# Supporting information for: Divalent metal ion activation of a guanine general base in the hammerhead ribozyme: insights from molecular simulations

Haoyuan Chen,[‡] Timothy J. Giese,[‡] Barbara L. Golden,[§] and Darrin M. York[*,‡]

*Center for Integrative Proteomics Research, Laboratory for Biomolecular Simulation Research and Department of Chemistry & Chemical Biology, Rutgers University, 174 Frelinghuysen Road, Piscataway, New Jersey 08854-8076, United States , and Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907, United States*

E-mail: Darrin.York@rutgers.edu

RUNNING HEADER: Divalent metal ion activation of a guanine general base in the hammerhead ribozyme: insights from molecular simulations

[*]To whom correspondence should be addressed

[‡]Center for Integrative Proteomics Research, Laboratory for Biomolecular Simulation Research and Department of Chemistry & Chemical Biology, Rutgers University, 174 Frelinghuysen Road, Piscataway, New Jersey 08854-8076, United States

[§]Department of Biochemistry, Purdue University, West Lafayette, Indiana 47907, United States

## Derivation of the correction for ion concentration and binding affinity in DFT calculations of $Mg^{2+}$-guanine model complexes

In the DFT calculations of the $Mg^{2+}$-induced $pK_a$ shift of guanine in model complexes, the free energy differences between species need to be further corrected by accounting for the metal ion concentration and the metal-guanine binding affinity to yield the prediction of $pK_a$ shift values. Consider the kinetic model below,

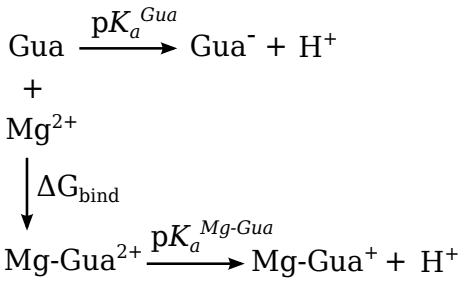


Figure S1: Kinetic model used in the $pK_a$ shift correction. $pK_a$ values are directly converted from the free energy differences in DFT calculations while $\Delta G_{bind}$ is the experimentally measured binding affinity.

we have

$$K_a(Mg-Gua, corr.) = \frac{c_{Gua} \times 10^{-pK_a(Gua)} + c_{Mg-Gua} \times 10^{-pK_a(Mg-Gua, uncorr.)}}{c_{Gua} + c_{Mg-Gua}} \tag{1}$$

in which all concentrations have been divided by the standard molar concentration 1.0 M and therefore are dimensionless. Since

$$c_{Mg-Gua} = c_{Gua} \times c_{Mg} \times e^{-\beta \Delta G_{bind}} \tag{2}$$

where $\beta = \frac{1}{k_B T}$ , Eq. (1) becomes

$$K_a(Mg{-}Gua, corr.) = \frac{c_{Gua} \times 10^{-pK_a(Gua)} + c_{Gua} \times c_{Mg} \times e^{-\beta\Delta G_{bind}} \times 10^{-pK_a(Mg-Gua,uncorr.)}}{c_{Gua} + c_{Gua} \times c_{Mg} \times e^{-\beta\Delta G_{bind}}} \tag{3}$$

where $c_{Gua}$ could be eliminated from both numerator and denominator, which leads to

$$K_a(Mg - Gua, corr.) = \frac{10^{-pK_a(Gua)} + c_{Mg} \times e^{-\beta\Delta G_{bind}} \times 10^{-pK_a(Mg-Gua,uncorr.)}}{1 + c_{Mg} \times e^{-\beta\Delta G_{bind}}} \tag{4}$$

that gives Eq. (2) in the main text. The corrected $Mg^{2+}$-induced p$K_a$ shift on guanine is therefore

$$\Delta pK_a(Mg, corr.) = -log_{10}(\frac{10^{-pK_a(Gua)} + c_{Mg} \times e^{-\beta\Delta G_{bind}} \times 10^{-pK_a(Mg-Gua,uncorr.)}}{1 + c_{Mg} \times e^{-\beta\Delta G_{bind}}}) - pK_a(Gua) \tag{5}$$

## Comparison between N7 and O6 binding sites

Recent studies from Auffinger and co-workers have suggested that inner sphere $Mg^{2+}$-Guanine:N7 binding might not be very favorable(S1, S2). Here, we performed the electronic structure calculations and thermodynamic integration simulations on the HHR system with other $Mg^{2+}$-Guanine binding modes as well. As shown in the table below, inner sphere coordination with either N7 or O6 leads to a significant and similar p$K_a$ shift on N1.

