). However, in simulations with ff99 and ff99 + parmbsc0, Gsw^{loop} shows in general larger RMSD values than Gsw^{apt}, and this effect is most pronounced in the absence of Mg^{2+}. This result reflects the destabilized loop-loop interactions in Gsw^{loop} (Buck *et al.*, 2010; Buck *et al.*, 2011) (see also below). In contrast, in simulations with ff10, Gsw^{apt} and Gsw^{loop} show very similar structural deviations at the respective Mg^{2+} concentrations.

4.3.3 Mobility of the aptamer domain

In order to investigate the influence of the force field on the dynamics of the RNA, we calculated atomic root mean-square fluctuations (RMSF) and averaged them per nucleotide (Figure 4A-F). The general shapes of the curves obtained for each combination of force field

and number of Mg^{2+} ions are very similar: The stem regions P2 and P3 are the least mobile, in contrast to high fluctuations observed for the aptamer terminal helix P1. Furthermore, the J2/3 region is found to be mobile, in agreement with SHAPE experiments (Stoddard *et al.*, 2008) and the suggestion that the J2/3 region acts as an entry gate for the ligand to the binding site (Gilbert *et al.*, 2006). Finally, pronounced mobility is observed in the L2 and L3 regions where tertiary interactions lead to the formation of two base quadruples.

As to differences between the curves, qualitatively, MD simulations of Gsw^{apt} (Figure 4A, C, E) show in general lower RMSF of the aptamer domain than MD simulations of Gsw^{loop} (Figure 4B, D, F); within one system, the RMSF become the lower the more Mg^{2+} ions have been added. Considering mean RMSF values over the 80% of core nucleotides (see Methods section) of the aptamer domain, these are statistically significantly different in the case of ff99 and ff99 + parmbsc0 for Gsw^{apt} *vs.* Gsw^{loop} without Mg^{2+}, as well as for Gsw^{apt} and Gsw^{loop} without *vs.* with 20 Mg^{2+} respectively (Figure 4G). These findings match perfectly with experimental observations that Gsw^{loop} without Mg^{2+} does not show formation of tertiary interactions (Buck *et al.*, 2010), with adverse effects on the structural stability of the aptamer, whereas Gsw^{apt} does; furthermore, addition of 20 Mg^{2+} results in a significant stabilization especially of the tertiary interactions in the case of Gsw^{apt} (Buck *et al.*, 2010) or the formation of the tertiary interactions in the case of Gsw^{loop} (Buck *et al.*, 2010).

Considering ff99 alone, two further differences are found significant (Figure 4G): Between Gsw^{apt} and Gsw^{loop} with 20 Mg^{2+} ions and between Gsw^{loop} without and with 12 Mg^{2+} ions. The first case reflects that tertiary interactions in Gsw^{apt} are already formed without Mg^{2+} and are further stabilized upon Mg^{2+} addition (Buck *et al.*, 2010); in contrast, in Gsw^{loop} the tertiary interactions start to form only at [Mg^{2+}]:[RNA] > 18:1 (Buck *et al.*, 2010), suggesting that the Gsw^{loop} ensemble may still be more heterogeneous even in the presence of 20 Mg^{2+}. The second case agrees with the observation that addition of a few Mg^{2+} ions (ratio ~1:5 to 1:12 (Buck *et al.*, 2011)) to Gsw^{loop} leads to only transiently formed tertiary interactions but an already ligand-

binding competent aptamer ensemble (Buck *et al.*, 2010), suggesting the beginning of structural compaction.

Likely the most unexpected observation is made for ff10 (Figure 4G). First, MD simulations with ff10 result overall in the lowest mean RMSF when comparing to results for ff99 and ff99 + parmbsc0 and the different combinations of Gsw^{apt} / Gsw^{loop} and number of Mg^{2+} ions. Second, none of the differences between the simulated systems is significant in the case of ff10, in stark contrast to what can be expected with respect to experiment (see above). Thus, this suggests that MD simulations with ff10 lead to too strongly damped aptamer motions and/or to an overly stable aptamer domain.

[Insert Figure 4 here]

In order to investigate to what extent the differences in the force fields contribute to the differences in the aptamer mobility, we computed frequency distributions of the torsion angles α, γ, and χ for Gsw^{loop} simulations in the absence of Mg^{2+}. As expected (Perez *et al.*, 2007), ff99 + parmbsc0 and ff10 show much reduced populations of the g^+ region for α (Figure 5A) and of the t region for γ (Figure 5B) in comparison to ff99, and in turn higher populations of the g^- and g^+ regions, respectively. The locations of the respective distribution maxima remained largely unchanged (α: ff99: -75°; ff99 + parmbsc0: -80°; ff10: -75°), however. Likewise expected (Zgarbova *et al.*, 2011), ff10 shows a much reduced population of the high-*anti* region of χ in comparison to ff99 and ff99 + parmbsc0 (Figure 5C), which is a result of shifting the location of the distribution maximum by ~5° towards the *anti* region plus a narrowing of the distribution function of the main population. The population changes have a pronounced influence on the spread of the torsion angles, measured in terms of the circular variance (eq. 1; Figure 5D). Note that the circular variance is bounded in the interval [0,1], with 1 denoting that all torsion angle values are spread out evenly around a circle. Highly significant

differences between circular variances are found for α and γ torsions between ff99 *vs.* ff99 + parmbsc0 or ff10; similarly, highly significant differences are found for the χ torsion between ff99 or ff99 + parmbsc0 *vs.* ff10 (Figure 5D). These findings suggest that the modifications in the α and γ torsions are not the cause for the low aptamer mobility observed with ff10, because the mean RMSF for Gsw^{loop} is not significantly ($p = 0.819$) different between ff99 and ff99 + parmsc0 (Figure 4G). In contrast, a significantly reduced circular variance of the χ torsion is observed between those simulations of Gsw^{loop} that also show a significant reduction in the aptamer mobility (Figure 4G; ff99 *vs.* ff10: $p = 0.005$; ff99 + parmbsc0 *vs.* ff10: $p = 0.003$).

