Published in final edited form as:

J Am Chem Soc. 2016 March 9; 138(9): 3058–3065. doi:10.1021/jacs.5b12061.

Ribozyme catalysis with a twist: the active state of the twister ribozyme in solution predicted from molecular simulation

Colin S. Gaines and Darrin M. York*

Center for Integrative Proteomics Research and Department of Chemistry & Chemical Biology, Rutgers University, 174 Frelinghuysen Road, Piscataway, NJ 08854-8076, USA

Abstract

We present results from molecular dynamics simulations and free energy calculations of the twister ribozyme at different stages along the reaction path in order to gain insight into its mechanism. Results, together with recent biochemical experiments, provide support for a mechanism involving general acid catalysis by a conserved adenine residue in the active site. Although adenine has been previously implicated as a general acid acting through the N1 position in other ribozymes such as the hairpin and VS ribozymes, in the twister ribozyme there may be a twist. Biochemical experiments suggest that general acid catalysis may occur through the N3 position, which has never before been implicated in this role; however, there currently lacks a detailed structural model for the active state of the twister ribozyme in solution that is consistent with these and other experiments. Simulations in a crystalline environment reported here are consistent with X-ray crystallographic data, and suggest that crystal packing contacts trap the RNA in an inactive conformation with U-1 in an extruded state that is incompatible with an in-line attack to the scissile phosphate. Simulations in solution, on the other hand, reveal this region to be dynamic and able to adopt a conformation where U-1 is stacked with G33. In this state, the nucleophile is in line with the scissile phosphate, and the N1 position of G33 and N3 position of A1 are poised to act as general base and acid, respectively, as supported by mutational experiments. Free energy calculations further predict the electrostatic environment causes a shift of the microscopic pK_a at the N3 position of A1 toward neutrality by approximately 5 pK_a units. These results offer a unified interpretation of a broad range of currently available experimental data that points to a novel mode of general acid catalysis through the N3 position of an adenine nucleobase, thus expanding the repertoire of known mechanistic strategies employed by small nucleolytic ribozymes.

Introduction

The twister ribozyme is a recently discovered member of the group of small nucleolytic ribozymes that catalyze the site-specific cleavage of the RNA phosphodiester backbone.

*To whom correspondence should be addressed, Darrin.York@rutgers.edu.

Supplemental Information

Additional details on residue numbering, crystal simulation, molecular dynamics simulations of the impact of U-1 stacking on the active site, and data used to predict apparent pK_a values.

Competing Financial interests

The authors declare no competing financial interests.

This group of self-cleaving ribozymes serves as an important testbed in the study of the mechanisms of RNA catalysis,^{1–6} and has impacted in the design of new biomedical technology^{7–12} and theories of evolution.^{13,14}

The twister ribozyme was originally identified from bioinformatics, and subsequently has been studied extensively by X-ray crystallography,^{15–17} mutagenesis,^{15,16,18} and other biochemical experiments.^{15,16} The twister ribozyme is of special interest in that experiments suggest that metal ions, although important for folding, do not play a direct role in catalysis.¹⁸ This implies that catalysis derives mainly from the electrostatic engineering of the active site and participation of nucleobase functional groups, which are generally considered as fairly chemically inert when compared to amino acids.^{3,19} In the case of the twister ribozyme, mutational studies have implicated that the N3 position of a conserved adenine residue, A1, is critical for activity, and may serve, in protonated form, as a general acid in the reaction. If this supposition is true, it would represent an intriguing new twist on general acid catalysis, and therefore be important in furthering our understanding of the mechanistic strategies employed by RNA enzymes.

Crystallographic data^{15–17} are suggestive in that the N3 position of A1 is in reasonably close proximity to the O5' leaving group to potentially act as a general acid. However, none of the currently available crystal structures^{15–17} are likely representative of the active state of the twister ribozyme, obfuscating their functional interpretation. Thus, there remain significant gaps in our understanding of the twister ribozyme based on currently available experimental data that prevent definitive conclusions to be drawn about its mechanism. Specifically, there currently does not exist a structural (or dynamical) model of the active state of the twister ribozyme in solution that is consistent with the current body of available experimental data. As a result, there remain several key questions as to: 1) what is the origin of the inactive state observed crystallographically^{15,17}? 2) what conformational events are required to form an active in-line conformation in solution, and what is the probability of finding the ribozyme in this conformation? 3) what residues likely act as the general base and acid? and 4) how do electrostatics shift the microscopic pK_a values of implicated general acid and base residues, and help to stabilize the transition state?

