Supporting information for: Magnesium Fluctuations Modulate RNA Dynamics in the SAM-I Riboswitch

Ryan L. Hayes,[†] Jeffrey K. Noel,[†] Udayan Mohanty,[‡] Paul C. Whitford,[¶] Scott P. Hennelly,[¶] José N. Onuchic,[§] and Karissa Y. Sanbonmatsu[*,¶]

Center for Theoretical Biological Physics and Department of Physics, University of California San Diego, San Diego, CA, Department of Chemistry, Boston College, Chestnut Hill, MA, Theoretic Biology and Biophysics, Theoretic Division, Los Alamos National Labs, Los Alamos, NM, and Center for Theoretical Biological Physics and Department of Physics & Astronomy, Rice University, Houston, TX

E-mail: kys@lanl.gov

1 Correlations and Autocorrelations

To determine the characteristic times for fluctuations of normal modes, fluctuations in density, or residence times, autocorrelations were computed. For some quantity q of interest, the normalized autocorrelation $AC(\Delta t)$ is given by

$$\tilde{q}(t) = q(t) - \langle q \rangle \tag{S1}$$

$$AC(\Delta t) = \frac{\langle \tilde{q}(t)\tilde{q}(t+\Delta t)\rangle_t}{\langle \tilde{q}^2 \rangle} \tag{S2}$$

[*]To whom correspondence should be addressed

[†]Center for Theoretical Biological Physics and Department of Physics, University of California San Diego, San Diego, CA

[‡]Department of Chemistry, Boston College, Chestnut Hill, MA

[¶]Theoretic Biology and Biophysics, Theoretic Division, Los Alamos National Labs, Los Alamos, NM

[§]Center for Theoretical Biological Physics and Department of Physics & Astronomy, Rice University, Houston, TX

where the averages $\langle\rangle$ are computed over time. For a given Δt, autocorrelation values close to unity indicate the system is unlikely to have changed, while autocorrelations near zero indicate that the system has little memory over this time interval. The autocorrelation provides a measure of the duration of fluctuations of the system. A representative time for fluctuations can be obtained by performing the integral

$$\tau = \int_0^\infty AC(t)dt \tag{S3}$$

For an exponentially decaying autocorrelation, Eq. (S3) gives the time constant of the exponential. In the case that the data is noisy or has an unreliable tail, this time constant can alternatively be obtained from the time interval for the autocorrelation to fall to $1/e$.

Correlations were computed to establish connections between Mg^{2+} dynamics and RNA dynamics. For two time-varying quantities p and q of interest, the correlation is given by

$$C = \frac{\langle \tilde{p}(t)\tilde{q}(t)\rangle}{\sqrt{\langle \tilde{p}^2\rangle \langle \tilde{q}^2\rangle}} \tag{S4}$$

where $\tilde{p}$ and $\tilde{q}$ have had had their means subtracted.

2 Normal Modes

As outlined in the main text, global coordinates are needed to characterize the global dynamics of the RNA. That is, it is useful to project our explicit solvent molecular dynamics trajectories onto a set of global coordinates for analysis. Both normal modes[1–3] and principal components[4] can capture the global dynamics, but normal modes are used because they provide a consistent measure between all simulations. To compute normal modes, we used a structure-based potential.[5,6] The mass weighted Hessian matrix H_{ij} of second derivatives of the potential about the minimum was computed.

$$H_{ij} = \frac{1}{\sqrt{m_i m_j}} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} U \tag{S5}$$

This matrix was diagonalized to produce the eigenfrequencies and modes of the system, with the

lower frequency modes corresponding to the larger global motions of the RNA.