Finally, we projected the per-nucleotide frequency with which a χ torsion is found in the high-*anti* region in MD simulations of Gsw^{loop} with ff99 and ff10 onto the aptamer domain (Figure 5E, F). For some nucleotides, increased frequencies are found in both trajectories (U36, U47, U48, U61, G62, A64, and A65) with frequencies > 80% for A64. Remarkably, most of the nucleotides with increased frequencies found in MD simulations with either force field also have a χ torsion in the high-*anti* region in one or both of the ligand bound Gsw^{apt} and Gsw^{loop} crystal structures (Figure 5G) (U47, G62, A64, and A65). The nucleotides with increased frequencies of their χ torsion in the high-*anti* region in MD simulations with either force field are found in those regions of the aptamer domain that are known to be the most mobile: The J2/3 region, which is suggested to act as an entry gate for the ligand to the binding site (Gilbert *et al.*, 2006) and the loop region (L2 and L3).

Differences for nucleotides with increased frequencies of the χ torsion in the high-*anti* region between MD simulations with ff99 and ff10 are found in the extended J2/3 region around A52, in the J3/1 region around C74, which forms one side of the binding site and gets stabilized upon ligand binding, and to a lesser extent in one strand of the P3 stem adjacent to the very mobile L3 loop. Here, simulations with ff99 show frequencies of a χ torsion in the high-*anti*

region of at most 25% (Figure 5E). None of these regions show χ torsions in the high-*anti* state in the Gsw^{apt} and Gsw^{loop} crystal structures (Figure 5G). Note, however, that the crystal structures represent ligand bound aptamer domains with stabilized binding sites, whereas in the MD simulations analyzed here without Mg^{2+} the aptamer domains are in the least stable states. In that respect, it appears unexpected that MD simulations with ff10 do agree so well with the crystal structures in that they show no χ torsions in the high-*anti* state in these regions either (Figure 5F). Phrased differently, together with our above results, our analyses indicate that a small propensity for visiting high-*anti* states of χ torsions is required for appropriately modeling the dynamics of known mobile regions of the Gsw aptamer domain.

[Insert Figure 5 here]

4.3.4 Loop-loop interactions

The occupancy of hydrogen bonds between bases of the two base quadruples in the L2/L3 region are reported in Table 3 for MD simulations of Gsw^{apt} and Gsw^{loop} in the absence of Mg^{2+} ions and using different force fields. The values are averages over all hydrogen bonds found between two bases. The force field dependence of the results is similar to what has been found for the force field influence on the mobility of the aptamer domain. First, for Gsw^{apt}, ff10 results in the highest occupancies, with almost permanent formation of hydrogen bonds between the bases of the upper (nucleotides 34, 37, 61 and 65; Figure 1C, D) and lower (nucleotides 33, 38, 60 and 66; Figure 1E, F) base quadruples, respectively. ff99 + parmbsc0 shows equally high occupancies for the upper quadruple but ~10 to ~20% lower occupancies for the lower quadruple. The lowest occupancies are found with ff99, with values < 90% for the upper quadruple and < 70% for the lower quadruple. Note that in this case the SEM is as high as 12%; closer inspection revealed that one out of the three independent simulations then differs with

respect to the other two, suggesting that longer simulation times will be required to achieve better converged results for the more mobile system. The exception to the above are hydrogen bonds between C61 and G37 as well as between G38 and C60, which have occupancies > 97% irrespective of the force field; the two Watson-Crick base pairs have been described as the most crucial components of the tertiary interaction (Gilbert, Love, Edwards & Batey, 2007).

Second, for Gsw^{loop}, a decrease in the occupancies with respect to Gsw^{apt} is found in general for each force field used. However, for ff10, only in the case of U61-A65 an occupancy below 60% is found, whereas such low occupancies are found for four and three base pairs in the case of ff99 + parmbsc0 and ff99, respectively. Moreover, only in the case of ff99 do we find significantly lower occupancies of ~70-77% for the two base pairs U61-A37 and G38-C60 that showed very persistent hydrogen bonds in Gsw^{apt} (see above). Finally, we find a destabilization of the lower base quadruple as a result of the mutation in the upper one only in the case of ff99 and ff99 + parmbsc0; in contrast, for ff10, we find that the occupancies of the lower base quadruple are rather comparable to the ones found for Gsw^{apt}.

The G37A/C61U mutation was introduced in Gsw^{loop} to disturb the native hydrogen bond network connecting the L2 and L3 loops (Noeske, Buck, *et al.*, 2007). Our observations of a reduced stability of tertiary interactions in Gsw^{loop} for ff99 and ff99 + parmbsc0 are in line with this. Later, NMR experiments showed that the loop-loop interactions are not formed in the absence of Mg^{2+} ions in Gsw^{loop} (Buck *et al.*, 2010), which is at variance with our non-zero occupancies. However, one needs to consider that we start from a Gsw^{loop} crystal structure that is almost indistinguishable in the L2/L3 region from Gsw^{apt} (Figure 1C-F) and that our simulation times of 200 ns are still short with respect to movements of an RNA on a rugged energy landscape (Sponer *et al.*, 2014). Finally, for Gsw^{apt}, it has been reported that, even if the tertiary interactions are already present in the absence of Mg^{2+}, addition of Mg^{2+} leads to a significant stabilization especially for the loop-loop interaction (Buck *et al.*, 2010; Buck *et al.*, 2011). Considering the hydrogen bond occupancies, it is only possible to observe such a

stabilization when using ff99: Hydrogen bond occupancies of the base quadruples in Gsw^{apt} raise to > 90% in the presence of 20 Mg^{2+} then (except for A66 and G38, where the occupancies are low throughout all simulations).

4.4 Concluding remarks

In this study, we aimed at identifying conditions for MD simulations that allow to appropriately model experimentally demonstrated differences in the structural stability of Gsw^{apt} and Gsw^{loop}, including the pronounced influence of Mg^{2+} on the structural dynamics of the aptamer domain. The most outstanding result is the strong force field dependence of the structural dynamics, with the original Amber force field ff99 yielding the best agreement with experimental observations whereas the use of ff10, implementing the parmbsc0 and parmχ_{OL} modifications, results in strongly damped aptamer motions and overly stable tertiary loop-loop interactions; ff99 + parmbsc0 yields an intermediate behavior.

