Protonation States of the Key Active Site Residues and Structural Dynamics of the *glmS* Riboswitch As Revealed by Molecular Dynamics

Pavel Banáš,[†,‡] Nils G. Walter,[§] Jiří Šponer,*[,†,‡] and Michal Otyepka*[,†,‡]

Department of Physical Chemistry, Faculty of Science, Palacky University, 17. Listopadu 12, 771 46 Olomouc, Czech Republic, Institute of Biophysics, Academy of Sciences of the Czech Republic, Kralovopolska 135, 612 65 Brno, Czech Republic, and Single Molecule Analysis Group, Department of Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109-1055

Received: November 18, 2009; Revised Manuscript Received: March 29, 2010

The *glmS* catalytic riboswitch is part of the 5′-untranslated region of mRNAs encoding glucosamine-6-phosphate (GlcN6P) synthetase (*glmS*) in numerous Gram-positive bacteria. Binding of the cofactor GlcN6P induces site-specific self-cleavage of the RNA. However, the detailed reaction mechanism as well as the protonation state of the *glmS* reactive form still remains elusive. To probe the dominant protonation states of key active site residues, we carried out explicit solvent molecular dynamic simulations involving various protonation states of three crucial active site moieties observed in the available crystal structures: (i) guanine G40 (following the *Thermoanaerobacter tengcongensis* numbering), (ii) the GlcN6P amino/ammonium group, and (iii) the GlcN6P phosphate moiety. We found that a deprotonated G40$^-$ seems incompatible with the observed *glmS* active site architecture. Our data suggest that the canonical form of G40 plays a structural role by stabilizing an in-line attack conformation of the cleavage site A-1(2′-OH) nucleophile, rather than a more direct chemical role. In addition, we observe weakened cofactor binding upon protonation of the GlcN6P phosphate moiety, which explains the experimentally observed increase in K_m with decreasing pH. Finally, we discuss a possible role of cofactor binding and its interaction with the G65 and G1 purines in structural stabilization of the A-1(2′-OH) in-line attack conformation. On the basis of the identified dominant protonation state of the reaction precursor, we propose a hypothesis of the self-cleavage mechanism in which A-1(2′-OH) is activated as a nucleophile by the G1(*pro*-R_p) nonbridging oxygen of the scissile phosphate, whereas the ammonium group of GlcN6P acts as the general acid protonating the G1(O5′) leaving group.

Introduction

Riboswitches are RNA motifs embedded in messenger RNAs (mRNAs) that regulate gene expression in response to binding of a specific small molecule ligand.[1−6] The *glmS* catalytic riboswitch (or ribozyme) is part of the 5′-untranslated region (5′-UTR) of mRNAs encoding glucosamine-6-phosphate synthetase (*glmS*), which catalyzes the conversion of glutamine and fructose-6-phosphate to glucosamine-6-phosphate (GlcN6P) and glutamate in numerous Gram-positive bacteria (Figure 1).[7−9] Riboswitches typically modulate gene expression on the transcriptional or translational level by undergoing structural rearrangements upon ligand binding that involve a transcription (anti-) terminator or ribosome binding site, respectively.[10−12] However, GlcN6P binding to the *glmS* riboswitch does not lead to any detectable structural rearrangements.[8,13−15] Instead, site-specific self-cleavage of the *glmS* riboswitch is activated directly by GlcN6P binding.[16] Self-cleavage consigns the RNA to the bacterial degradation pathway, which relies on the action of RNase J1 and ultimately results in down-regulation of *glmS* expression and GlcN6P production in a negative feedback loop.[17] The mechanism of the self-cleavage involves the nucleophilic attack of the A-1(2′-OH) hydroxyl group on its neighboring scissile phosphate of G1, thereby generating the 2′,3′-cyclic and 5′-OH termini of the reaction products. The same general mechanism is found in all ribozymes classified as "small" with typically less than 100 nucleotides in the catalytic core, yet the details of how the reaction participants are activated differ substantially.[18−23]

The *glmS* catalytic riboswitch is unique in that it completely depends on an external ligand. Mechanistic studies proved that GlcN6P is absolutely required to activate the *glmS* riboswitch, accelerating the cleavage rate by more than 10^5-fold over background hydrolysis.[8,16,24] The related compound glucose-6-phosphate, which contains a hydroxyl group in place of the 2-amine of GlcN6P (Scheme 1), is a competitive inhibitor of the self-cleavage reaction.[16] Weak activity was observed for glucosamine, serinol, L-serine, tris(hydroxymethyl)aminomethane, and ethanolamine, suggesting that a primary amine and hydroxyl group in a vicinal position is the structural feature required for all effectors activating *glmS* riboswitch.[16] The cleavage rate depends on the pK_a of the amino group. Thus, the GlcN6P ligand acts as a cofactor, and its 2-amino group is directly involved in catalysis, whereas the vicinal hydroxyl group is presumably required for proper effector positioning.[16]

Crystal structures of the *glmS* riboswitch from *Thermoanaerobacter tengcongensis*[25−27] and *Bacillus anthracis*[28] in various functional states show the RNA fold with a rigid core formed by a pseudoknot motif among the P2, P2.1, and P2.2 stems that is nested within another pseudoknot (Figure 1). The global structure does not rearrange upon cofactor or inhibitor binding. In addition, the riboswitch conformation is identical over a large pH range (5.5−8.5).[27] The pH-rate profiles show

* Corresponding authors. (M.O.) Fax: +420 585634761. E-mail: michal.otyepka@upol.cz. (J.S.) Phone: +420 541517133. E-mail: sponer@ncbr.chemi.muni.cz.
† Palacky University.
‡ Academy of Sciences of the Czech Republic.
§ University of Michigan.

10.1021/jp9109699
Published on Web 06/10/2010

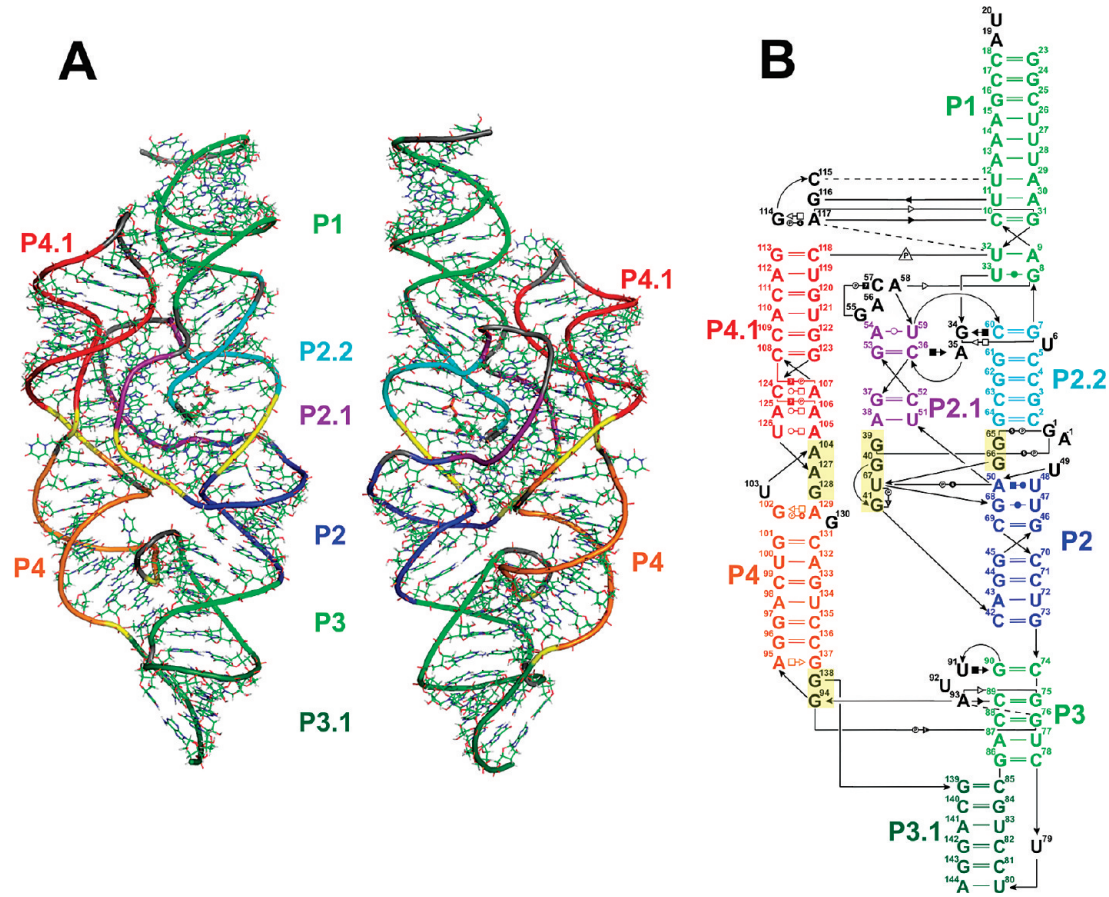
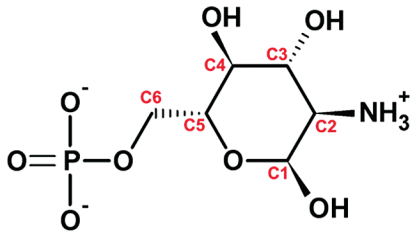


Figure 1. Structure of the *glmS* riboswitch from *T. tengcongensis*. (A) Rear and front views of the three-dimensional structure of the *glmS* riboswitch. Double-helical stems are shown in different colors, base stacking modules are in yellow, and unstructured parts and GNRA tetraloop are in gray. (B) The sequence and secondary structure of the *glmS* riboswitch. The colors of structural elements match those in panel A. Base pairs are annotated using standard classification.[71,72]