Table S1: Comparison of calculated p$K_a$s of G12:N1 in different $Mg^{2+}$ binding modes. I and O in parentheses stand for inner sphere and outer sphere binding, respectively.

| Binding Mode | p$K_a$ (QM) | p$K_a$ (TI) |
|---|---|---|
| N7 (I), O6 (O) | 2.2 | 8.0 ± 0.4 |
| N7 (O), O6 (I) | 3.1 | 6.6 ± 0.8 |
| N7 (O), O6 (O) | – | 13.4 ± 0.7 |

## Choice of the basis set in QM/MM simulations

Table S2: M06-2X single point energies computed using different basis sets at the geometries optimized using M06-2X/6-31+G(d) level of theory. G and MG stands for guanine and $Mg^{2+}$-bound guanine, respectively. Absolute energies (column 2-5) are in hartree while relative energies (column 6-8) are in kcal/mol.

| Basis Set | E(G) | E($G^-$) | E($MG^{2+}$) | E($MG^+$) | $\Delta E$(G) | $\Delta E$(MG) | $\Delta\Delta E$ |
|---|---|---|---|---|---|---|---|
| 6-31G(d) | -542.3443 | -541.7819 | -1124.0730 | -1123.7497 | 352.93 | 202.86 | -150.07 |
| 6-31G(d,p) | -542.3580 | -541.7923 | -1124.1382 | -1123.8118 | 354.95 | 204.85 | -150.10 |
| 6-311G(d,p) | -542.4946 | -541.9351 | -1124.4169 | -1124.0938 | 351.10 | 202.76 | -148.34 |
| 6-31+G(d) | -542.3653 | -541.8196 | -1124.1038 | -1123.7885 | 342.42 | 197.81 | -144.61 |
| 6-31+G(d,p) | -542.3788 | -541.8300 | -1124.1669 | -1123.8486 | 344.38 | 199.73 | -144.65 |
| 6-311+G(d,p) | -542.5073 | -541.9588 | -1124.4328 | -1124.1144 | 344.15 | 199.75 | -144.40 |
| 6-31++G(d) | -542.3655 | -541.8198 | -1124.1047 | -1123.7895 | 342.44 | 197.79 | -144.65 |
| 6-31++G(d,p) | -542.3790 | -541.8302 | -1124.1677 | -1123.8495 | 344.40 | 199.70 | -144.70 |
| 6-311++G(d,p) | -542.5074 | -541.9590 | -1124.4332 | -1124.1149 | 344.18 | 199.73 | -144.45 |

The recent developed ambient potential Ewald method(S3) that was implemented in Amber allows much more efficient *ab initio* QM/MM simulations. However, it's still important to carefully choose the level of theory to balance accuracy and performance since using *ab initio* calculations in dynamics is still slow in general. In particular, the choice of basis set for the QM calculations usually affects both accuracy and performance. Since our QM/MM simulations in this work were to estimate the free energy cost of a proton transfer reaction in which a deprotonated guanine (with/without $Mg^{2+}$ bound) was the proton acceptor, we first performed some single point energy calculations to evaluate the proton affinity of deprotonated guanine with/without $Mg^{2+}$ bound. As seen in Table 1, the diffuse functions seem to play a much bigger role than the polarization and additional split-valence functions. In terms of the relative proton affinity $\Delta\Delta E$ value (which is really what we care about in the QM/MM simulations in this work), 6-31+G(d) basis set could already give very close answer comparing to the largest basis set 6-311++G(d,p). Therefore, we chose 6-31+G(d) as the basis set for QM/MM simulations.

We then conducted QM/MM simulations of the general base proton transfer reaction, using PBE0 density functional with 6-31+G(d) basis set and a smaller 6-31G(d) basis set