Normal modes are typically computed in proteins with the Tirion potential. The Tirion potential is an ultra-simple forcefield based on the geometry of the C_α atoms. Tirion's original paper[7] demonstrated that it was the geometry of the protein which controlled the lowest frequency normal modes, and not the details of the potential, so the Tirion potential gave good agreement with normal modes obtained from more detailed potentials. It is unclear how to apply the Tirion potential for proteins to an RNA system, and whether one should include one, three, or many beads per residue to adequately capture the geometry of RNA, due to its larger residues. In the present case, an established all atom structure-based potential is used for computing normal modes. While this potential is substantially more detailed than the Tirion potential, the primary motivation to use it is that it is simpler to apply to RNA.

Principle components, also known as essential dynamics, are another possible global coordinate.[4] Principle components are computed by diagonalizing the mass weighted covariance matrix

$$C_{ij} = \left\langle \sqrt{m_i}\left(x_i(t) - \langle x_i \rangle\right) \sqrt{m_j}\left(x_j(t) - \left\langle x_j \right\rangle\right) \right\rangle \tag{S6}$$

Since the covariance matrix is dependent on the trajectory, principal components are dependent on the degree of sampling and are not necessarily consistent between simulations.

Normal modes and principal components are related in that for simulations at very low temperatures, they will coincide because there will not be enough thermal energy to explore the landscape beyond where the harmonic approximation implicit in the Hessian matrix breaks down. At higher temperatures normal modes are advantageous because there is no need to wait for convergence, but principal components are advantageous because they capture the global motions of the system beyond the harmonic basin of attraction.

Because of the need to compare global motions between simulations, we use normal modes. We compare the subspace overlap between normal modes and principle components to more precisely relate normal modes to principal components (Figure S1). A subspace overlap C_i determines

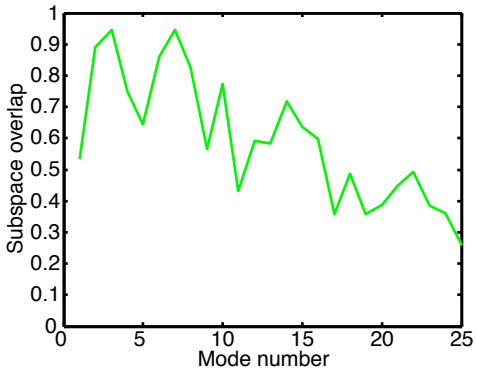

Figure S1: The subspace overlap between each principal component and the subspace of the first 50 normal modes. The first 10 principal components are well represented, so normal modes are able to represent these most important motions of the RNA.

what portion of a given mode can is captured by a linear combination of modes of another basis.

$$C_i = \sum_{j=1}^{50} \left(\vec{p}_i \cdot \vec{q}_j\right)^2 \tag{S7}$$

Here we look at how completely a given principal component $\vec{p}_i$ is represented within the subspace of the first 50 normal modes $\vec{q}_j$. The first 10 principal components are well represented, after that the overlap begins to decline. Thus normal modes are able to represent the first 10 principal components which give the most important motions of the RNA.

3 Time Scales

Further time scale details omitted in Figure 2 are shown in Figure S2. The normal mode autocorrelations were averaged together over the first 50 normal modes, to simplify presentation of the data (Figure S2a, inset). These curves show the long non-exponential tails of the autocorrelation, and also reflect the slowing of RNA fluctuations with increasing Mg^{2+}. Integration of these curves (Eq. (S3)) biases longer time scales, and gave time scales that increased from 6.8 ns for $0^{\dagger}$ mM to 12.7 ns for 10 mM. Likewise, the association site local density autocorrelations were also averaged together (Figure S2b, inset). The resulting autocorrelation curves are much more exponential in character. The time scale of local density fluctuations is 3.7 ns with the 10 Å shell and 3.4 ns with