SCHEME 1: Structure and Atom Numbering of the Cofactor GlcN6P



that the cleavage rate increases with pH,[16] wherein the low activity at low pH originates mainly from a K_m (cofactor binding affinity) rather than from a k_{cat} (catalytic) effect.[28] This observation can also explain why Ferré-D'Amaré and co-workers were unable to crystallographically observe the complex of GlcN6P with the *glmS* riboswitch at pH 5.5.[26] These authors were, however, able to demonstrate that the P2.2 stem (Figure 1) plays an important role in cofactor binding, with G1(N1) significantly contributing by interacting with the phosphate moiety of GlcN6P.[27] The substitution of G1 with several purine analogs resulted in a switched cofactor specificity between GlcN6P and glucosamine at neutral pH, suggesting that G1 helps position the cofactor.[27]

G57C and G57A (G65C and G65A according to the *T. tengcongensis* numbering, Figure 1B) mutations in the consensus-type *glmS* riboswitch resulted in minimal and no catalytic activity, respectively.[29] Strobel and co-workers found that G57 (G65 in *T. tengcongensis*) directly interacts with the A-1 nucleobase and explained the abolished activity of the mutants by a loss of the proper in-line attack conformation of the A-1 nucleophile (Figure 2).[28] These authors also identified another substantial contact, G1(2′-OH)···G30(N7) (G1(2′-OH)···G37(N7) in *T. tengcongensis*), that stabilizes the reactive in-line attack conformation of A-1(2′-OH) relative to the scissile phosphate,[28] consistent with the observed loss of activity upon 7-deazaguanine substitution of this guanosine.[30] The active-site G40 (G33 in *B. anthracis*), which is in hydrogen bond distance to the A-1(2′-OH) nucleophile, was also identified as a nucleobase essential for riboswitch activity (Figure 2).[25] A G40A mutation reduces the cleavage rate constant ~10^5-fold, however, the crystal structure of the G40A mutant does not reveal any significant distortion of the active site (AS) compared to that of the wild-type;[25] the only difference found is an increased distance between the A-1(2′-OH) nucleophile and the N1 nitrogen of A40 compared to A-1(2′-OH)···G40(N1) distance, but the in-line attack conformation of A-1(2′-OH) remains intact.[25]

Two mechanisms were suggested for *glmS* riboswitch self-cleavage (Figure 3). The conserved G40 guanine (G33 in *B. anthracis*) was suggested to be deprotonated in the reactive state so that $G40^-$ acts as the general base accepting the proton from the A-1(2′-OH) nucleophile. In this mechanism, the ammonium form of GlcN6P is bound to the AS to act as the general acid (Figure 3A).[20,26] Ferré-D'Amaré and co-workers proposed an alternative mechanism, in which the amino form of GlcN6P is bound to the AS to act as the general base accepting the proton of the A-1(2′-OH) nucleophile via two tightly bound water

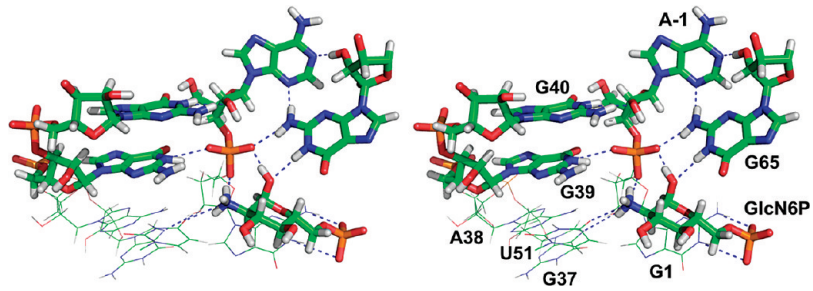


Figure 2. Stereo view of key nucleotides in the *glmS* riboswitch active site taken from the first snapshot of the G40/GlcN^{+}6P^{2-} simulation, showing a stabilized reactive in-line attack conformation of the nucleophile A-1(O2′) and proper binding of GlcN6P by a specific hydrogen bond network (*T. tengcongensis* numbering). See supplemental Figure S10 for enlarged but nonstereo version of this figure.

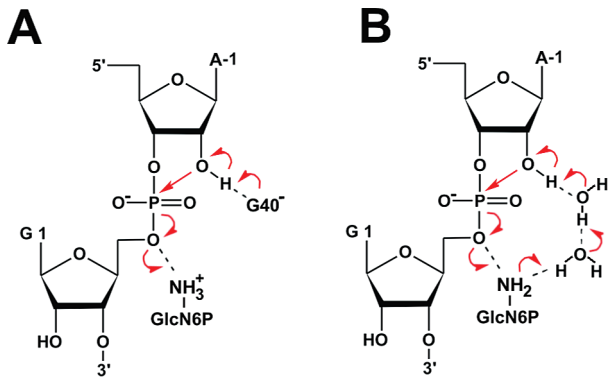


Figure 3. Proposed mechanisms for *glmS* riboswitch self-cleavage. (A) The conserved G40 is deprotonated and acts as the general base and GlcN6P acts as the general acid. (B) Glucosamine-6-phosphate acts both as the general base, accepting a proton from the A-1(O2′) nucleophile via two tightly bound waters, and as the general acid, transferring this proton to the leaving G1(O5′) oxygen. Red arrows denote electron flow during the reaction.

molecules. Subsequently, the now-protonated ammonium form of GlcN6P acts as the general acid by transferring its proton to the G1(O5′) leaving group (Figure 3B).[25]

The X-ray crystallography provides an inherently static and averaged view, and the crystal structure is not able to give direct information about protonation states (e.g., of key AS residues). However, the crystal structure can be effectively complemented by molecular dynamics (MD), providing structural and time-resolved information. MD with explicit inclusion of water molecules and ions is a computational tool that can be used to obtain direct atomic-resolution insights into the structural dynamics of an RNA. The applicability of the MD simulation technique is limited by the approximations of the underlying classical force field and the accessible time scale of simulations. When used wisely, however, MD can provide very valuable structural information.[31−37] Among the most appropriate applications for MD simulations of RNA are the analysis of the basic structural dynamics and structural substates associated with experimentally observed, inherently averaged, and static atomic structures;[38−41] the prediction of binding sites for monovalent ions;[42−45] the description of specific hydration sites, including the elucidation of the role of structured long-residency waters;[46−48] and the prediction of structural and dynamic consequences of base modifications and substitutions, including modifications of the protonation states.[43,49−51] Although classical MD simulations cannot be used to directly address catalysis, they can provide plausible mechanistic hypotheses and starting structures for direct investigation of catalysis by hybrid quantum chemical/molecular mechanical methods.[18,52−54]

Here, we apply MD simulations to the *glmS* riboswitch from *T. tengcongensis* with the native cofactor GlcN6P and different protonation states of the essential residues in the AS. MD is suitable to suggest the protonation states of the critical nucleobases corresponding to the crystalline conditions. This was previously shown, for example, for protonated cytosines in a HDV ribozyme and a frameshifting pseudoknot.[50,55] The aim of our simulations is to suggest dominant protonation states of AS residues, to describe the cofactor binding, and to study the dynamic behavior of AS to get ideas about the plausible reaction state. In particular, our data indicate that G40 is not deprotonated and likely plays a structural role in stabilizing the in-line attack conformation of the cleavage site A-1(2′-OH). We propose that A-1(2′-OH) could be activated as the nucleophile by the G1(*pro*-R_p) nonbridging oxygen of the scissile phosphate, whereas the ammonium group of GlcN6P acts as the general acid that neutralizes the leaving group 5′-oxygen (to avoid any confusion, we will further use terms *pro*-S_p for O1P and *pro*-R_p for O2P nonbridging oxygens according to IUPAC terminology). The presented results should be considered within the context of common limitations (mentioned above) of the contemporary simulation methods.

Methods

Preparation of Starting Structures. The starting geometries were based on the currently available crystal structures. The native structure of the *glmS* riboswitch precursor from *T. tengcongensis* resolved at 2.9 Å (PDB ID 2HO7) was used as a template.[26] This structure was obtained with the competitive inhibitor glucose-6-phosphate (Glc6P) bound in the AS. The structure of the *glmS* riboswitch precursor from *B. anthracis* with native cofactor GlcN6P bound in the AS and inhibited by 2′-*O*-methyl substitution on A-1 residue (PDB ID 2NZ4) shows that the competitive inhibitor Glc6P and native cofactor have the same binding pattern.[28] Thus, the 2NZ4 and 2HO7 structures were superimposed and modified to derive the native *glmS* structure with the bound native cofactor. The coordinates of the native cofactor as well as two Mg^{2+} ions bound to the phosphate of the cofactor with their first solvation shells were pasted from the 2NZ4 structure into the 2HO7 structure. A structural Mg^{2+} ion with three inner-shell contacts to the phosphates of C2, G36, and G37 in the 2HO7 structure was retained in the starting MD structure as the third divalent ion. The classical empirical nonpolarizable force fields describe divalents inaccurately, and divalents sample poorly in simulations.[31,37,42,55,56] The three retained Mg^{2+} ions in the *glmS* structure represent specific binding patterns supported by unequivocal experimental data;[26,28] however, other Mg^{2+} ions seen in the X-ray structure do not appear to be of any functional importance, and thus, to avoid force field artifacts arising from the inaccurate description of divalents, they were not considered in the simulations. Finally, the system was neutralized by sodium ions, as described below.

TABLE 1: Overview of the MD Simulations Performed Here[a]

simulation name	cofactor	G40	GlcN6P amino group	GlcN6P phosphate	simulation length (ns)
$G40^-/GlcN^+6P^{2-}$	yes	−1	+1	−2	20
$G40^-/GlcN^+6P^-$	yes	−1	+1	−1	20
$G40/GlcN^06P^{2-}$	yes	0	0	−2	20
$G40/GlcN^+6P^{2-}$	yes	0	+1	−2	50
$G40/GlcN^+6P^-$ (G1)[b]	yes	0	+1	−1	20
$G40/GlcN^+6P^-$ (P2.2)[b]	yes	0	+1	−1	20
$G40/GlcN^+6P^-$ (bulk)[b]	yes	0	+1	−1	20
G40/free	no	0			50

[a] The presence of ligand in the active site and the charge of the acid−base groups differing in the protonation state are indicated. [b] The labels "G1", "P2.2", or "bulk" in parentheses indicate initial orientation of the hydroxyl of the $GlcN^+6P^-$ phosphate toward G1, P2.2 stem, and bulk solvent, respectively.