In this paper, we utilize molecular dynamics simulations in a crystalline environment and in solution, along with free energy calculations, to gain insight into these key questions. Our results provide a unified molecular-level interpretation of a broad range of experimental data^{15–18} that sheds new light onto the mechanism of the twister ribozyme.

Results and Discussion

Crystal packing causes extrusion of U-1 from the active site leading to a conformation that is not catalytically relevant

Crystallographic data provides valuable information about ribozyme structure and function; however, it remains an open question as to the degree to which a static picture of a deactivated ribozyme in a crystalline environment may have direct catalytic relevance for the active ribozyme in solution.²⁰ In the case of the twister ribozyme, with each of the three available crystal structures (PDB ID: 40JI,¹⁵ 4RGE,¹⁶ and 4QJD/4QJH¹⁷), U-1 both lacks

the O2' nucleophile and is seen to be extruded from the active site. As a consequence, these structures do not reflect a conformation where the nucleophile is both in line (so as to be able to attack the scissile phosphate) and directly interacting, through hydrogen bonding, with a residue with the potential to act as the general base. These features are generally considered to be a requirement for catalysis.²¹ This observation then begs the question: is the structure in the crystal catalytically relevant? In order to address these questions, we performed molecular dynamics simulations of the twister ribozyme, departing from the high resolution structure¹⁵ (PDBID: 4OJI), both in a crystalline environment and in solution (see Methods). We first set out to ascertain the consistency of our crystal simulations with available crystallographic data, and second, to characterize the dynamics in the region of the active site and, in particular, the local geometry associated with the orientation of U-1.

Overall, the crystal simulations were in very close agreement with available experimental data. The average structure from the crystal simulation was almost identical to the crystal structure (Figure 1), particularly in the active site where the heavy-atom root-mean-square deviation (RMSD) was 0.44 Å. Furthermore, the B-values determined from the simulated mean-square positional fluctuations (Figure 2), represent the degree of structural variation, and agree closely with those determined from crystallographic data (linear correlation coefficient of 0.90). Additional details on the observed structural fluctuations within the crystal simulation can be found in Figure S2. Taken together, this provides credence for the reliability of the simulations, and also underscores the supposition that the region of residues 17 and 18 are highly disordered as interpreted by the lack of electron density in the $2F_o - F_c$ map.¹⁵ The crystal simulations predict that U-1 is stabilized by a crystal packing contact with G14 in a symmetry related monomer, where it remains trapped for the duration of the simulation and does not sample configurations that allow the 2' position to come in line with the scissile phosphate (Figure S3). Together with the relatively modest experimental and predicted fluctuations in this region, this suggests U-1 does not sample catalytically relevant conformations in the crystal.

Simulations in solution suggest that U-1 is able to transition between three distinct conformational states

Over the course of a 120 ns unrestrained simulation of the twister ribozyme in solution, we identified three distinct conformational states of U-1 (Figure S4) that we denote as "Extruded", "Stacked" and "Triple". In the "Extruded" state, U-1 is completely solvent exposed in an orientation similar to that of the crystal (but with greater dynamical fluctuations since it is not stabilized by crystal packing contacts). In the "Stacked" state, U-1 stacks under G33 while forming a hydrogen bond with a non-bridging oxygen of A34, and in the "Triple" state the additional hydrogen bonds characteristic of a WC/H/WC (U-1/A34/A19) base triple are observed. We found these different states could be distinguished by a single coordinate, the distance between U-1:N3 - A34:N7, henceforth denoted as D_{stack} : "Extruded" ($D_{stack} > 7.0$ Å), "Stacked" (3.75 Å $< D_{stack} < 7.0$ Å) and "Triple" ($D_{stack} < 3.75$ Å). The average structures from a set of three 75 ns simulations restrained to each of the aforementioned states are shown in Figure 3 (with a similar set of structures for G33 deprotonated shown in Figure S5). Additionally, this coordinate can be used to assess the occupancy of each state in solution, and estimate the kinetics of transitions between these

states. Figure 4 shows the free energy profile along the stacking coordinate for simulations with G33, which has been implicated as a general base, in its neutral form and deprotonated at the N1 position as would be required for general base catalysis.