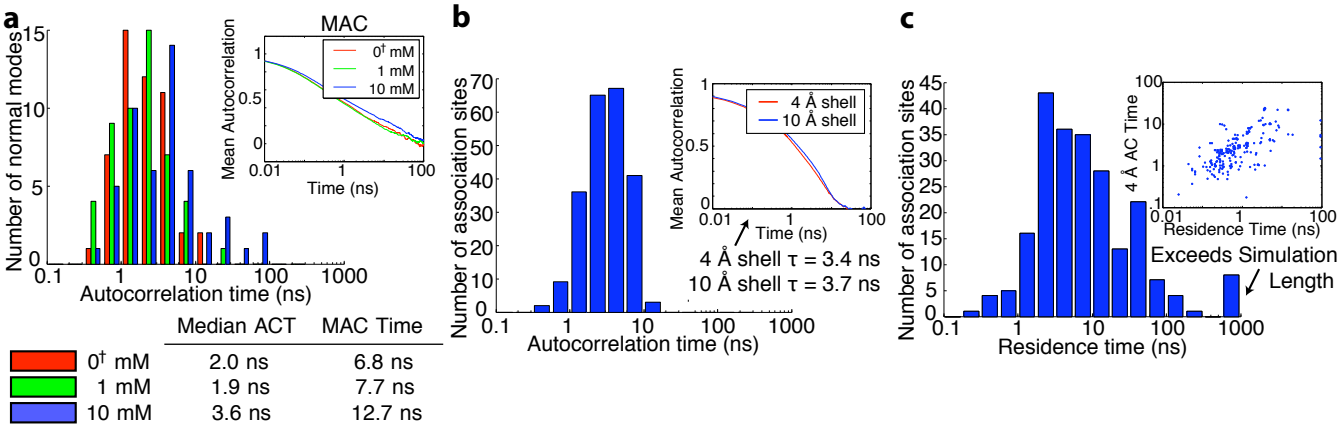

Figure S2: Mg^{2+} affects RNA kinetics. (a) The distribution of autocorrelation times for the MD trajectories projected onto the lowest 50 normal modes. The inset shows the autocorrelation averaged over all normal modes. The time scales of RNA fluctuations increase with Mg^{2+} concentration in accordance with experiment. This occurs because Mg^{2+} tightens RNA tertiary contacts and stabilizes compact conformations. (b) The distribution of autocorrelation times for the local Mg^{2+} density around several association sites. The inset shows the mean of the autocorrelation of the local Mg density averaged over all sites. The time scales of RNA fluctuations are comparable to the time scale of the RNA fluctuations, which allows them to kinetically couple. (c) The residence time for the association sites. The inset shows the comparison between residence time and autocorrelation time for each site.

the 4 Å shell. Thus, both 10 Å and 4 Å have comparable time scales, even though a 4 Å shell only captures if a Mg^{2+} is in that site, and 10 Å captures a slightly broader picture of the local density. Comparison was made between density fluctuation times from panel b and residence times in panel c (Figure S2c, inset), and correspondence was found for 4 Å spheres, as mentioned in the main text.