Seven starting structures differing in protonation of the important acid−base groups in the AS were prepared on the basis of this structure. These include systems with both the amino and ammonium form in combination with both the singly protonated monocharged phosphate and deprotonated double-charged phosphate of GlcN6P. The abbreviation $GlcN^06P^{2-}$ denotes the amino form of GlcN6P with deprotonated double-charged phosphate, and $GlcN^+6P^{2-}$ and $GlcN^+6P^-$ stand for the ammonium form of GlcN6P with deprotonated double-charged and singly protonated monocharged phosphate, respectively. The abbreviation GlcN6P generally represents the glucosamine-6-phosphate without specification of its protonation states. Two simulations were carried out with deprotonated guanine $G40^-$ in the AS, and the remaining five were executed with a canonical G40. Thus, the following simulations were prepared (Table 1): $G40^-/GlcN^+6P^{2-}$, $G40^-/GlcN^+6P^-$, $G40/GlcN^06P^{2-}$, $G40/GlcN^+6P^{2-}$, and three independent simulations of $G40/GlcN^+6P^-$ differing in the starting orientation of the hydroxyl group at phosphate moiety of the GlcN6P cofactor. The simulation in which the hydroxyl group was initially oriented toward the G1 nucleobase is labeled as $G40/GlcN^+6P^-$(G1). $G40/GlcN^+6P^-$(P2.2) represents the simulation with the hydroxyl group oriented toward the P2.2 stem, whereas the hydroxyl group of simulation $G40/GlcN^+6P^-$(bulk) was initially exposed to the bulk solvent (Supporting Information, Figure S3). In addition, one reference structure was prepared with a ligand-free AS and a canonical G40 (G40/free). The structural triple inner-shell bound Mg^{2+} ion was retained, and the other two Mg^{2+} ions neighboring the cofactor were deleted.

Molecular Dynamics Simulations. All structures were neutralized with Na^+ counterions (Na^+ radius 1.868 Å and well depth 0.0028 kcal/mol) that were iteratively placed into the minima of the electrostatic potential calculated on a grid with spacing 1 Å using the program Leap of AMBER 9.0.[57] The structures with a ligand in the AS contained three Mg^{2+} ions and 136−138 Na^+ ions, depending on the charge of the AS residues. The reference structure contained one structural Mg^{2+} ion and 140 Na^+ ions. All structures were immersed in a rectangular box with at least a 10-Å-thick layer of TIP3P water molecules around the solute. The size of the boxes was ∼130 × 80 × 70 $Å^3$. The complete structures contained ∼55 000 atoms, including ∼6500 water molecules. The overall concentration of monovalent ions was ∼0.33 mol/L, which is entirely sufficient to provide stable RNA simulation trajectories with realistic local counterion accumulation around the solute molecule.

The whole RNA−solvent system was minimized prior to the simulations as follows. Minimization of the riboswitch hydrogen atoms was followed by minimization of counterions and water molecules. Subsequently, the riboswitch was constrained, and solvent molecules with ions were allowed to move during a 10 ps long MD run. The nucleobases were allowed to relax in several minimization runs with decreasing force constants applied to the backbone phosphate atoms. After full relaxation, each system was slowly heated to 298.15 K over 100 ps using 2-fs time steps and the NpT conditions. The simulations were evolved under periodic boundary conditions in the NpT ensemble (298.15 K, 1 atm) with 2-fs time steps.

The particle-mesh Ewald method was used to calculate electrostatic interactions, and a 9.0-Å cutoff was applied for Lennard-Jones interactions. The SHAKE algorithm was applied to all bonds containing hydrogen atoms. The PMEMD module of AMBER 9.0[57] with the Cornell et al. force field parm99[58,59] was used for all simulations. The length of each simulation was 20 ns, except for the $G40/GlcN^+6P^{2-}$ and G40/free simulations, which were expanded to 50 ns. The parameters of all nonstandard residues were determined by the RESP procedure of Cornell et al.[60] The ab initio calculations required for the parametrization of GlcN6P in various protonation states and deprotonated guanine were carried out using Gaussian03 (see the Supporting Information for details and parameters).[61,62]

Results

We carried out MD simulations of the *glmS* riboswitch to characterize cofactor binding and to study the dynamic behavior of the *glmS* riboswitch as a whole. Nonetheless, the main aim was to elucidate the dominant protonation states of the key AS residues and to obtain insights about plausible reaction mechanisms. Three different groups in the AS were identified as of an uncertain protonation state: (i) the N1 nitrogen of guanine G40 that was suggested to be deprotonated in a precursor state to act as the general base,[26,28] (ii) the amino group of GlcN6P that was suggested to act as either a general base[26] or general acid[28] during cleavage, and (iii) the phosphate moiety of GlcN6P (see Methods and Table 1 for details).

Basic Structural Dynamics Shows Extremely Rigid Pseudoknot Core of *glmS* Riboswitch. The structural dynamics and flexibility of the *glmS* riboswitch were monitored as B-factors of the *glmS* nucleotides. The pseudoknot core formed by the P2, P2.1, and P2.2 stems and including the AS was extremely rigid in all simulations. This is in agreement with the experimentally observed low B-factors in the crystal structures.[26,28] By contrast, the P1, P3, P3.1, and P4 stems represent more flexible regions of the RNA fold with correspondingly high B-factors (Figure 4). The rigidity of the pseudoknot core likely originates from a structural stabilization of the core arrangement by the coaxial P4 and P4.1 stems. They stabilize the pseudoknot fold by two interactions: (i) a ribose zipper motif[63] formed between the GNRA tetraloop[64] closing P4.1 and the P1 stem, which contains a type I A-minor interaction[65] between A117 and the C10═G31 base pair, and (ii) the oblique interaction of the G128|A127|A104|A105|A106 purine stack with the minor groove of P2.1 (Figure 5).[26] This salient tertiary interaction is entirely stable and very rigid in all simulations and might represent the key stabilizing interaction of the P2.1 stem fold. The flexibility of the P1 stem is nonuniform. The ribose zipper motif between

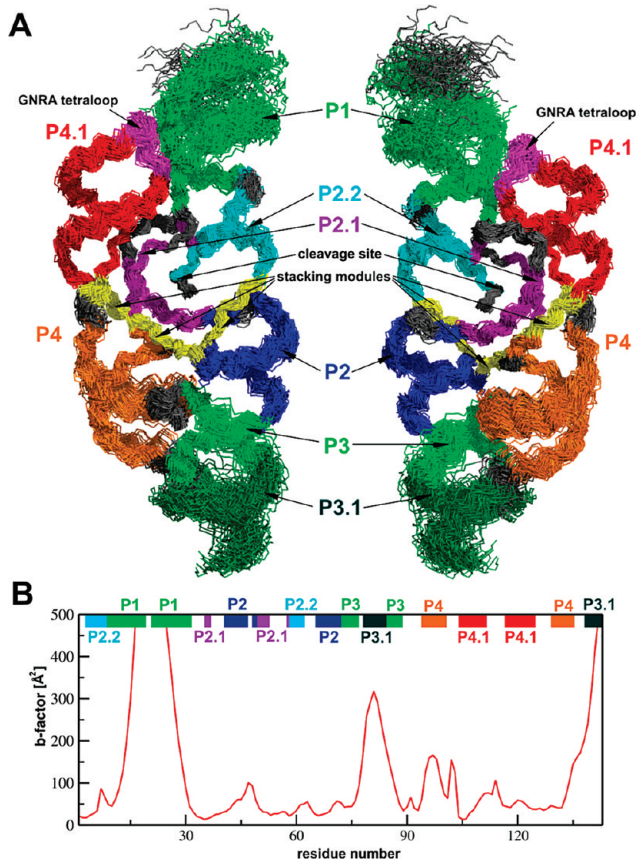


Figure 4. (A) The superimposed snapshots of G40/GlcN$^+$6P^{2-} simulation taken at each nanosecond demonstrate flexibility of P1, P3, P3.1 and P4 stems and rigidity of the pseudoknot core. (B) The thermal B-factors of the backbone atoms of each residue calculated from MD simulation (for comparison with X-ray B-factors, see the Supporting Information). The coloring of the stripes at the top of the plot matches the colors of the stems (shown above).

the GNRA tetraloop and P1 stem buttresses the P1 segment adjacent to the pseudoknot core up to the U12-A29 base pair, making it as rigid as the pseudoknot core itself. The remainder of the P1 stem is significantly more flexible (Figure 4). The pseudoknot core is also stabilized by the structural Mg^{2+} ion. As noted above, inclusion of divalent ions into simulations is hampered by considerable limitations.[31,56,66] The cation dynamics is explained in detail in the Supporting Information).

Four stacking modules were found in the *glmS* riboswitch structure (a term stacking module will be further used for stacks within a single-stranded region): the above-mentioned G128|A127|A104|A105|A106 stacking module between the P4 and P4.1 stems, G94|G138 stacking module connecting the P3 and P4 stems, and two G65|G66 and G39|G40|U67|G41 stacking modules occurring in the AS. These stacking modules form a base-zipper motif or partial base-zipper motifs associated with a crossover of single strands.[67−70] We found that these stacking modules are involved in cofactor binding, AS preorganization and structural stabilization of the overall fold (see the Supporting Information for details). Furthermore, the stacking module regions are often accompanied by non-Watson−Crick (non-WC) and base−phosphate (BPh) interactions (see Figure 1, Supporting Information for analysis of non-WC and BPh interactions, and refs 71 and 72 for their classification).