In both cases, simulations predict that U-1 rotation from the extruded state to the stacked state (the global free energy minimum) is essentially unhindered in solution. While the triple state is higher in free energy than the stacked state by less than 2.5 kcal/mol with G33 in its neutral form, it is considerably higher (greater than 8 kcal/mol) for G33 in the deprotonated form. Further, there is a considerable forward (~3.5 kcal/mol) and reverse (6-11.5 kcal/mol) barriers to transitions between the triple and stacked states. Overall, the stacked state is predicted to be long-lived and the dominant state in solution. As will be discussed below, this has important implications for catalysis.

Stacked state of U-1 promotes active in-line conformations and supports the role of G33 as a general base

To explore the impact of the different stacking states on the active site structure where the nucleophile adopts a catalytically active in-line conformation, we performed sets of simulations restrained to each of these states, considering G33 both in the neutral (protonated at N1) and anionic (deprotonated at N1) forms (see Methods).

Figure 5 shows a profile for the probability of finding the nucleophile in an active in-line attack conformation for each of the stacking states with G33 in its ground state neutral form. Whereas active in-line attack conformations are not observed in the extruded state, and are partially observed (51.0% occupancy) in the triple state, they are dominantly clustered in the stacked state. This clustering is preserved (96.4% occupancy) in simulations with G33 activated (deprotonated) for base catalysis (Figure 5: G33N1⁻ Stacked), and further analysis reveals this state is marked by preservation of key hydrogen bonds between the endocyclic and exocyclic nitrogens of G33 with the nucleophile 2'OH and non-bridging oxygen of the scissile phosphate, respectively (Figure 5). This provides strong computational support that the twister ribozyme forms in-line attack conformations while in the dominant stacked state in solution, and upon deprotonation of G33, forms a hydrogen bond network that facilitates nucleophile activation with the N1 position of G33 acting as a general base catalyst.

With G33 poised to act as a base (G33N1⁻), the sampling of in-line conformations significantly increases for both the "Triple" and "Extruded" states, with occupancies of 93.1% and 57.5%, respectively (Figure S6). However, the sampling of in-line conformations is only one of the bases for developing a model of the twister ribozyme's mechanism. With this in mind, the "Triple" state is disfavored as the most relevant active state, firstly, because of the higher free energy associated with sampling that state relative to the "Stacked" or "Extruded" states (Figure 4) and secondly, because of the loss of the hydrogen bond between G33N2 and the non-bridging phosphoryl oxygen of A1, which aids in neutralizing the negative charge on the phosphate (Figure 3c and Figure S5c). In contrast, with G33 deprotonated, the "Extruded" state does exhibit the important hydrogen bonding interactions between G33 and A1. However, the "Stacked" state is both lower in free energy and has a significantly higher probability of the nucleophile being positioned in line, when compared to the "Extruded" state.

In addition to the six restrained simulations exploring the impact of U-1 stacking on the sampling of in-line conformations and the potential role of G33 as both a general base and hydrogen bond donor to the scissile phosphate, one additional restrained simulation was run to examine the interaction between the proposed general acid (A1:N3) and the 5'-oxygen leaving group. As with the other simulations of U-1 in the "Stacked" state, the probability of finding the ribozyme in an in-line conformation (Figure 6a) is extremely high (99.1% occupancy). Additionally, the hydrogen bonding between G33 and either the nucleophile and the pro-R non-bridging oxygen are observed in a majority of simulation snapshots. As seen in Figure 6b, the distribution of distances between A1:N3 and A1:O5' is bimodal, indicating that there are two distinct states being sampled. The first state (Figure 6c), is strikingly similar to the G33N1⁻ "Stacked" state (Figure S5b) where the general acid is neutral, and there is no direct interaction between the presumed acid and the leaving group. However, a small fluctuation in the position of A1 (Figure 6d) allows for N3 to donate a hydrogen bond to the 5'-oxygen, enhancing it as a leaving group. This data is consistent with the proposed active state model, where U-1 stacks beneath the general base G33 and A1:N3 is poised to act as the general acid.

Transition state mimic simulations further suggest A1:N3 acts as a general acid

Simulations of a transition state mimic (TSM) allow the prediction of the active site architecture and hydrogen bonding network at a critical point along the reaction pathway. Simulations of the TSM were performed with the twister ribozyme in a protonation state that would support general acid/base catalysis in accordance with implicated residues identified through mutagenesis and biochemical studies.¹⁵ Specifically, both the N3 position of A1 (the presumed general acid) and the N1 position of G33 (the presumed general base) were protonated, representing a transition state that follows from the active state (designated as AH^+EB^-) where G33 was negatively charged and has accepted a proton from the 2'-oxygen nucleophile. Under these assumptions, examination of the TSM simulations serves as a predictive test; if the integrity of the active site and catalytic requirements of the hydrogen bond network is preserved, this provides support for the roles of A1 and G33 as general acid and base, respectively. On the other hand, if the TSM simulations are not stable, this creates strong evidence against the working hypothesis.