We investigated the aptamer domain of the Gsw from *B. subtilis* here. The function of riboswitches inherently depends on their ability to switch between the inactive and the active state, and it requires that the aptamer domain maintains ligand binding ability also in the inactive state in order to allow for a quick response on the presence of ligands. Especially transcriptionally acting riboswitches such as the one investigated here have only a limited time frame to respond to ligand binding (Stoddard *et al.*, 2010). In order to achieve this quick response, the aptamer domain must be marginally stable, such that it can either follow the inactive route if no ligand binds, or it can quickly bind a ligand and transmit the information to the expression platform. These properties, its intricate tertiary structure, yet sufficiently small size, and the wealth of available experimental information on structure, stability, and dynamics under various conditions (Batey, 2012) make this system ideally suited for such investigations.

Our conclusions are based on detailed investigations of differences in the mobility of nucleotides and occupancies of tertiary loop-loop interactions. Ideally, our simulation results should have been compared to more direct measures of RNA dynamics, e.g. from NMR relaxation measurements (Juneja, Villa & Nilsson, 2014). Such information is not yet available however on a broad per-residue level for the unbound Gsw. Still, we believe that our conclusions are well grounded: First, we performed three independent simulations for each

investigated system leading to an aggregate simulation time of > 11 μs, which allows us to estimate the uncertainty in the simulation results; only for occupancies of tertiary interactions in the most mobile systems (Gsw^{apt} and Gsw^{loop} in the absence of Mg^{2+} and using ff99) did we find indications that our MD simulations have not yet converged. Second, we used thorough statistical analysis for evaluating the significance of differences in mean values. Third, we compared our results to detailed experimental information on the structural stability of the aptamer domain and the formation of the loop-loop interactions (Buck *et al.*, 2010; Buck *et al.*, 2011; Stoddard *et al.*, 2008). Finally, and most importantly, we analyze *differences* in the characteristics of two related RNA systems (Gsw^{apt} and Gsw^{loop}) and under various Mg^{2+} concentrations, rather than absolute characteristics of different RNA systems. Our approach should thus inherently profit from error cancelation.

The analysis of the spread of torsions angles sampled by the three force fields revealed a significantly reduced circular variance of the χ torsion between those simulations of Gsw^{loop} that also show a significant reduction in the aptamer mobility. Together with the population analysis of the χ torsion, the reduced variance is a result of a ~5-fold lower population of the high-*anti* region in ff10 than in ff99 and ff99 + parmbsc0. This is a consequence of an increased slope of the χ torsion profile in the high-*anti* region in ff10 compared to ff99 and ff99 + parmbsc0 (Zgarbova *et al.*, 2011), which was successfully introduced to suppress ladder-like RNA structures (Banas *et al.*, 2012; Sklenovsky *et al.*, 2011).

The study by Zgarbova *et al.* (Zgarbova *et al.*, 2011) demonstrated already that overstabilization of the χ *anti* region has adverse effects on the geometry of A-RNA. Compared with the original effort by Zgarbova *et al.* (Zgarbova *et al.*, 2011) to suppress ladder-like RNA structures in long MD simulations, our results suggest that even the moderate stabilization of the χ *anti* region in ff10 can have an unwanted damping effect on the functionally relevant structural dynamics of marginally stable RNA systems. Our suggestion is supported by the

finding of a high coincidence of χ torsions with high-*anti* values in the Gsw^{apt} and Gsw^{loop} crystal structures in those regions that are known to be most mobile; in contrast, MD simulations with ff10 result in very low frequencies of occurrence of χ torsions with high-*anti* values in some of the mobile regions. From the perspective of future force field development, our results suggest to include systems such as the aptamer domain of the Gsw in such studies and to extent the objectives for force field optimization beyond criteria of structural closeness to the native structure or the stabilization of signature interactions seen in experimental structures (Banas *et al.*, 2010; Perez *et al.*, 2007; Sklenovsky *et al.*, 2011; Zgarbova *et al.*, 2011) towards agreement of (differences in) the structural dynamics with experiment.

Finally, from our investigations on the influence of the setup of Mg^{2+} ions on the structural dynamics of Gsw^{apt} we strongly recommend not to use the default procedure implemented in the *leap* program of the Amber suite for placing Mg^{2+} around a nucleic acid structure but rather to use *leap* to place larger dummy ions that are later replaced by hexahydrated Mg^{2+}. Regarding the use of different Mg^{2+} ion parameters, we observed only a small influence on the structural dynamics of Gsw^{apt}. Our observation of no inner-sphere contact formation in any of the MD simulations when using the recently developed Mg^{2+} parameters of Allner *et al.* was unexpected. However, further investigations including other RNA systems are required to confirm this finding.

5 Acknowledgements

We gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUROPA at Jülich Supercomputing Center (JSC) (NIC project 4722). Additional computational support was provided by the "Center for Information and Media Technology" (ZIM) at the Heinrich-Heine-University of Düsseldorf (Germany).

6 References

Allner, O., Nilsson, L., & Villa, A. (2012). Magnesium ion-water coordination and exchange in biomolecular simulations. *Journal of Chemical Theory and Computation, 8*(4), 1493-1502

Aqvist, J. (1990). Ion water interaction potentials derived from free-energy perturbation simulations. *Journal of Physical Chemistry, 94*(21), 8021-8024

Aqvist, J. (1992). Modeling of ion ligand interactions in solutions and biomolecules. *Theochem-Journal of Molecular Structure, 88*, 135-152

Auffinger, P. (2012). Ions in molecular dynamics simulations of RNA systems. In N. Leontis & E. Westhof (Eds.), *RNA 3D structure analysis and prediction* (pp. 299-318): Springer.