Only Certain Protonation States of Specific Active Site Residues Are Consistent with Experimental Active Site Architecture. An initial insight into the structural integrity of the AS with different protonation states was assessed by monitoring the mass-weighted root-mean-square deviations (rmsd) of three AS segments: (i) the sugar−phosphate backbone between the A-1(C4′) and G1(C4′) atoms encompassing the scissile phosphate, (ii) the preceding segment with the G40 base, and (iii) the preceding segment including the cofactor molecule with the exception of its phosphomethyl group (Figure 6). The geometry of each segment was fit with its reference crystal structure conformation, and the rmsd value was calculated. Subsequently, all the structural deviations from the X-ray structure were analyzed in detail by monitoring of key hydrogen bond interactions and sugar−phosphate backbone torsions in the AS.

The AS conformation remained close to the crystal structure throughout the entire G40/GlcN$^+$6P^{2-} and G40/GlcN$^+$6P$^-$(bulk) simulations. Backbone atoms neighboring the scissile phosphate and G40 base fluctuated near the crystal structure geometry (Figure 6, green and red lines, respectively), and the cofactor was bound tightly (Figure 6, black line).

A rearrangement of the AS took place in both the G40/GlcN$^+$6P$^-$(G1) and G40/GlcN$^+$6P$^-$(P2.2) simulations (Figures 6 and 7) representing the same protonation state as does the G40/GlcN$^+$6P$^-$(bulk) simulation, but with a different initial orientation of the hydroxyl group of GlcN6P phosphate moiety (Methods and Supporting Information Figure S3). The structural changes in the AS in the G40/GlcN$^+$6P$^-$(P2.2) simulation were initiated by an α/γ flip of the G41 nucleobase from gauche(−)/gauche(+) to trans/trans conformation, rather than by the protonation state of the cofactor phosphate moiety (Supporting Information Figure S2). This α/γ flip caused a disruption of the G40(2′-OH)···U67(O2′) hydrogen bond and subsequent rearrangement of the G40 base. The G40(2′-OH)···U67(O2′) hydrogen bond was not reestablished during the remainder of the simulation, even after the α/γ torsions of G41 flipped back (see the Supporting Information).

It cannot be ruled out that the α/γ flip may be related to the use of the parm99 force field. This force field has been recently rendered outdated for DNA simulations due to the accumulation of pathological γ-trans substates with concomitant B-DNA degradation in longer simulations and has been replaced by the parmbsc0 force field.[73] However, the α/γ flip issue of parm99 is assumed to be irrelevant for RNA simulations in which the α/γ t/t substates do occur in X-ray structures and do not accumulate during MD simulations in a manner that would destabilize the RNA structure.[42,74,75] Nevertheless, we still cannot rule out that the parm99 force field modestly exaggerates the propensity to adopt the γ-trans substate for this particular nucleotide. Note that even an application of the parmbsc0 force field would not necessarily provide an ultimate answer in this case, since parmbsc0 may instead overstabilize the canonical backbone conformation and thus eliminate correct sampling of γ-trans substates.[76]

The disruption of the AS architecture in the G40/GlcN$^+$6P$^-$(G1) simulation was clearly initiated by a rotation of the GlcN6P phosphate moiety. The phosphate's hydroxyl group originally pointing toward the N1 and N2 nitrogen atoms of the G1 base was repelled early after 0.38 ns, which resulted in a rupture of the stabilizing contact between G65 and A-1 (Figure 2) and subsequent AS rearrangement (see the Supporting Information). This structural change is evidently caused by an unfavorable contact between the GlcN6P phosphate's hydroxyl group and the G1 in the starting structure.

The rmsd analysis indicated large AS rearrangements in both simulations with deprotonated guanine (G40$^-$/GlcN$^+$6P^{2-} and G40$^-$/GlcN$^+$6P$^-$), primarily due to changes in the relative

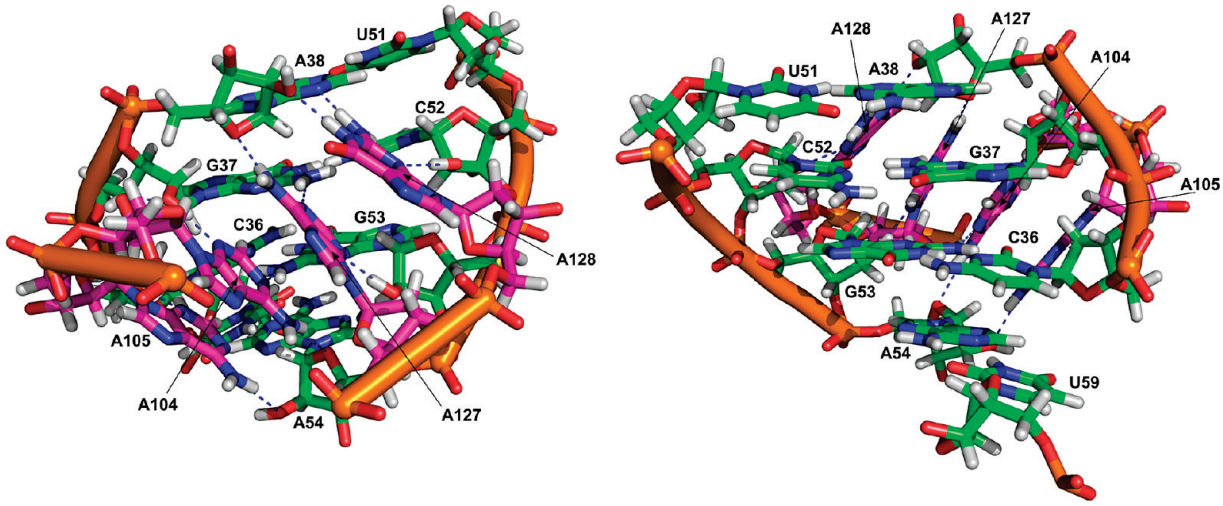


Figure 5. The front and rear view of the oblique interaction of the G128|A127|A104|A105|A106 purine stack with the minor groove of the P2.1 stem.

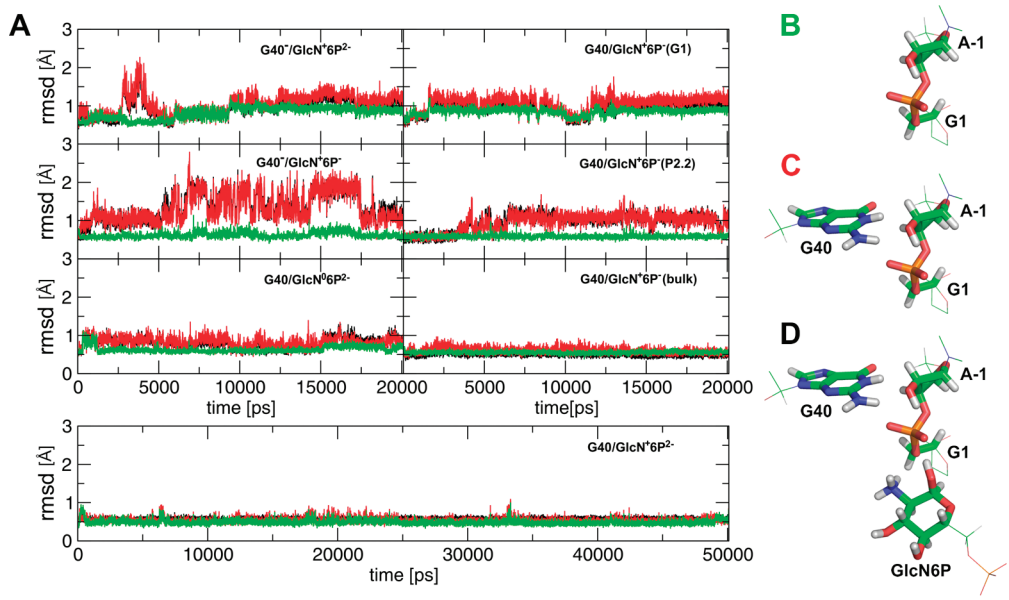


Figure 6. The rmsd profiles of the selected active site region. (A) Time dependence of the rmsd from the initial crystal structure calculated for three different regions of the *glmS* riboswitch active site as depicted (B) for the green line, (C) for the red line, and (D) for the black line. The specific atoms used for the rmsd calculation are shown in stick representation. The lowest rmsd of ~0.5 Å is characteristic of structures with no geometrical changes as compared to the starting X-ray geometries. rmsd values above ~0.5 Å indicate changes from the starting crystal structure (for a detailed description, see the Supporting Information).

position of G40$^-$ and the backbone around the scissile phosphate, with some internal reconfiguration of the backbone (Figure 6). The loss of the hydrogen bond between A-1(2′-OH) and G40$^-$(N1), and subsequent expulsion of G40$^-$ from the AS was observed in both simulations (Figure 7D). The G41 residue (or any other residue in proximity to AS) does not undergo any α/γ flip, and thus, the shift of G40$^-$ away from the scissile phosphate is unambiguously caused by the effect of G40 deprotonation (see the Supporting Information). The G40$^-$ is not compatible with the experimental crystal structures.

A slight and almost insignificant increase in rmsd was observed in the G40/GlcN^{0}6P^{2-} simulation (Figure 6); however, the detailed structural analysis showed that the cofactor amino group was not able to establish a stable hydrogen bond with the U51(O4) carbonyl, in contrast to the simulations involving the protonated ammonium form of the cofactor, in which this hydrogen bond was stable over the entire simulation (Figure 7B). The weak binding between the GlcN^{0}6P^{2-}(N1) amino group and U51(O4) in the G40/GlcN^{0}6P^{2-} simulation caused increased flexibility of the amino group with subsequent disruption of the G65(N2)···G1(*pro*-R$_p$) hydrogen bond. Simultaneously, U51 shifted out of the AS and influenced the conformation of the functionally important G40 base. Ultimately, also, the interaction between G39(N1, N2) and the G1(*pro*-S$_p$) nonbridging oxygen (Figure 2) was disrupted, the backbone around the scissile phosphate was rearranged, and the AS remained distorted over the remainder of the simulation (see the Supporting Information).