Overall, simulations of the TSM in the stacked state preserve the key hydrogen bonds necessary for catalysis, while providing important electrostatic stabilization of the negatively charged transition state. Figure 7 shows that G33 maintains hydrogen bonding with the pro-R oxygen of the scissile phosphate, while at the same time, A1 donates a hydrogen bond via the protonated N3 position to the O5' leaving group. Results of the TSM simulations in the triple state (Figure S7b) are generally similar with those of the stacked state. Whereas, comparison of the TSM simulations in the extruded state (Figure S7a) are markedly different. In particular, there is a notable rotation of the backbone alpha and gamma torsion angles for the scissile phosphate. The alpha/gamma torsion angles flip from g^-/g^- in the "Extruded" state to t/g^+ and g^+/t for the "Triple" and "Stacked" states, respectively. The rearrangement of the backbone to t/g^+ or g^+/t allows for the 5' leaving group to be positioned in close proximity to A1. These results suggest U-1 must stack underneath G33 in order to preserve the necessary hydrogen bond network to promote general acid/base catalysis by

A1/G33, consistent with the "Stacked" state as the preferred working model of catalysis for the twister ribozyme.

Local electrostatic environment shifts the pK_a of A1:N3 toward neutrality

Although the present simulations, along with crystallographic data,¹⁵ indicate that A1:N3 is positioned to donate a hydrogen bond to the O5' leaving group, the microscopic pK_a of the N3 site for adenosine in solution has been predicted to be quite low (1.5 ± 0.3),²² raising possible objections to its effectiveness as an acid catalyst near neutral pH.¹⁷ This poses a question as to whether the electrostatic environment in the active site might cause the pK_a of A1:N3 to shift toward neutrality, and if so, to what extent. The observation of such a large shift in the pK_a of a nucleic acid functional group would not be unprecedented in either experimental^{23,24} or computational studies.²⁵ Therefore, we have performed free energy calculations to predict the apparent pK_a values for the proposed general acid (A1:N3) and base (G33:N1) in order to address this key question.

In keeping with the microscopic kinetic model for general-acid base catalysis²⁶ (Figure 8), the free energy (and thus the pK_a) for each leg of the cycle was calculated using a series of thermodynamic integration calculations (see Methods). From these free energy calculations, predicted fractions of each microstate could be determined from the partition function Q ,

$$Q = 1 + 10^{(pK_{a,B}^{AH^+} - pH)} + 10^{(pK_{a,B}^{AH^+} - pK_{a,A}^{BH})} + 10^{(pH - pK_{a,A}^{B^-})} \quad (1)$$

with the fraction of the active state, $_{AH^+}E_B^-$, determined as

$$f_{(AH^+/B^-)} = 1/Q \quad (2)$$

The observed rate, which is pH-dependent, is presumed to be directly proportional to the active fraction according to the relation:

$$k_{obs} = k_{cl} \times f_{(AH^+/B^-)} \quad (3)$$

where k_{cl} is a proportionality constant, representing the pH independent rate of cleavage. Furthermore, the simulated active fraction ($_{AH^+}E_B^-$) can be fit to a two state apparent pK_a model²⁷ that assumes the pK_a values of the general acid and base are uncorrelated. This model has the three free parameters k_{cl} , $pK_{a,A}$, and $pK_{a,B}$, with the latter two parameters representing the apparent pK_a of the acid and base participating in catalysis, respectively:

$$k_{obs} = k_{cl} \times 1 / \left[1 + 10^{(pK_{a,B} - pH)} + 10^{(pK_{a,B} - pK_{a,A})} + 10^{(pH - pK_{a,A})} \right] \quad (4)$$

By fitting the function as described in Eq. (4) to the simulated active fraction (i.e., in the same manner that experimental activity-pH curves are analyzed and interpreted), one obtains a predicted pH-activity curve and apparent pK_a values that can be compared with experiment (Figure 9). The predicted apparent pK_a values of 6.5 ± 0.4 and 9.0 ± 0.4 for A1:N3 and G33:N1 acting as the general acid and base respectively, are very close to the experimentally observed values of 6.9 and 9.5, respectively, and the overall shape of the simulated and experimental pH-activity profiles are very similar.¹⁵ These calculations support the hypothesis that the local environment within the active site of the twister ribozyme significantly shifts the pK_a of A1:N3 (by ≈ 5 pK_a units) towards neutrality, to provide a new mode of general acid-base catalysis.