Auffinger, P., Bielecki, L., & Westhof, E. (2003). The Mg^{2+} binding sites of the 5S rRNA loop E motif as investigated by molecular dynamics simulations. *Chemistry & Biology, 10*(6), 551-561

Banas, P., Hollas, D., Zgarbova, M., Jurecka, P., Orozco, M., Cheatham III, T. E., *et al.* (2010). Performance of molecular mechanics force fields for RNA simulations: Stability of UUCG and GNRA hairpins. *Journal of Chemical Theory and Computation, 6*(12), 3836-3849

Banas, P., Sklenovsky, P., Wedekind, J. E., Sponer, J., & Otyepka, M. (2012). Molecular mechanism of $preQ_1$ riboswitch action: A molecular dynamics study. *Journal of Physical Chemistry B, 116*(42), 12721-12734

Batey, R. T. (2012). Structure and mechanism of purine-binding riboswitches. *Quarterly Reviews of Biophysics, 45*(3), 345-381

Batey, R. T., Gilbert, S. D., & Montange, R. K. (2004). Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature, 432*(7015), 411-415

Besseova, I., Otyepka, M., Reblova, K., & Sponer, J. (2009). Dependence of A-RNA simulations on the choice of the force field and salt strength. *Physical Chemistry Chemical Physics, 11*(45), 10701-10711

Blackburn, G. M., & Gait, M. J. (1996). *Nucleic acids in chemistry and biology*. Oxford: Oxford University Press.

Buck, J., Noeske, J., Wöhnert, J., & Schwalbe, H. (2010). Dissecting the influence of Mg^{2+} on 3D architecture and ligand-binding of the guanine-sensing riboswitch aptamer domain. *Nucleic Acids Research, 38*(12), 4143-4153

Buck, J., Wacker, A., Warkentin, E., Wöhnert, J., Wirmer-Bartoschek, J., & Schwalbe, H. (2011). Influence of ground-state structure and Mg^{2+} binding on folding kinetics of the guanine-sensing riboswitch aptamer domain. *Nucleic Acids Research, 39*(22), 9768-9778

Case, D. A., Cheatham III, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., *et al.* (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry, 26*(16), 1668-1688

Case, D. A., Darden, T. A., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., *et al.* (2010). AMBER 11. University of California, San Francisco.

Case, D. A., Darden, T. A., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., *et al.* (2012). AMBER 13. University of California, San Francisco.

Cheatham III, T. E., & Case, D. A. (2013). Twenty-five years of nucleic acid simulations. *Biopolymers, 99*(12), 969-977

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *Journal of Chemical Physics, 98*(12), 10089-10092

Dixon, N., Duncan, J. N., Geerlings, T., Dunstan, M. S., McCarthy, J. E. G., Leys, D., *et al.* (2010). Reengineering orthogonally selective riboswitches. *Proceedings of the National Academy of Sciences of the United States of America, 107*(7), 2830-2835

Draper, D. E., Grilley, D., & Soto, A. M. (2005). Ions and RNA folding. *Annual Review of Biophysics and Biomolecular Structure, 34*, 221-243

Foloppe, N., & MacKerell Jr., A. D. (2000). All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry, 21*(2), 86-104

Gilbert, S. D., Love, C. E., Edwards, A. L., & Batey, R. T. (2007). Mutational analysis of the purine riboswitch aptamer domain. *Biochemistry, 46*(46), 13297-13309

Gilbert, S. D., Stoddard, C. D., Wise, S. J., & Batey, R. T. (2006). Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *Journal of Molecular Biology, 359*(3), 754-768

Gohlke, H., Kuhn, L. A., & Case, D. A. (2004). Change in protein flexibility upon complex formation: Analysis of Ras-Raf using molecular dynamics and a molecular framework approach. *Proteins-Structure Function and Bioinformatics, 56*(2), 322-337

Harned, H. S., & Hudson, R. M. (1951). The diffusion coefficient of magnesium sulfate in dilute aqueous solution at 25°. *Journal of the American Chemical Society, 73*(12), 5880-5882

Hayes, R. L., Noel, J. K., Mohanty, U., Whitford, P. C., Hennelly, S. P., Onuchic, J. N., *et al.* (2012). Magnesium fluctuations modulate RNA dynamics in the SAM-I riboswitch. *Journal of the American Chemical Society, 134*(29), 12043-12053

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics, 79*(2), 926-935

Joung, I. S., & Cheatham, T. E., III. (2008). Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *Journal of Physical Chemistry B, 112*(30), 9020-9041

Juneja, A., Villa, A., & Nilsson, L. (2014). Elucidating the relation between internal motions and dihedral angles in an RNA hairpin using molecular dynamics. *Journal of Chemical Theory and Computation, 10* (8), 3532-3540

Li, P. F., Roberts, B. P., Chakravorty, D. K., & Merz, K. M. (2013). Rational design of particle mesh Ewald compatible lennard-jones parameters for +2 metal cations in explicit solvent. *Journal of Chemical Theory and Computation, 9*(6), 2733-2748

MacArthur, M. W., & Thornton, J. M. (1993). Conformational analysis of protein structures derived from NMR data. *Proteins, 17*(3), 232-251

MacKerell Jr., A. D., & Banavali, N. K. (2000). All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *Journal of Computational Chemistry, 21*(2), 105-120

Mills, R., & Lobo, V. M. M. (1989). *Self-diffusion in electrolyte solutions: A critical examination of data compiled from the literature.* Amsterdam: Elsevier.

Mlynsky, V., Banas, P., Hollas, D., Reblova, K., Walter, N. G., Sponer, J., *et al.* (2010). Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H$^+$ forms are most consistent with crystal structures of hairpin ribozyme. *Journal of Physical Chemistry B, 114*(19), 6642-6652

Moller, K. B., Rey, R., Masia, M., & Hynes, J. T. (2005). On the coupling between molecular diffusion and solvation shell exchange. *Journal of Chemical Physics, 122*(11)

Noeske, J., Buck, J., Fürtig, B., Nasiri, H. R., Schwalbe, H., & Wöhnert, J. (2007). Interplay of 'induced fit'and preorganization in the ligand induced folding of the aptamer domain of the guanine binding riboswitch. *Nucleic Acids Research, 35*(2), 572-583

Noeske, J., Schwalbe, H., & Wöhnert, J. (2007). Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain. *Nucleic Acids Research, 35*(15), 5262-5273

Ohtaki, H., & Radnai, T. (1993). Structure and dynamics of hydrated ions. *Chemical Reviews, 93*(3), 1157-1204

Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham III, T. E., Laughton, C. A., *et al.* (2007). Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers. *Biophysical Journal, 92*(11), 3817-3829