Cofactor Binding Is Weakened by Protonation of GlcN6P Phosphate. The binding of the GlcN6P cofactor in the AS was stable in all simulations except for G40$^-$/GlcN$^+$6P^{2-}, in which a shift of the cofactor toward U51 and G39 and a disruption of the GlcN$^+$6P^{2-}(N2)···G1(O5′) hydrogen bond were observed at 9 ns (Figure 7B), accompanying the overall structural destabilization of the AS in the presence of the deprotonated G40$^-$. In all other simulations carrying the ammonium form of

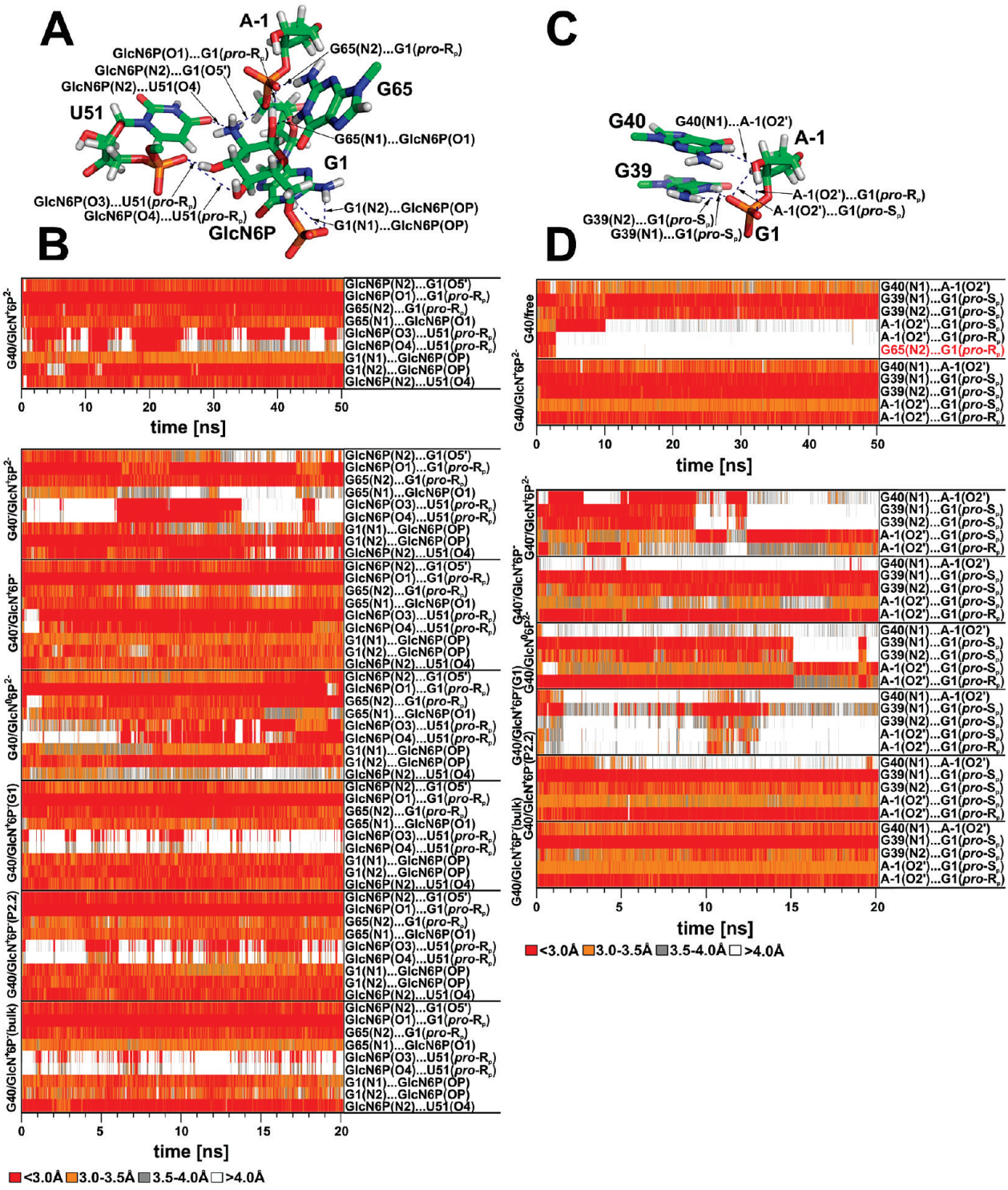


Figure 7. Time evolution of key hydrogen bonds in our MD simulations of the *glmS* riboswitch, documenting cofactor binding and stability of the active site architecture. (A) Glucosamine-6-phosphate bound in the glmS active site with its hydrogen bonding network; (B) time evolution of the hydrogen bonds of panel A; (C) part of the active site, including the scissile phosphate and two neighboring guanines with their network of hydrogen bonds; (D) time evolution of the hydrogen bonds of panel C (the G65(N2)···G1(pro-R_p) distance marked in red is shown in panel A).

GlcN6P, the cofactor remained in the conformation corresponding to its crystallographic binding pattern with a stable hydrogen bond between the ammonium group of the cofactor and the G1(O′5) backbone oxygen. This hydrogen bond is assumed to be important in self-cleavage of the *glmS* riboswitch.[16] In addition, the GlcN6P ammonium group was stabilized in its position by a hydrogen bond to the U51(O4) carbonyl (Figure 7B), whereas the third hydrogen of the ammonium group was bound to a water molecule.

The stabilizing interaction between U51(O4) and the cofactor was absent in the G40/Glc$N^0$6P^{2-} simulation (Figure 7B), where the cofactor bore the unprotonated amino group incapable of such a hydrogen bond, causing a shift of the amino group toward the nonbridging oxygens of the scissile phosphate and subsequent AS rearrangement.

A firm and stable hydrogen bond was observed between the GlcN6P C1−OH hydroxyl group and the G1(*pro*-R_p) oxygen in all simulations. This interaction was supported by two

additional G65(N1)···GlcN6P(C1−OH) and G63(N2)···G1(*pro*-R_p) hydrogen bonds (Figure 7B) which, however, were completely lost in simulation $G40^-/GlcN^+6P^{2-}$ and partially in simulations $G40^-/GlcN^+6P^-$ and $G40/GlcN^06P^{2-}$ accompanied by a rearrangement of the AS. The local interaction of G65, the scissile phosphate, and the C1-OH hydroxyl group of GlcN6P represents a compact binding pattern, which stabilizes the contact between the cofactor ammonium group and G1(O5′).

A strong bifurcated hydrogen bond between G1(N1, N2) and the cofactor phosphate group was observed in all simulations with a deprotonated double-charged phosphate. The same interaction was observed in simulations with a singly protonated monocharged GlcN6P phosphate (one with deprotonated $G40^-$ and three with canonical G40). However, an early reorientation of the cofactor phosphate group was observed in the $G40^-/GlcN^+6P^-$ and $G40/GlcN^+6P^-$(G1) simulations that started with a phosphate hydroxyl group oriented toward the G1 base. This hydroxyl group was exposed to the bulk solvent during the first nanosecond of the respective simulations; that is, the hydroxyl group switched to the conformation corresponding to the starting structure of the $G40/GlcN^+6P^-$(bulk) simulation. $G40/GlcN^+6P^-$(P2.2) was the only simulation, in which the cofactor phosphate hydroxyl group was not oriented toward the bulk solvent but, rather, interacted directly with the G64 or G65 bases, whereas in other simulations, the GlcN6P phosphate interacted with guanines G64 and G65 via a solvated Mg^{2+} ion.

Partially occupied hydrogen bonds were observed between the cofactor hydroxyl groups C3−OH and C4−OH and the U51 (*pro*-R_p) nonbridging oxygen. These hydrogen bonds are fluctuating and do not appear to be essential for GlcN6P binding (Figure 7).

In summary, both a deprotonated $G40^-$ and an unprotonated amino form of cofactor ($GlcN^06P$ instead of $GlcN^+6P$) result in significant destabilization of cofactor binding. Further, protonation of the cofactor phosphate moiety ($GlcN6P^-$ instead of $GlcN6P^{2-}$) likely negatively affects the stability of the interaction between the cofactor phosphate and the G1 nucleobase. More specifically, we note that the $GlcN^+6P^-$ with the phosphate hydroxyl group oriented toward G1 represents an unstable binding motif. By contrast, the cofactor was firmly bound in the active site when the $GlcN^+6P^-$ phosphate hydroxyl group was oriented toward the bulk solvent. Thus, protonation of the GlcN6P phosphate moiety does not prevent cofactor binding for an appropriate geometry but restricts a conformational variability of the cofactor, which should be accompanied by an entropic penalty for cofactor binding.

Structural Dynamics of the *glmS* Riboswitch in the Absence of Cofactor Show That A-1(2′-OH) In-Line Attack Conformation Requires Cofactor Binding. Reference 50-ns-long simulation of the ligand-free *glmS* riboswitch (without GlcN6P) was started from the crystal structure geometry (i.e., the structure with bound GlcN6P) and revealed profound changes of the AS. This indicates that cofactor binding is needed for properly structuring the AS. The two-step rearrangement of the AS is described in the Supporting Information.

In contrast to the AS occupied by the cofactor, the stable ligand-free AS arrangement lacks the G65(N2)···G1(*pro*-R_p) hydrogen bond and the in-line attack conformation of A-1(2′−OH) (Supporting Information Figure S5 and Table 2). On the other hand, the base−phosphate interaction between G39(N1) and the G1(*pro*-S_p) nonbridging oxygen of the scissile phosphate as well as the G65(N2)···A-1(N3) and G65(2′-OH)···A-1(N1) hydrogen bonds were preserved (Supporting Information Figure S4). It seems that the position of the G1 base controls the conformation of the backbone between A-1 and G1. We suggest that cofactor binding can affect the position of G1 and, thus, is vital for a proper local conformation of the scissile phosphate. Furthermore, cofactor binding can help to establish the interaction between the scissile phosphate and G65 and thus stabilizes the A-1 nucleotide in a way that supports the A-1(2′-OH) in-line conformation.