4 Locations of Fluctuation Times

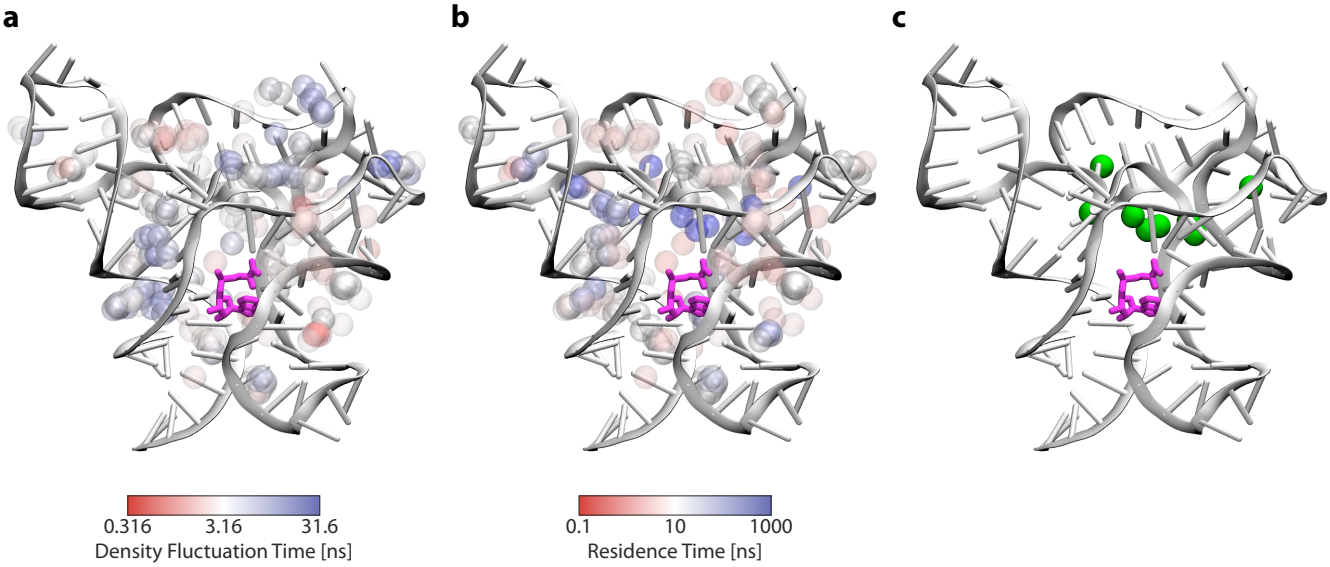

Figure S3: The fluctuation times and residence times plotted by site. Spheres are plotted with the Van der Waals radius for Mg^{2+} (a) Density Fluctuation Times. (b) Residence Times. (c) 8 association sites with residence times longer than the simulation. Two pairs of association sites are redundant, so only 6 distinct sites were identified.

5 Local Mg^{2+} Events Affect Local Dynamics

Local densities were computed within a 10 Å shell of the electronegative backbone atoms. These local densities were correlated with P-P distance to give a degree of correlation as a function of sequence along the RNA. While this method had some advantages over the analysis employed in Figure 3, it was eventually abandoned because spurious correlations were arising. The electronegative atoms around which the local density was computed were registering changes in local density as the atoms moved with respect to more fixed association sites. In this figure, both P-P distance and Mg density were smoothed with a 1 ns long moving average. Such smoothing was not necessary for the calculations within the main text.

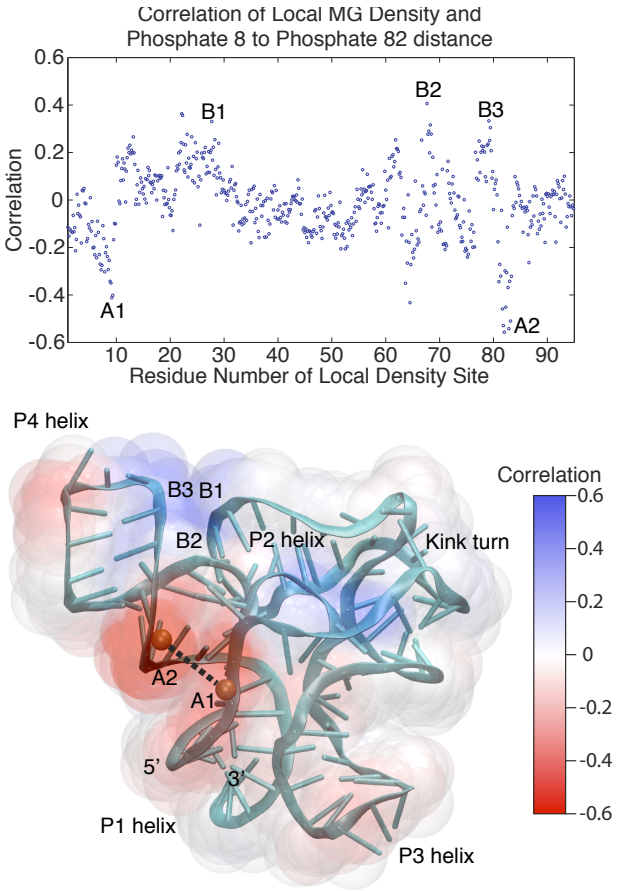

Figure S4: Correlations are shown between the phosphate 8 to phosphate 82 distance (dashed line) and the local magnesium density around 592 electronegative RNA atoms. Correlations are plotted by sequence in the top panel and by location in the structure in the bottom panel. A1 and A2 show one binding site while B1, B2, and B3 show a competing site.