Priyakumar, U., & MacKerell Jr., A. D. (2010). Role of the adenine ligand on the stabilization of the secondary and tertiary interactions in the adenine riboswitch. *Journal of Molecular Biology, 396*(5), 1422-1438

Pyle, A. M. (2002). Metal ions in the structure and function of RNA. *Journal of Biological Inorganic Chemistry, 7*(7-8), 679-690

Reblova, K., Lankas, F., Razga, F., Krasovska, M. V., Koca, J., & Sponer, J. (2006). Structure, dynamics, and elasticity of free 16S rRNA helix 44 studied by molecular dynamics simulations. *Biopolymers, 82*(5), 504-520

Reblova, K., Spackova, N., Stefl, R., Csaszar, K., Koca, J., Leontis, N. B., *et al.* (2003). Non-Watson-Crick basepairing and hydration in RNA motifs: Molecular dynamics of 5S rRNA loop E. *Biophysical Journal, 84*(6), 3564-3582

Robinson, H., Gao, Y. G., Sanishvili, R., Joachimiak, A., & Wang, A. H. J. (2000). Hexahydrated magnesium ions bind in the deep major groove and at the outer mouth of A-form nucleic acid duplexes. *Nucleic Acids Research, 28*(8), 1760-1766

Roe, D. R., & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation, 9*(7), 3084-3095

Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Physics, 23*(3), 327-341

Saini, J. S., Homeyer, N., Fulle, S., & Gohlke, H. (2013). Determinants of the species selectivity of oxazolidinone antibiotics targeting the large ribosomal subunit. *Biological Chemistry, 394*(11), 1529-1541

Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of Chemical Theory and Computation, 9*(9), 3878-3888

Schafmeister, C. E. A. F., Ross, W. S., & Romanovski, V. (1995). Leap. University of California, San Francisco

Serganov, A., & Nudler, E. (2013). A decade of riboswitches. *Cell, 152*(1-2), 17-24

Serganov, A., Yuan, Y. R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., *et al.* (2004). Structural basis for discriminative regulation of gene expression by adenine-and guanine-sensing mrnas. *Chemistry & Biology, 11*(12), 1729-1741

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*(9), 1673-1692

Sklenovsky, P., Florova, P., Banas, P., Reblova, K., Lankas, F., Otyepka, M., *et al.* (2011). Understanding RNA flexibility using explicit solvent simulations: The ribosomal and group I intron reverse kink-turn motifs. *Journal of Chemical Theory and Computation, 7*(9), 2963-2980

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods*: Iowa State University Press.

Sponer, J., Banas, P., Jurecka, P., Zgarbova, M., Kuhrova, P., Havrila, M., *et al.* (2014). Molecular dynamics simulations of nucleic acids. From tetranucleotides to the ribosome. *Journal of Physical Chemistry Letters, 5*(10), 1771-1782

Sponer, J., Cang, X. H., & Cheatham III, T. E. (2012). Molecular dynamics simulations of G-DNA and perspectives on the simulation of nucleic acid structures. *Methods, 57*(1), 25-39

Stoddard, C. D., Gilbert, S. D., & Batey, R. T. (2008). Ligand-dependent folding of the three-way junction in the purine riboswitch. *RNA, 14*(4), 675-684

Stoddard, C. D., Montange, R. K., Hennelly, S. P., Rambo, R. P., Sanbonmatsu, K. Y., & Batey, R. T. (2010). Free state conformational sampling of the SAM-I riboswitch aptamer domain. *Structure, 18*(7), 787-797

Stoddard, C. D., Widmann, J., Trausch, J. J., Marcano-Velazquez, J. G., Knight, R., & Batey, R. T. (2013). Nucleotides adjacent to the ligand-binding pocket are linked to activity tuning in the purine riboswitch. *Journal of Molecular Biology, 425*(10), 1596-1611

Tucker, B. J., & Breaker, R. R. (2005). Riboswitches as versatile gene control elements. *Current Opinion in Structural Biology, 15*(3), 342-348

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 Research, 37*(14), 4774-4786

Wang, J. M., Cieplak, P., & Kollman, P. A. (2000). How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry, 21*(12), 1049-1074

Welch, B. L. (1947). The generalization of students problem when several different population variances are involved. *Biometrika, 34*(1-2), 28-35

Woodson, S. A. (2005). Metal ions and RNA folding: A highly charged topic with a dynamic future. *Current Opinion in Chemical Biology, 9*(2), 104-109

Zgarbova, M., Otyepka, M., Sponer, J., Mladek, A., Banas, P., Cheatham III, T. E., *et al.* (2011). Refinement of the Cornell et al. Nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *Journal of Chemical Theory and Computation, 7*(9), 2886-2902

7 Tables

Table 1: List of MD simulations presented in this work.

Gsw variant	Number of Mg^{2+} ions [a]	Force field	Simulation time [e]	Mg^{2+} parameters
Gsw^{apt}	0	ff99 [b]	3 x 200	Aqvist [f]
	12	ff99 [b]	3 x 200	Aqvist [f]
	20	ff99 [b]	3 x 200	Aqvist [f]
Gsw^{loop}	0	ff99 [b]	3 x 200	Aqvist [f]
	12	ff99 [b]	3 x 200	Aqvist [f]
	20	ff99 [b]	3 x 200	Aqvist [f]
Gsw^{apt}	0	parmbsc0 [c]	3 x 200	Aqvist [f]
	12	parmbsc0 [c]	3 x 200	Aqvist [f]
	20	parmbsc0 [c]	3 x 200	Aqvist [f]
Gsw^{loop}	0	parmbsc0 [c]	3 x 200	Aqvist [f]
	12	parmbsc0 [c]	3 x 200	Aqvist [f]
	20	parmbsc0 [c]	3 x 200	Aqvist [f]
Gsw^{apt}	0	ff10 [d]	3 x 200	Aqvist [f]
	12	ff10 [d]	3 x 200	Aqvist [f]
	20	ff10 [d]	3 x 200	Aqvist [f]
Gsw^{loop}	0	ff10 [d]	3 x 200	Aqvist [f]
	12	ff10 [d]	3 x 200	Aqvist [f]
	20	ff10 [d]	3 x 200	Aqvist [f]
Gsw^{apt}	20	ff99 [b]	3 x 100	Allner *et al.* [g]
Gsw^{apt}	20	ff99 [b]	1 x 100	Aqvist [f,h]

[a] Per RNA molecule.
[b] (Wang *et al.*, 2000)
[c] (Perez *et al.*, 2007)
[d] (Zgarbova *et al.*, 2011)
[e] In ns.
[f] (Aqvist, 1992)
[g] (Allner *et al.*, 2012)
[h] The default placement procedure of *leap* was used for the Mg^{2+} ions in this simulation.