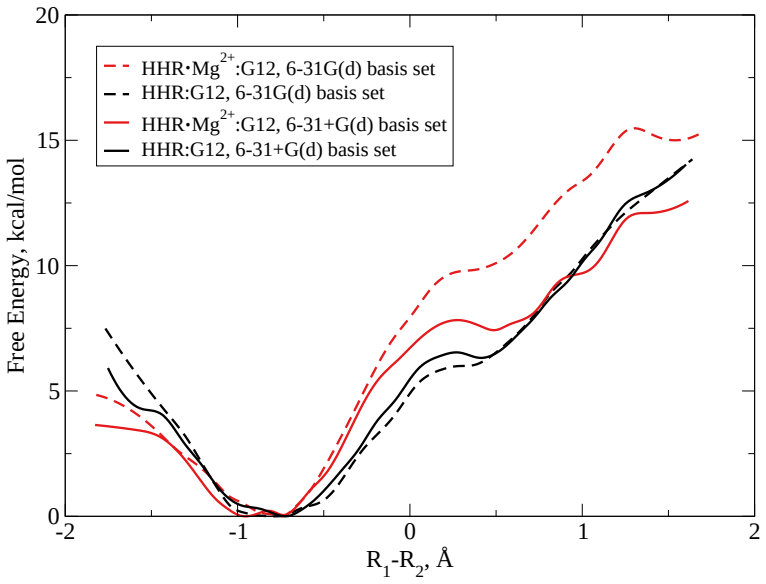


Figure S2: QM/MM free energy profile of the general base proton transfer using different basis sets.

for comparison. As seen in Figure 2, the 6-31G(d) profiles did not show clear minima for the product states while the 6-31+G(d) did. A possible explanation is that in the product state, the negative charge on the 2'O nucleophile could not be accurately modeled without diffuse functions. The overall free energies are also higher in the 6-31G(d) profiles, especially for the one with $Mg^{2+}$. Therefore, we used the data from 6-31+G(d) basis set to give the final evaluation of the free energy cost of the reaction. The simulation using 6-31+G(d) basis set is 2$\sim$3 times slower than using 6-31G(d) basis set.

## Detail of error estimation in QM/MM free energy profiles

For both simulations (with/without $Mg^{2+}$ bound at G12), bootstrapping method was used to give the error estimates. For each simulation, 1000 data points of the reaction coordinate value was saved in the 10 ps production run. In each bootstrapping step, one-fourth (250 points) of the data was randomly chosen and put in vFEP(S4) to generate a free energy profile. The step was repeated for 100 times for both simulations. Standard error of the energy difference between the two states (reactant and product) were then estimated from

the 100 samples for both profiles.

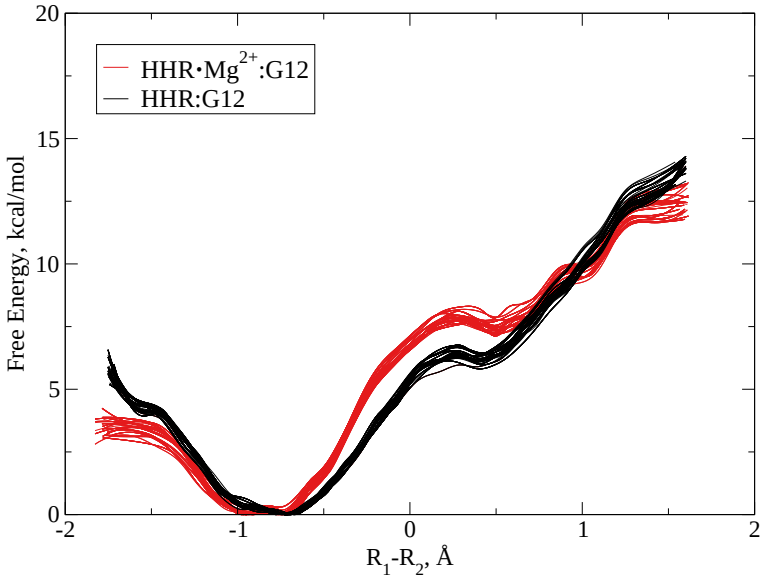


Figure S3: All the free energy profiles generated in bootstrapping.

## References

(S1) Leonarski, F., D'Ascenzo, L., and Auffinger, P. (2016) $Mg^{2+}$ ions: do they bind to nucleobase nitrogens?. *Nucleic Acid Res. 45*, 987–1004.

(S2) Leonarski, F., D'Ascenzo, L., and Auffinger, P. (2016) Binding of metals to purine N7 nitrogen atoms and implications for nucleic acids: A CSD survey. *Inorg. Chim. Acta 452*, 82–89.

(S3) Giese, T. J., and York, D. M. (2016) Ambient-Potential Composite Ewald Method for ab Initio Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulation. *J. Chem. Theory Comput. 12*, 2611–2632.

(S4) Lee, T.-S., Radak, B. K., Pabis, A., and York, D. M. (2013) A new maximum likelihood approach for free energy profile construction from molecular simulations. *J. Chem. Theory Comput. 9*, 153–164.