6 Local Mg^{2+} Events Affect Global Dynamics

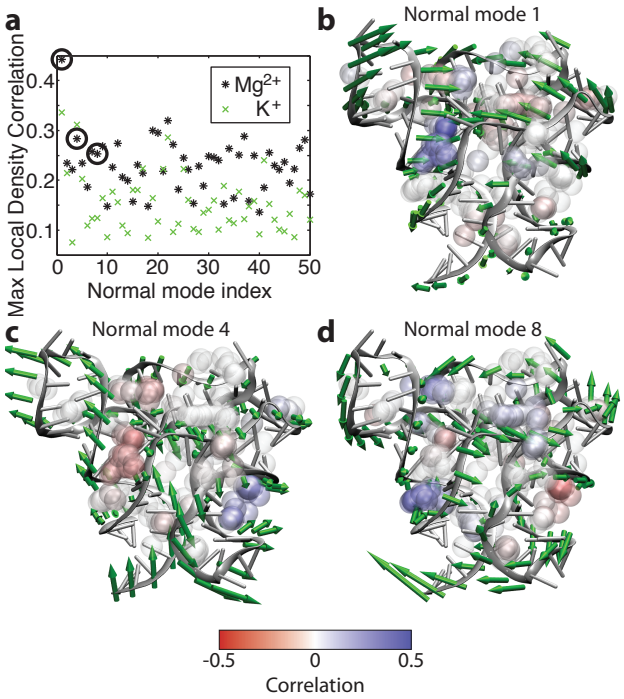

Figure S5: Mg^{2+} association and resulting local bridging interactions also affect the global RNA conformation. (a) Correlations between normal mode projections and the most strongly correlated local magnesium density for each normal mode. Several of the first few normal modes are strongly correlated. Circled modes are shown in (b-d). The positive direction of motion along the normal mode is plotted as green arrows. Densities in regions which close upon positive motion along the normal mode are generally positively correlated (blue), while densities in regions which open are generally negatively correlated (red). Balls represent association sites, and are plotted one quarter scale as in Figure 3.

Figure S5a shows the first 50 normal modes whereas Figure S5a only shows the first 25. Twenty[8] to thirty[7] normal modes is generally considered enough, but we include 50 here for completeness and consistency with the 50 normal modes considered in previous sections. Figure S5c&d are referenced in the main text, but were omitted for space considerations. They are included here for completeness and to demonstrate that the effect of Mg^{2+} bridging interactions tugging on normal modes is not unique to a single normal mode.

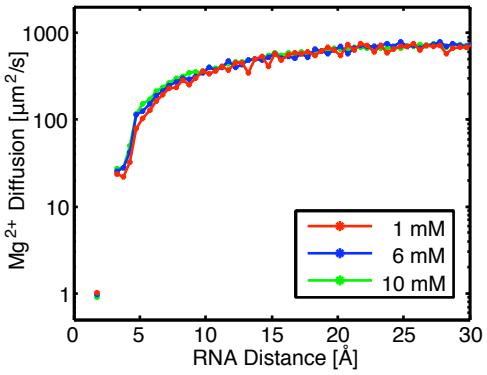

Figure S6: Diffusion as a function of distance for the three non-zero Mg^{2+} concentrations.

7 Diffusion of Mg^{2+} Ions

The diffusion effects mentioned in the main text are robust with respect to concentration (Figure S6). The different Mg^{2+} concentrations show the same diffusion in all but the outer-sphere regime. The 1 mM Mg^{2+} simulation shows slighly lower values of diffusion in the outer-sphere regime because Mg^{2+} is more scarce and fills longer lived association sites first, resulting in lower apparent diffusion.