TABLE 2: The Mean Values of A-1(2′-O)···G1(P)-G1(O5′) In-Line Attack Angle (IAA) in Degrees with Standard Deviations in All Presented Simulations

simulation	IAA	simulation	IAA
$G40^-/GlcN^+6P^{2-}$ [a,b]	155 ± 10	$G40/GlcN^+6P^-$ (G1) [a,c]	135 ± 10
$G40^-/GlcN^+6P^-$ [a]	155 ± 10	$G40/GlcN^+6P^-$ (P2.2) [a]	165 ± 5
$G40/GlcN^06P^{2-}$ [a,b]	160 ± 10	$G40/GlcN^+6P^-$ (bulk)	170 ± 5
$G40/GlcN^+6P^{2-}$	170 ± 5	G40/free [d]	130 ± 10

[a] Poor or missing interaction between G40 and A-1(2′-OH). [b] Poor or missing interaction between G39 and G1(*pro*-S_p). [c] Disruption of G65(N2)···A-1(N3) hydrogen bond. [d] Missing cofactor binding.

Discussion

We have carried out a set of explicit solvent MD simulations of the *glmS* riboswitch with the aim to elucidate its overall structural dynamics, the arrangement of its AS in the reactive state, and its major protonation state.

Because the X-ray crystallography is not able to identify proton positions, the MD simulation can be employed to assign a protonation state, which corresponds to the state reflected in the X-ray structure. Here, such identification is based on structural deviations of AS residues from the X-ray structure in MD simulation applying various protonation states of AS residues. It can be expected that the protonation state corresponding to that observed in the X-ray structure (here resolved at pHs ranging from 5.5 to 8.5) would exhibit the smallest structural deviation in MD. This represents an indirect way how to assign a protonation state that likely corresponds to the experimentally observed crystal structure.

A direct theoretical method to address the protonation state and estimate the pK_a of a given titrable group is a constant-pH molecular dynamics.[77−79] However, the constant-pH MD methods that are used and widely tested for proteins[77−80] typically utilize implicit solvent models, which reasonably limits their applicability in the case of nucleic acids. We used the constant-pH MD as implemented in AMBER[57,77] to estimate pK_a values of three discussed titrable groups of *glmS* riboswitch (see the Supporting Information). Unfortunately, we observed a rapid and massive degradation of the *glmS* riboswitch global fold during the first ∼10 ps of the thermalization phase in simulations using the implicit solvent (see the Supporting Information). Thus, we restrained the structure of the *glmS* riboswitch to its crystal-like geometry and estimated the pK_a using constant-pH MD. The sign of the observed pK_a shifts was in agreement with the conclusions presented in this work; however, the absolute value of pK_a shifts was severely overestimated (see Supporting Information). The overestimated pK_a shifts together with the rapid degradation of RNA structure in implicit solvent show that the implicit solvent methods (at least in current implementation) are not able to efficiently screen out the electrostatics of complex folded RNA molecules. It seems that the RNA molecule remains a challenge for MD simulations in combination with implicit solvent methods.

Overall Dynamics and Stability. Our simulations highlight a significant rigidity of the pseudoknot core of the *glmS*

riboswitch, formed by the P2, P2.1, and P2.2 stems, which is contrasted by the flexible outer parts, including the top of the P1 stem and the pseudoknot formed by stems P3 and P3.1. The core appears to be made rigid by interactions with coaxial P4 and P4.1 stems; namely, oblique interaction of the G128|A127|A104|A105|A106 stacking module with the minor groove of the P2.1 stem[26] and the interaction of the GNRA tetraloop at the tip of the P4.1 stem with the minor groove of the P1 stem that forms a ribose−zipper motif with a class I A-minor interaction. Our simulations further suggest that the stacking modules made by stacked single strands or base−zipper motifs play an important role in structural stabilization, AS preorganization, and cofactor binding. There are several base−phosphate interactions that are stable in MD simulations. These interactions are often connected with the noncanonical segments of the *glmS* riboswitch, especially the stacking modules, and likely play an important role in the structural stabilization of its noncanonical parts. All interactions above are very stable in MD simulations.

Protonation State and Proposed Functional Role of G40. Simulations of the *glmS* riboswitch were used to identify the likely dominant protonation state of three critical moieties in the AS. G40 was the first moiety of the uncertain protonation state, which was previously identified by biochemical data to play an essential role in *glmS* self-cleavage.[24,27] It was further suggested that the deprotonated G40$^-$ acts as the general base during self-cleavage;[25,26,28,81] however, expulsion of the G40$^-$ base from the AS and subsequent AS distortion were observed in both simulations with a deprotonated G40$^-$. The deprotonated G40$^-$ appears to be incompatible with the *glmS* AS architecture found in all available crystallographic studies. The same has been observed in simulations with the deprotonated form of G8, which was suggested to be involved in the hairpin ribozyme self-cleavage.[89] A structural rather than catalytic role of G8 in transition state stabilizing the self-cleavage reaction in the hairpin ribozyme was suggested by York et al. who proposed that the tautomeric form of the active site guanine G8 is not likely involved in the reaction chemistry of hairpin ribozyme.[52,82] We found that G40 structurally stabilizes the A-1(2′-OH) in-line attack conformation relative to the scissile phosphate, and thus, we suggest that G40 is important for structural and electrostatic stabilization of the reactive state rather than directly acting in the catalytic reaction as the general base.

Still, it is possible that the deprotonated form of G40$^-$ may be involved in catalysis; however, in such a case, it would likely correspond to a low-populated, transient, reactive state. In the 20+ ns scale simulations, the AS of systems with G40$^-$ is entirely unstable and rearranges swiftly to very different geometries. Assuming that the contact between A-1(O2′) and G40$^-$ was disrupted at the beginning of the simulation and was not reestablished for even one single 2-ps-long snapshot, a rough estimate would yield an occupation of the deprotonated G40$^-$ in the reactive state of less than 0.01%. The corresponding ΔG correction for the low-populated reactive state would be higher than 6 kcal/mol. In addition, the unperturbed pK_a of guanine of 9.2[83] suggests a free energy correction for formation of a guanine anion of 3 kcal/mol at pH 7. Thus, the reaction mechanism with G40$^-$ acting as the general base would have to be chemically favored in the transition state stabilization by at least 9 kcal/mol over other alternative pathways (with significantly populated reactive state). We cannot ultimately rule out that longer simulations may rearrange the AS into a suitable, more populated geometry involving G40$^-$ that would be separated from the presently sampled geometries by an energy barrier and thus not immediately accessible. However, we have no indication what type of geometry that would be and consider this possibility as less likely. In addition, it is common in the hydrolysis of the sugar−phosphate backbone that the departure of the O5′-alcoholate from the pentahedral intermediate is the rate-limiting step, rather than the 2′-OH nucleophilic attack.[84,85] Thus, it appears unlikely that the increased basicity of G40$^-$ affects the reaction rate sufficiently to counteract the 9 kcal/mol penalty in ΔG arising from a low-populated G40$^-$ reactive state and rare protonation state of G40$^-$.

Mechanistic and structural experiments have previously found that a G40A mutation abrogates most of the catalytic activity of the *glmS* riboswitch but does not extensively affect the arrangement of the AS in both *T. tengcongensis* and *B. anthracis*.[25,81] The only difference between G40A and wild-type AS conformation is a slight increase in the distance between the N1 of residue 40 and the A-1(O2′) nucleophile, whereas the in-line conformation remains intact.[25,81] One explanation is that G40 is deprotonated and acts as the general base.[25,26,28,81] Alternatively, donation of the hydrogen bond by canonical G40 to A-1(O2′) could be critical for the precise positioning of the nucleophile to a reactive in-line attack conformation or for electrostatic stabilization of the attacking A-1(2′-OH), whereas an in-line like position in G40A mutant is nonproductive. Preliminary MD simulations of the G40A mutant (data not shown) suggest that the mutation, consistent with the crystal structures, causes only slight structural changes in the AS. The A40 remains locked in the same position as G40 of the wild type, but A-1(O2′) does not establish any hydrogen bond with the A40(N1). Furthermore, the lack of a stabilizing G40(N1)···A-1(O2′) hydrogen bond allows tighter binding of A-1(O2′) to the G1(*pro*-R_p) nonbridging oxygen as compared with the wild type MD simulations.

On the basis of the above-mentioned arguments, we suggest that G40 is not directly involved in the chemical reaction but, rather, plays a role in electrostatic stabilization of the attacking A-1(2′-OH) nucleophile (i.e., in electrostatic transition state stabilization). The same role has also been proposed for the G8 involved in the hairpin ribozyme self-cleavage;[49,52,82] however, further clarification of the inhibition effect of G40A mutation is still required.

GlcN6P: Amino or Ammonium Group? The second chemical group in the AS with an uncertain protonation state is the amino/ammonium group of GlcN6P. The hydrogen bond between the cofactor ammonium group and U51(O4) was stable in all simulations with the ammonium form of GlcN6P; however, this hydrogen bond was missing in the simulation in which the cofactor contained the (uncharged) amino group, consistently with the reduced proton donor capability of the amino as compared with the ammonium group. Because the hydrogen bond contact between U43(O4) (U51(O4) in *T. tengcongensis*) and GlcN6P nitrogen was observed in crystal structures of *B. anthracis* with native GlcN6P cofactor,[28] the loss of this hydrogen bond detected in our MD simulations with the amino form of GlcN6P shows that the amino form of GlcN6P is not consistent with X-ray structures. Since the crystal structures of *B. anthracis* were obtained at pH 6.8[28] and the unperturbated pK_a of the GlcN6P equals 8.2,[29] our data do not support a significant pK_a shift of this amino/ammonium group in the active site. The simulations suggest that the ammonium form of the cofactor is preferentially bound to the AS at physiological pH and is bound more tightly as compared with the amino form. Thus, the ammonium form of the cofactor is capable of tightly binding G1(O5′) and is in a suitable position to act as the general acid of the reaction.