Table 2: RMSD of the core nucleotides of Gsw^{apt} and Gsw^{loop} using different force fields.[a]

Simulated system		ff99	bsc0	ff10
Gsw^{apt}	0 Mg^{2+} [b]	2.9 ± 0.4	3.5 ± 0.5	2.7 ± 0.3
	12 Mg^{2+} [b]	2.7 ± 0.4	2.8 ± 0.3	2.3 ± 0.3
	20 Mg^{2+} [b]	2.1 ± 0.3	3.0 ± 0.4	2.2 ± 0.3
Gsw^{loop}	0 Mg^{2+} [b]	3.7 ± 0.6	4.1 ± 0.7	2.7 ± 0.3
	12 Mg^{2+} [b]	2.3 ± 0.3	2.9 ± 0.3	2.5 ± 0.3
	20 Mg^{2+} [b]	2.5 ± 0.3	3.6 ± 0.3	2.6 ± 0.2

[a] In Å; mean ± SEM over three trajectories of 200 ns length. Nucleotides 5-21, 23-33, 36-47, and 51-64 are taken as core nucleotides. All atoms of the aptamer domain were considered.
[b] Number of Mg^{2+} ions per RNA molecule.

Table 3: Hydrogen bond occupancy in the L2/L3 loop region of Gsw^{apt} and Gsw^{loop} using different force fields.[a]

RNA	Pair of nucleotides	ff99	bsc0	ff10
Gsw^{apt}	U34 - A65	80.9 ± 10.6	96.9 ± 0.7	95.7 ± 1.0
	U34 - G37	86.1 ± 10.5	99.6 ± 0.2	99.2 ± 0.3
	C61 - A65	74.9 ± 12.0	98.4 ± 0.6	99.3 ± 0.2
	C61 - G37	99.6 ± 0.3	99.9 ± 0.0	99.9 ± 0.0
	A33 - A66	69.4 ± 12.6	88.0 ± 7.5	97.1 ± 1.4
	G38 - A66	57.5 ± 9.1	63.9 ± 3.4	80.2 ± 2.6
	G38 - C60	97.8 ± 0.8	97.2 ± 0.5	97.8 ± 0.5
Gsw^{loop}	U34 - A65	66.1 ± 11.9	16.0 ± 10.4	87.6 ± 10.1
	U61 - A65	13.1 ± 11.1	28.7 ± 14.0	56.9 ± 11.1
	U61 - A37	70.2 ± 7.9	95.9 ± 1.6	67.9 ± 6.1
	A33 - A66	50.0 ± 15.0	1.5 ± 1.7	85.2 ± 10.1
	G38 - A66	41.6 ± 15.1	54.8 ± 10.3	84.5 ± 7.7
	G38 - C60	77.0 ± 10.0	98.4 ± 1.0	99.2 ± 0.4

[a] In %; mean ± SEM over three trajectories of 200 ns length in the absence of Mg2+ ions. The nucleotides 34, 37, 61, and 65 belong to the upper base quadruple (Figure 1C, D) where the G37A/C61U mutation is located; the nucleotides 33, 38, 60, and 66 form the lower base quadruple (Figure 1E,F).

Figure captions

Figure 1: A: Secondary structure and sequence used in this study of the aptamer domain of Gsw^{apt}; the secondary structure elements are assigned according to Batey *et al.* (Batey *et al.*, 2004). In addition, the ligand binding site is indicated by the purple oval and denoted by LIG. B: Tertiary structure of the aptamer domain of the Gsw bound to hypoxanthine (LIG) (PDB ID 4FE5 (Stoddard *et al.*, 2013)). The secondary structure elements are assigned as in panel A. The black rectangle indicates the location of the base quadruples. C, D: Upper base quadruple formed by the L2 and L3 loops for Gsw^{apt} (C) and Gsw^{loop} (D). Here, the G37A/C61U mutation is located (highlighted by the red labels). E, F: Lower base quadruple formed by the L2 and L3 loops for Gsw^{apt} (E) and Gsw^{loop} (F).

Figure 2: A: Initial placement of 20 Mg^{2+} ions (green spheres) and surrounding water molecules (sticks) around the structure of the aptamer domain of the Gsw (with the backbone displayed as grey line). The ions were placed prior to the water molecules at electrostatically favorable locations as implemented in *leap* (Case *et al.*, 2005) using the default procedure and are thus located very closely to the RNA backbone. B,E: Positions of 20 Mg^{2+} ions during 100 ns of MD simulation. Different colors represent different Mg^{2+} ions. The ions were initially placed by *leap* either using the default procedure (B) or using the improved placement procedure (see Methods section) considering a first hydration shell (E). C: Close-up of a Mg^{2+} ion connecting the OP1 atoms of nucleotides G27 and G68 (red sticks connected to the RNA backbone in grey). Four water molecules (red and white sticks) form the hydration shell of the Mg^{2+} ion. D: Distance calculated between the P atoms of nucleotides G27 and G68 shown in (C) during MD simulations of 200 ns length. The blue dotted line represents the distance in a simulation with the default placement of the ions; the three red solid lines, which were smoothed for better visibility, represent the distances in simulations with the improved placement procedure.