8 Spatial Characterization of Transient Surface Mg^{2+}

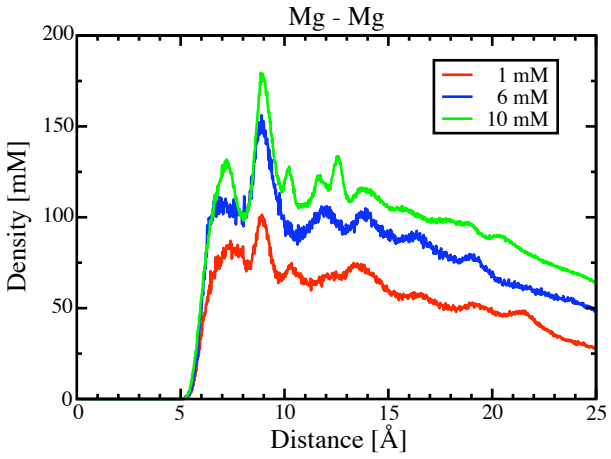

Figure S7: The RDF of Mg^{2+} with other Mg^{2+} ions. Mg^{2+} spacing was about 6.5 Å.

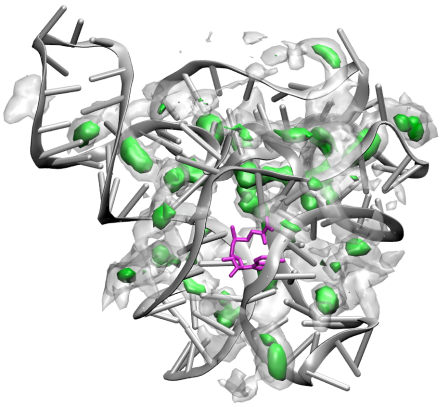

Figure S8: Spatial distribution functions (SDF) of Mg^{2+} about SAM-I. Mg^{2+} density surfaces of 1.4 M and 7 M are shown in transparent white and solid green respectively (10 mM simulation with SAM). The Mg^{2+} distribution around the whole riboswitch is shown here. More localized distributions can be found in Figure 7 of the main text.

For computation of SDFs in Figures 7 and S8, each frame of the trajectory was fitted to each of the 93 phosphates, to the P1 helix, or to the whole riboswitch, and the magnesium locations were binned. Bin sizes were 0.5 Å for the phosphate SDF and 1.0 Å for the helix and riboswitch SDFs. The resulting probability density can be expressed as molarity ($1\ \mathrm{M} = 0.6022\ \mathrm{nm}^{-3}$). Regions of 1.4 M and 7 M were plotted in VMD using isosurfaces.[9] We express these values in molarity to emphasize the immense enrichment of 700-fold for outer-sphere Mg^{2+} over bulk Mg^{2+} concentrations.

Figure S8 shows the Mg^{2+} about the entire riboswitch. The sites in Figure 7b are visible on the back side of this view. There is a preference for the backbones and major grooves, along with the central regions where there is more negative RNA charge to balance.

On the order of 25 regions of Mg^{2+} enrichment exceeding 7 M concentration are visible. These regions are connected by regions of lower enrichment. From the Table 2 in the main text, 22.3 of the magnesium ions are in the inner-sphere regime on average, so on average most of these regions will b occupied. We expect a similar occupation pattern would apply at lower concentration, with the exception that longer lived sites should fill up preferentially. That is, the bluer regions of Figure S3b should be more full. Thus the 8.6 outer-sphere Mg^{2+} in 1 mM Mg, should preferentially fill up the

P1 to P4 major groove and the aptamer core.