The Protonation of GlcN6P Phosphate Weakened Cofactor Binding. The third moiety of the uncertain protonation state is the phosphate of GlcN6P. From the presented simulations, it was not possible to clearly suggest whether the phosphate group of the cofactor bound to the AS is deprotonated (double-charged) or singly protonated (monocharged). Both protonation forms appear to remain bound to the AS, yet tighter binding of the cofactor phosphate moiety with G1 base was observed in the deprotonated double-charged, as compared with the singly protonated monocharged, form. In addition, the binding of a singly protonated phosphate may be disadvantageous because the phosphate hydroxyl group cannot bind to the G1 base and instead prefers the orientation to the bulk solvent. This conformational restriction is likely associated with an additional entropic penalty for binding the GlcN6P with the protonated phosphate moiety. The increase in K_m with decreasing pH[28] supports our observation that GlcN6P with a protonated phosphate binds more weakly as compared with the double-charged deprotonated phosphate of GlcN6P. A recent kinetic analysis shows that the pH dependence of cofactor binding is consistent with the pK_a of the GlcN6P phosphate moiety, suggesting that the protonation of the phosphate moiety of GlcN6P may inhibit cofactor binding.[81] This experimental finding, together with the MD data, lends support to the notion that GlcN6P with a fully deprotonated phosphate binds most efficiently to the *glmS* riboswitch.

Possible Role of G65. Biochemical analysis has suggested an essential role for the G65 base (or equivalent G57 in *B. anthracis*) in catalysis.[28,29] G65 was suggested to stabilize a proper conformation of the A-1 base and to neutralize the negative charge on the G1(*pro*-R_p) nonbridging oxygen of the scissile phosphate. We found that both these interactions are, indeed, required for stabilization of the AS in a reactive A-1(2′-OH) in-line attack conformation. In addition, we suggest that a further crucial role of G65 is to stabilize the hydrogen bond between the C1−OH group of GlcN6P and the G1(*pro*-R_p) oxygen that in turn helps maintain the hydrogen bond between the cofactor ammonium group and G1(O5′). The GlcN6P(C1−OH)···G1(*pro*-R_p) contact is the most rigid hydrogen bond in the AS of the *glmS* riboswitch. The G65 base stabilizes this hydrogen bond and forms the core of the GlcN6P binding motif. Our findings are consistent with the observation that ethanolamine is the minimal motif binding to the AS of the *glmS* riboswitch to activate self-cleavage.[16] In addition, on the basis of the simulation with ligand-free AS, we suggest that cofactor binding is crucial for establishing the interactions among the scissile phosphate, the cofactor, and G65 that structurally support the A-1(2′-OH) in-line attack conformation. In other words, we suggest that there are two roles of the GlcN6P cofactor: (i) the A-1(2′-OH) in-line attack conformation is induced by cofactor binding, and (ii) the ammonium group of the cofactor acts as the general acid in the self-cleavage reaction.

Hypothesis of the Reaction Mechanism Based on the Suggested Active Site Protonation State. Taken together, presented MD simulations suggest a plausible AS conformation of the reactive state, including the protonation states of key AS residues that are compatible with X-ray structures. On the basis of this reactive conformation, we can obtain insight into the possible mechanism of the self-cleavage reaction. Ribozymes are thought to employ four strategies to achieve catalysis:[86,87] (i) they can stabilize the in-line attack conformation of the nucleophile toward the scissile phosphate; (ii) they can activate the nucleophile by deprotonation of the cleavage site 2′-hydroxyl either prior to or simultaneously with the nucleophile attack; (iii) they can neutralize the increased electron density of the scissile phosphate during catalysis, making it more susceptible to nucleophilic attack; and (iv) they can help to protonate the leaving 5′-O^- alcoholate group. Here, we argue that canonical G40 plays the role in the structural and electrostatic stabilization of the transition state rather than acting as the general base in the deprotonated form ($G40^-$). We consequently suggest that the proton of the A-1(2′-OH) group could be transferred to the G1(*pro*-R_p) nonbridging oxygen of the scissile phosphate (Figure 8). It is worth noting that the nonperturbed pK_a of the nonbridging oxygen is equal to ∼1;[83] however, the pK_a of the nonbridging oxygen of the phosphorane intermediate is equal to ∼6.5[84] so that the basicity of the nonbridging oxygens of the scissile phosphate is increasing during the course of the advancing nucleophilic attack. Thus, the nonbridging oxygen is capable of acting as the general base during the reaction and accepting the proton from the 2′-OH nucleophile.

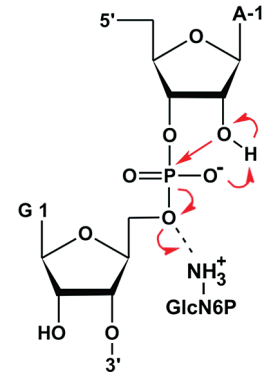


Figure 8. Proposed mechanism for *glmS* riboswitch self-cleavage based on presented MD simulations. Simultaneously with the nucleophilic attack, the G1(pro-R_p) nonbridging oxygen acts as the general base, accepting a proton from the A-1(O2′) nucleophile, and the GlcN6P acts as the general acid to donate its proton to the leaving oxygen G1(O5′). Red arrows denote electron flow during the reaction.

A similar reaction scenario has been proposed for the hairpin ribozyme.[52,88] Simultaneously with the transfer of the proton to the G1(*pro*-R_p) oxygen, the electron density located on the oxygen atom would be polarized toward the incoming hydrogen, making the scissile phosphate even more susceptible to nucleophilic attack. In addition to this effect, the 5BPh interaction of G39 with the scissile phosphate and strong hydrogen bond pattern between C1−OH hydroxyl of GlcN6P, G65, and the G1(*pro*-R_p) nonbridging oxygen further draws the electron density from the scissile phosphate. Finally, the ammonium group of the GlcN6P cofactor is perfectly positioned to act as a general acid, protonating the leaving G1(O5′) group (Figure 8).

Conclusions

MD simulations suggest that a deprotonated $G40^-$ is incompatible with the active site architecture observed in all *glmS* riboswitch crystal structures. We therefore propose that canonical G40 stabilizes the A-1(2′-OH) in-line conformation, plays the key role in electrostatic stabilization of transition state rather than directly participating in reaction chemistry, or both. We suggest a possibility that the A-1(2′-OH) nucleophile is activated and deprotonated by the G1(*pro*-R_p) nonbridging oxygen of the scissile phosphate.

The simulations reveal that the protonated ammonium form of the cofactor is bound in the AS more tightly and is more consistent with crystal structures than its uncharged amino form.

In addition, the ammonium group is in a suitable position to act as the general acid.

$GlcN^{+}6P^{-}$ with a singly protonated phosphate binds to the G1 nucleobase more weakly as compared with the double-charged, deprotonated phosphate of $GlcN^{+}6P^{2-}$, which is consistent with experimental data.[28,81]

We suggest that alongside a role in structural stabilization of the A-1(2′-OH) in-line attack conformation, G65 might play a crucial role in cofactor binding by stabilizing the hydrogen bond between the C1−OH hydroxyl of GlcN6P and the G1(*pro*-R_p) oxygen that, in turn, helps maintain the hydrogen bond between the cofactor ammonium group and G1(O5′). Thus, two roles of the GlcN6P cofactor in self-cleavage are suggested: (i) the A-1(2′-OH) in-line attack conformation is induced by the cofactor binding, and (ii) the ammonium group of the cofactor acts as the general acid in the reaction.

Acknowledgment. This study was supported by Grants LC512, LC06030, and MSM6198959216 from the Ministry of Education of the Czech Republic, Grants 203/09/1476 and 203/09/H046 from the Grant Agency of the Czech Republic, Grants IAA400040802 and 1QS500040581 from the Grant Agency of the Academy of Sciences of the Czech Republic, Grants AV0Z50040507 and AV0Z50040702 from the Academy of Sciences of the Czech Republic, and NIH Grant GM62357 (to N.G.W.). We thank S. R. Das for helpful comments and discussions.