Figure 3: A: Number of Mg^{2+} ions having six water molecules during Gsw^{apt} simulations with Mg^{2+} ions. The black dashed lines at 20 represent results for the three simulations using the Mg^{2+} parameters from Allner *et al.* (Allner *et al.*, 2012); the grey solid and dotted lines represent results for the three simulations using the Mg^{2+} parameters from Aqvist (Aqvist, 1992). B: Per-nucleotide RMSF values obtained from MD simulations of Gsw^{apt} without Mg^{2+} ions (green dotted line with filled circle), with 20 Mg^{2+} ions with Mg^{2+} parameters from Aqvist (blue dashed line with filled triangle), or with 20 Mg^{2+} ions with Mg^{2+} parameters from Allner *et al.* (red solid line with 'X' symbol); the mean ± SEM over three trajectories of 100 ns length is given. C: Moving average of the diffusion coefficient (± SEM) over 20 Mg^{2+} ions calculated for varying time windows along three trajectories of 100 ns length of Gsw^{apt} simulations. The black solid line corresponds to the simulations using the Mg^{2+} parameters from Aqvist; the grey dashed line corresponds to the simulations using the Mg^{2+} parameters from Allner *et al.*

Figure 4: A-F: Per-nucleotide RMSF obtained from MD simulations of Gsw^{apt} (A, C, and E) and Gsw^{loop} (B, D, and F) using different force fields and different numbers of Mg^{2+} ions; the mean ± SEM over three trajectories of 200 ns length is given. ff99: red solid line with 'X' symbol; parmbsc0: blue dashed line with filled triangle; ff10: green dotted line with filled circle. A, B: absence of Mg^{2+} ions; C, D: 12 Mg^{2+} ions per RNA molecule; E,F: 20 Mg^{2+} ions per RNA molecule. G: Mean RMSF ± SEM of core nucleotides for the simulations shown in panels A-F grouped by the used force field. The stars indicate statistically significant differences between the mean values (*: $p < 0.05$; **: $p < 0.01$).

Figure 5: A,B,C: Normalized histograms of torsion angles α (A), γ (B), and χ (C) in MD simulations of Gsw^{loop} in the absence of Mg^{2+} ions using different force fields. Red: ff99; blue: parmbsc0; green: ff10. D: Mean circular variance (eq. 1) ± SEM in the torsion angles α, γ, and χ for the force fields ff99, parmbsc0 (bsc0) and ff10. The stars indicate statistically significant differences between the mean values (*: $p < 0.001$). g^+, g^-, and t regions are assigned according to (Perez *et al.*, 2007), the *anti* and high-*anti* regions are assigned according to (Blackburn & Gait, 1996). E, F: Per-nucleotide frequency of the χ torsion angle being in the high-anti region (between -110° and -60°) for simulations using the ff99 (E) and the ff10 (F) force field. Frequency values between 0 (blue) and 50% (red) are mapped onto the structure of Gsw^{loop} according to the color scale. Nucleotides A52 and A64 (G62, A64, and A65) showed frequencies > 50% and up to 90% (90%) in the case of ff99 (ff10). G: Occupation of the high-*anti* region of the χ torsion angle in the ligand bound crystal structures of Gsw^{apt} (PDB ID: 4FE5 (Stoddard *et al.*, 2013)) and Gsw^{loop} (PDB ID: 3RKF (Buck *et al.*, 2011)). Nucleotides whose χ torsion angles are not in the high-*anti* region in either of the crystal structures are colored in blue. The other residues are colored according to the crystal structure in which their χ torsion angle is in the high-*anti* region: Gsw^{apt} only (light pink), Gsw^{loop} only (magenta), or both (red).**