9 Mg^{2+} Concentration Affects RMSF

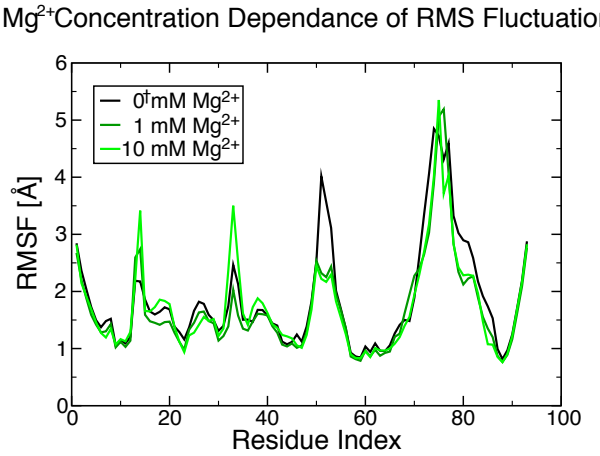

Figure S9: Shown are the root mean square fluctuations by residue of each phosphate in the riboswitch. Fluctuations are especially high in the kink turn (residues 15 and 33) and in the loop regions of P3 and P4 (G50 to A53 and G74 to A77), though magnesium concentration has varying effects on these regions.

As a first test, we checked the changes in root mean squared fluctuations of the phosphates as a function of Mg^{2+} concentration. Results show that some regions display enhanced fluctuations, and others quenched fluctuations in the presence of Mg^{2+} (Figure S9). Residues near 15 and 33 (the kink turn break between helices P2a and P2b) show enhanced fluctuation in the presence of Mg^{2+} because there is a binding region that Mg^{2+} pops in and out of, resulting in greater fluctuations at the base of the P2 helix. On the contrary, residues 50-53 (tip of P3 helix) and 79-82 (base of P4 helix) show decreased fluctuations because Mg^{2+} association enhances RNA stability and keeps these regions anchored down.

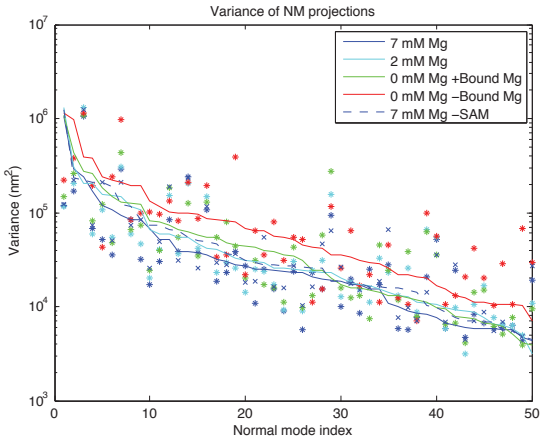

Figure S10: Plotted are the variances of the normal modes. Dots represent the variance along a particular normal mode, plotted by the index of the normal mode; while lines represent the variances sorted by magnitude. It is of note that the removal of all bulk magnesium (green line) affects the stability as much as the removal of the SAM ligand (dashed blue line)

10 Higher Mg^{2+} Concentration Quenches Normal Mode Fluctuations

Analysis of the normal modes shows that Mg^{2+} quenches global fluctuations of the RNA. Figure S10 shows the variance of the normal modes (dots) plotted by the normal mode index. The variances are also sorted by magnitude (lines) to ease comparison between varying concentrations. The variances are about twice as large in the 0 mM Mg^{2+} simulation as in the 7 mM Mg^{2+} simulation. Removal of the inner-sphere bound magnesium that connects A9 and U63 causes and additional factor of two increase in the variances. This destabilization is almost as great as the removal of SAM, which also results in nearly a factor of two increase in variance within the 7 mM simulations. The drift of the mean of the normal modes from the crystal structure reflects a similar destabilization with decreasing Mg^{2+}. Simulations without as much Mg^{2+} drift further in the same amount of time. These simulations are not long enough for unfolding to occur to compare relative free energies, but the changes in fluctuations hint at changes in stability because of increased attempt rates reflected by larger variances and drift of the mean.