Supporting Information Available: The content of the Supporting Information includes force field parameters of nonstandard residues, a detailed analysis of the structural dynamic of the *glmS* riboswitch active site, details of the structural dynamics of the *glmS* riboswitch without cofactor, analysis of constant-pH MD simulations using implicit solvent methods, and some other material. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Mandal, M.; Boese, B.; Barrick, J. E.; Winkler, W. C.; Breaker, R. R. *Cell* **2003**, *113*, 577.
(2) Tucker, B. J.; Breaker, R. R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 342.
(3) Winkler, W. C.; Breaker, R. R. *Annu. Rev. Microbiol.* **2005**, *59*, 487.
(4) Henkin, T. M. *Genes Dev.* **2008**, *22*, 3383.
(5) Coppins, R. L.; Hall, K. B.; Groisman, E. A. *Curr. Opin. Microbiol.* **2007**, *10*, 176.
(6) Winkler, W. C. *Curr. Opin. Chem. Biol.* **2005**, *9*, 594.
(7) Barrick, J. E.; Corbino, K. A.; Winkler, W. C.; Nahvi, A.; Mandal, M.; et al. *Proc. Natl. Acad. Sci.* **2004**, *101*, 6421.
(8) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. *Nature* **2004**, *428*, 281.
(9) Milewski, S. *Biochim. Biophys. Acta* **2002**, *1597*, 173.
(10) Batey, R. T.; Gilbert, S. D.; Montange, R. K. *Nature* **2004**, *432*, 411.
(11) Corbino, K. A.; Barrick, J. E.; Lim, J.; Welz, R.; Tucker, B. J., *Genome Biol.* **2005**, *6*, R70.
(12) Grundy, F. J.; Henkin, T. M. *Crit. Rev. Biochem.* **2006**, *41*, 329.
(13) Hampel, K. J.; Tinsley, M. M. *Biochemistry* **2006**, *45*, 7861.
(14) Fedor, M. J. *Annu. Rev. Biophys.* **2009**, *38*, 271.
(15) Tinsley, R. A.; Furchak, J. R. W.; Walter, N. G. *RNA* **2007**, *13*, 468.
(16) McCarthy, T. J.; Plog, M. A.; Floy, S. A.; Jansen, J. A.; Soukup, J. K.; et al. *Chem. Biol.* **2005**, *12*, 1221.
(17) Collins, J. A.; Irnov, I.; Baker, S.; Winkler, W. C. *Genes Dev.* **2007**, *21*, 3356.
(18) Banas, P.; Rulisek, L.; Hanosova, V.; Svozil, D.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2008**, *112*, 11177.
(19) Bevilacqua, P. C.; Yajima, R. *Curr. Opin. Chem. Biol.* **2006**, *10*, 455.
(20) Cochrane, J. C.; Strobel, S. A. *Acc. Chem. Res.* **2008**, *41*, 1027.
(21) Strobel, S. A.; Cochrane, J. C. *Curr. Opin. Chem. Biol.* **2007**, *11*, 636.
(22) Lilley, D. M. J.; Eckstein, F. *Ribozymes and RNA Catalysis*; The Royal Society of Chemistry: Cambridge, 2008.
(23) Walter, N. G. *Mol. Cell* **2007**, *28*, 923.
(24) Roth, A.; Nahvi, A.; Lee, M.; Jona, I.; Breaker, R. R. *RNA* **2006**, *12*, 607.
(25) Klein, D. J.; Been, M. D.; Ferre-D'Amare, A. R. *J. Am. Chem. Soc.* **2007**, *129*, 14858.
(26) Klein, D. J.; Ferre-D'Amare, A. R. *Science* **2006**, *313*, 1752.
(27) Klein, D. J.; Wilkinson, S. R.; Been, M. D.; Ferre-D'Amare, A. R. *J. Mol. Biol.* **2007**, *373*, 178.
(28) Cochrane, J. C.; Lipchock, S. V.; Strobel, S. A. *Chem. Biol.* **2007**, *14*, 95.
(29) Soukup, G. A. *Nucleic Acids Res.* **2006**, *34*, 968.
(30) Jansen, J. A.; McCarthy, T. J.; Soukup, G. A.; Soukup, J. K. *Nat. Struct, Mol. Biol.* **2006**, *13*, 517.
(31) Banas, P.; Jurecka, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *Methods* **2009**, *49*, 202.
(32) McDowell, S. E.; Spackova, N.; Sponer, J.; Walter, N. G. *Biopolymers* **2007**, *85*, 169.
(33) Auffinger, P.; Hashem, Y. *Curr. Opin. Struct. Biol.* **2007**, *17*, 325.
(34) Hall, K. B. *Curr. Opin. Chem. Biol.* **2008**, *12*, 612.
(35) Sponer, J.; Lankas, F. *Computational Studies of RNA and DNA*; Springer: Dordrecht, The Netherlands, 2006.
(36) Cheatham, T. E. *Curr. Opin. Struct. Biol.* **2004**, *14*, 360.
(37) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. *Acc. Chem. Res.* **2010**, *42*, 40.
(38) Razga, F.; Koca, J.; Mokdad, A.; Sponer, J. *Nucleic Acids Res.* **2007**, *35*, 4007.
(39) Almlof, M.; Ander, M.; Aqvist, J. *Biochemistry* **2007**, *46*, 200.
(40) Villa, A.; Wöhnert, J.; Stock, G. *Nucleic Acids Res.* **2009**, *37*, 4774.
(41) Lee, T. S.; Lopez, C. S.; Giambasu, G. M.; Martick, M.; Scott, W. G.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 3053.
(42) Krasovska, M. V.; Sefcikova, J.; Reblova, K.; Schneider, B.; Walter, N. G.; et al. *Biophys. J.* **2006**, *91*, 626.
(43) Reblova, K.; Spackova, N.; Stefl, R.; Csaszar, K.; Koca, J.; et al. *Biophys. J.* **2003**, *84*, 3564.
(44) Auffinger, P.; Bielecki, L.; Westhof, E. *J. Mol. Biol.* **2004**, *335*, 555.
(45) Lee, T. S.; Giambasu, G. M.; Sosa, C. P.; Martick, M.; Scott, W. G.; et al. *J. Mol. Biol.* **2009**, *388*, 195.
(46) Martick, M.; Lee, T. S.; York, D. M.; Scott, W. G. *Chem. Biol.* **2008**, *15*, 332.
(47) Rhodes, M. M.; Reblova, K.; Sponer, J.; Walter, N. G. *Proc. Natl. Acad. Sci.* **2006**, *103*, 13380.
(48) Razga, F.; Koca, J.; Sponer, J.; Leontis, N. B. *Biophys. J.* **2005**, *88*, 3466.
(49) Ditzler, M. A.; Sponer, J.; Walter, N. G. *RNA* **2009**, *15*, 560.
(50) Csaszar, K.; Spackova, N.; Stefl, R.; Sponer, J.; Leontis, N. B. *J. Mol. Biol.* **2001**, *313*, 1073.
(51) Lee, T. S.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 7168.
(52) Nam, K. H.; Gao, J. L.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 4680.
(53) Trobro, S.; Aqvist, J. *Proc. Natl. Acad. Sci.* **2005**, *102*, 12395.
(54) Trobro, S.; Aqvist, J. *Mol. Cell* **2007**, *27*, 758.
(55) Krasovska, M. V.; Sefcikova, J.; Spackova, N.; Sponer, J.; Walter, N. G. *J. Mol. Biol.* **2005**, *351*, 731.
(56) Gresh, N.; Sponer, J. E.; Spackova, N.; Leszczynski, J.; Sponer, J. *J. Phys. Chem. B* **2003**, *107*, 8669.
(57) Case, D. A.; Darden, T. A.; Cheatham, I.; Simmerling, C. L.; Wang, J., *AMBER 9*; University of California: San Francisco, 2006.
(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(59) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
(60) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.
(61) Dennington, R. I.; Keith, T.; Millam, J.; Eppinnett, K.; Hovell, W. L., *GaussView, Version 3.0*; Semichem, Inc.: Shawnee Mission, KS, 2003.
(62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.;

Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.

(63) Tamura, M.; Holbrook, S. R. *J. Mol. Biol.* **2002**, *320*, 455.

(64) Hsiao, C.; Mohan, S.; Hershkovitz, E.; Tannenbaum, A.; Williams, L. D. *Nucleic Acids Res.* **2006**, *34*, 1481.

(65) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci.* **2001**, *98*, 4899.

(66) Sponer, J.; Sabat, M.; Gorb, L.; Leszczynski, J.; Lippert, B.; Hobza, P. *J. Phys. Chem. B* **2000**, *104*, 7535.

(67) Chou, S. H.; Zhu, L. M.; Reid, B. R. *J. Mol. Biol.* **1994**, *244*, 259.

(68) Shepard, W.; Cruse, W. B. T.; Fourme, R.; de la Fortelle, E.; Prange, T. *Structure* **1998**, *6*, 849.

(69) Spackova, N.; Berger, I.; Sponer, J. *J. Am. Chem. Soc.* **2000**, *122*, 7564.

(70) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J. P.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644.

(71) Leontis, N. B.; Stombaugh, J.; Westhof, E. *Nucleic Acids Res.* **2002**, *30*, 3497.

(72) Zirbel, C. L.; Sponer, J. E.; Sponer, J.; Stombaugh, J.; Leontis, N. B. *J. Biol. Struct. Dyn.* **2009**, *26*, 819.

(73) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; et al. *Biophys. J.* **2007**, *92*, 3817.

(74) Reblova, K.; Lankas, F.; Razga, F.; Krasovska, M. V.; Koca, J.; et al. *Biopolymers* **2006**, *82*, 504.

(75) Besseova, I.; Otyepka, M.; Reblova, K.; Sponer, J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10701.

(76) Fadrna, E.; Spackova, N.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E., III; Kulinski, T.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5*, 2514.

(77) Mongan, J.; Case, D. A.; McCammon, J. A. *J. Comput. Chem.* **2004**, *25*, 2038.

(78) Khandogin, J.; Brooks, C. L. *Biophys. J.* **2005**, *89*, 141.

(79) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. *Proteins: Struct. Funct. Bioinf.* **2004**, *56*, 738.

(80) Meng Y.; Roitberg, A. E. *J. Chem. Theory Comput.* **2010**, *6*, 1401−1412.

(81) Cochrane, J. C.; Lipchock, S. V.; Smith, K. D.; Strobel, S. A. *Biochemistry* **2009**, *48*, 3239.

(82) Nam, K.; Gao, J. L.; York, D. M. *RNA* **2008**, *14*, 1501.

(83) Bevilacqua, P. C.; Brown, T. S.; Nakano, S.; Yajima, R. *Biopolymers* **2004**, *73*, 90.

(84) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **1997**, *36*, 432.

(85) Zhou, D. M.; Taira, K. *Chem. Rev.* **1998**, *98*, 991.

(86) Emilsson, G. M.; Nakamura, S.; Roth, A.; Breaker, R. R. *RNA* **2003**, *9*, 907.

(87) Breaker, R. R.; Emilsson, G. M.; Lazarev, D.; Nakamura, S.; Puskarz, I. J.; et al. *RNA* **2003**, *9*, 949.

(88) Liu, L.; Cottrell, J. W.; Scott, L. G.; Fedor, M. J. *Nat. Chem. Biol.* **2009**, *5*, 351.

(89) Mlynsky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2010**, *114*, 6642–6652.

JP9109699