8 Figures

Figure 1 color

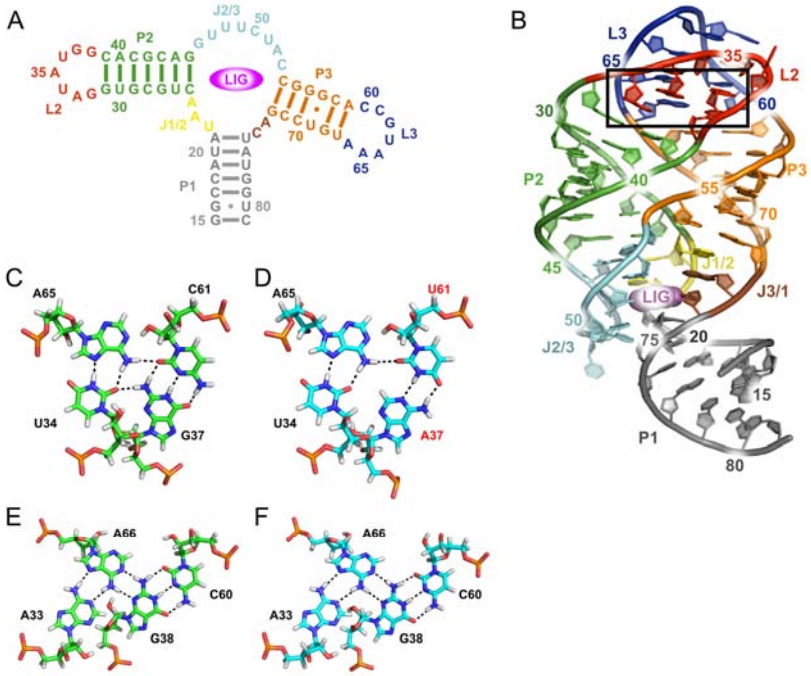

Figure 2 color

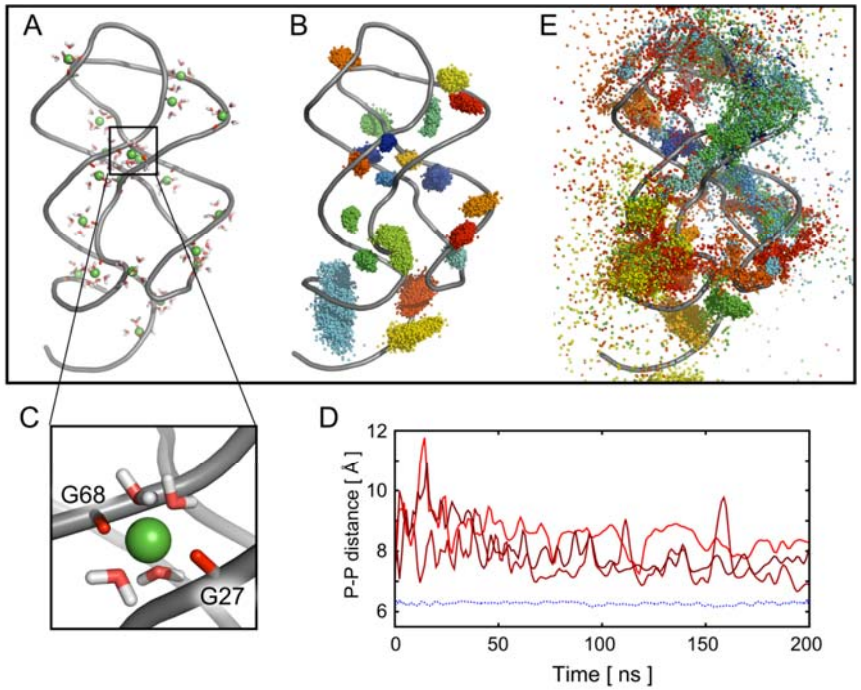

Figure 3 color

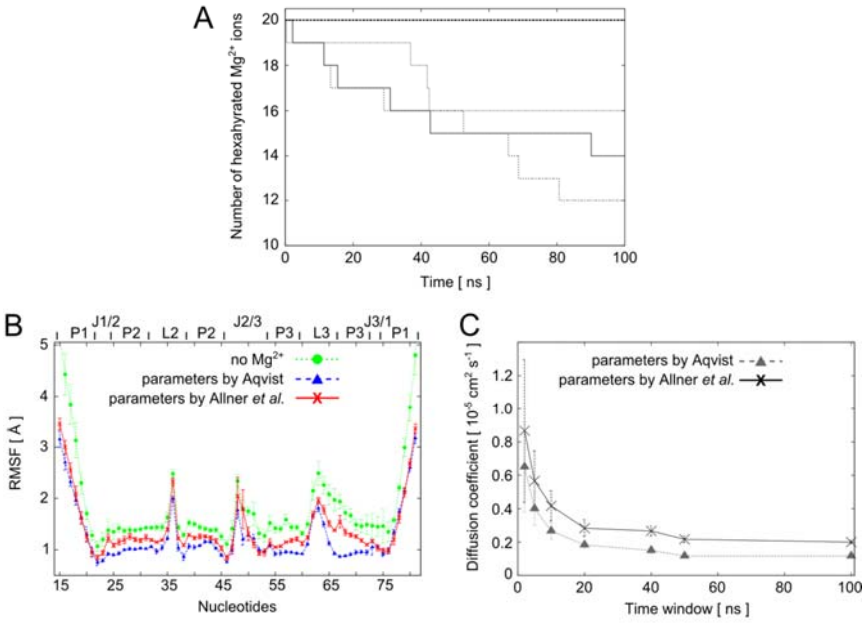

Figure 4 color

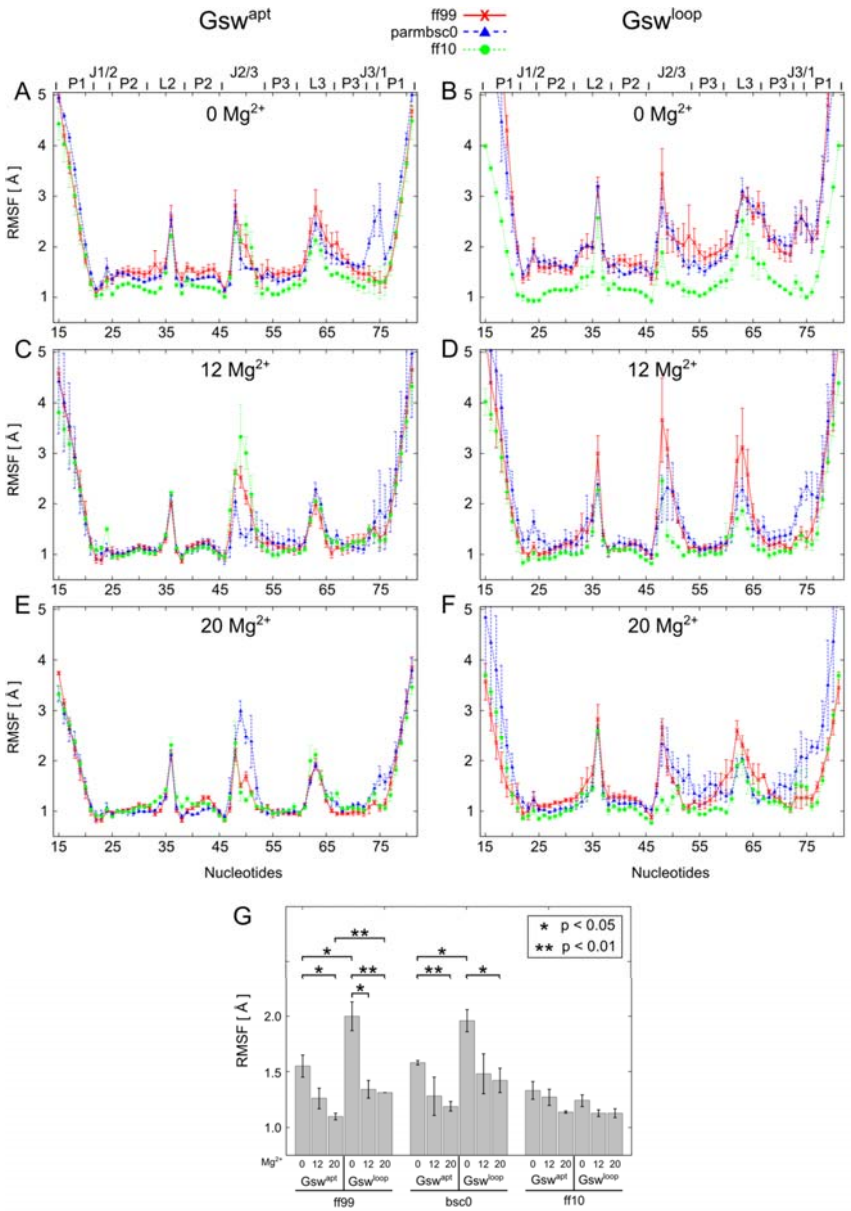

Figure 5 color