11 K^+ Controls

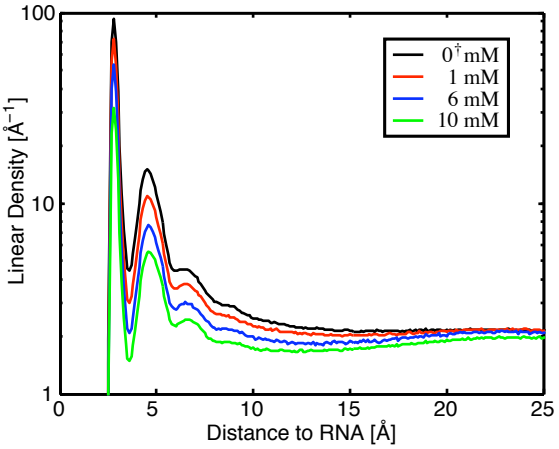

Figure S11: The populations of K^+ ions in various distance regimes from the RNA.

Table S1: Shows the number and percentage of the preferred interaction coefficient for K^+ ions in each hydration layer from Figure S11.

	$0^\dagger$ mM	1 mM	6 mM	10 mM
Γ_{K^+}	75.4	57.3	41.7	26.6
	Number of K^+			
Inner-Sphere	34.6	26.2	19.1	11.7
Outer-Sphere	21.9	16.1	11.5	8.6
Diffuse	19.0	15.1	11.1	6.3
	Percent of Γ_{K^+}			
Inner-Sphere	45.8	45.7	45.8	44.0
Outer-Sphere	29.0	28.0	27.6	32.5
Diffuse	25.2	26.3	26.7	23.5

The populations of K^+ (Figure S11 and Table S1) are much more uniform between regimes than the populations of Mg^{2+} (Table 2). The substantial enrichment of the outer-sphere population is a feature unique to Mg^{2+}.

12 Box Size Controls

Table S2: Dependence of corrected concentrations and preferential interaction coefficients on simulation periodic box size. Following Chen, et al.,[10] we have performed simulations with different box sizes. In this test study, we performed 100 ns simulations. We note that our other simulations described in this paper were 2000 ns trajectories.

	100 Å	120 Å	150 Å
$[Mg^{2+}]$	0.9 mM	0.7 mM	0.7 mM
$[K^{+}]$	101.1 mM	102.5 mM	103.4 mM
$[Cl^{-}]$	102.8 mM	103.9 mM	104.8 mM
$\Gamma_{Mg^{2+}}$	10.5	10.3	10.7
$\Gamma_{K^{+}}$	58.1	57.4	55.7
$\Gamma_{Cl^{-}}$	-12.8	-14.1	-15.0

References

(1) Levitt, M.; Sander, C.; Stern, P. S. *International Journal of Quantum Chemistry: Quantum Biology Symposium* **1983**, *10*, 181–199.

(2) Gō, N.; Noguti, T.; Nishikawa, T. *Proceedings of the National Academy of Sciences of the United States of America* **1983**, *80*, 3696–3700.

(3) Brooks, B.; Karplus, M. *Proceedings of the National Academy of Sciences of the United States of America* **1983**, *80*, 6571–6575.

(4) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins: Structure, Function, and Bioinformatics* **1993**, *17*, 412–425.

(5) Whitford, P. C.; Schug, A.; Saunders, J.; Hennelly, S. P.; Onuchic, J. N. *Biophysical Journal* **2009**, *96*, L7–L9.

(6) Noel, J. K.; Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N. *Nucleic Acids Research* **2010**, *38*, W657–W661.

(7) Tirion, M. M. *Physical Review Letters* **1996**, *77*, 1905–1908.

(8) Krebs, W. G.; Alexandrov, V.; Wilson, C. A.; Echols, N.; Yu, H.; Gerstein, M. *Proteins: Structure, Function, and Bioinformatics* **2002**, *48*, 682–695.

(9) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33–38.

(10) Chen, A. A.; Draper, D. E.; Pappu, R. V. *Journal of Molecular Biology* **2009**, *390*, 805–819.