Conclusion

The simulations presented here provide a uniform interpretation of a large body of experimental data in terms of a molecular-level description of catalytic mechanism for the twister ribozyme. The results depict a mechanism whereby U-1, which is trapped in an extruded state in the crystal due to crystal packing contacts, adopts a stable stacked state with G33 in solution. In this stacked state, the active site is able both to adopt an active in-line attack conformation of the nucleophile with the scissile phosphate and to position nucleotide residues to act as general acid and base catalysts. The transition state mimic simulations support a model whereby G33:N1 maintains hydrogen bonding with the O2' position (which has a partial covalent bond to phosphorus) and A1:N3 donates a hydrogen bond to the O5' position of the leaving group. Finally, free energy calculations suggest that the active site provides electrostatic stabilization of the transition state, and shifts the pK_a of A1:N3 toward neutrality by almost 5 pK_a units. Together, these results support the hypothesis that the twister ribozyme employs a new mode of general acid-base catalysis involving the N3 position of a protonated adenine, adding to the growing body of knowledge about the diversity of strategies employed by catalytic RNAs.

Methods

Crystal Simulation

In keeping with the methodology set forth in the literature,^{28–31} we completed a 120 ns simulation of the ribozyme in the crystalline environment in an effort to both validate our computational models and provide the foundation for investigating twister in solution. The simulated crystal was generated by applying the appropriate symmetry operations to the A conformation of the X-ray structure deposited as PDB ID: 4OJI,¹⁵ such that the unit cell contained 12 monomers arranged in the space group P 6₅ 2 2. The system was then solvated with TIP4P-Ew³² water, and equilibrated over 20 ns at constant temperature and pressure conditions to ensure that the system volume was within 0.3% of the experimental value. In a continued effort to replicate the experimental conditions, ammonium acetate and magnesium chloride were added to the system to match the experimental bulk concentrations of 0.1M and 0.02M, respectively. Following approximately 20 ns of equilibration, the simulation was carried out using the GPU accelerated version of PMEMD14^{33,34} at 280K and 1 atm using the AMBER FF14SB force field³⁵ with ion parameters³⁶ for use with the TIP4P-Ew water

model. The average monomer structure from this simulation was generated by first reversing the symmetry operations for each of the 12 monomers in the unit cell in each 10 ps snapshot, then averaging over the coordinates of all monomers as well as over all snapshots in the last 100 ns of the 120 ns trajectory. Analysis of the data was conducted using both the XtalAnalyze³⁰ module in AMBER14 and in-house scripts.

General solution phase simulation parameters

For each of the solution phase simulations, the ribozyme was solvated in a truncated octahedral box of TIP4P-Ew water with 12 Å between the solute and box, additionally including the crystallographically observed Mg²⁺ as well as 140 mM NaCl. All of the following solution phase simulations were carried out at 300K and 1 atm, using the AMBER FF14SB force field. The simulated annealing and solute equilibration procedure applied here has been used for other ribozymes and is detailed elsewhere.³⁷ Details on restraints applied and data collection specific to the various simulations completed is discussed below.

Free Energy of stacking U-1 beneath G33

Two free energy profiles were generated via umbrella sampling, by first minimizing representative structures of the "Stacked" state (Figure 3b or Figure S5b) while restraining the U-1:N3 - A34:N7 distance to 5.5 Å. This distance restraint was applied using a harmonic potential with a force constant of 50 kcal/mol-Å². A short equilibration step of 1 ns was performed under constant temperature and pressure conditions, while continuing to enforce the above restraint. The structures used for each of the umbrella sampling windows were spaced every 0.5 Å from 3.0 to 11.0 Å and were equilibrated in serial from the previously equilibrated, adjacent window (starting from the initial window at 5.5 Å). The distance distribution data for each of the 17 windows, used for generating the free energy profile, was gathered every picosecond from the last 10 ns of 15 ns trajectories. For the data collection stage of the simulations, the restraint weight was reduced to 10 kcal/mol-Å² to provide adequate overlap between adjacent sampling windows. One additional window at 3.75 Å was needed for the G33N1⁻ profile to improve sampling at the top of the barrier between the "Stacked" and "Triple" states. The free energy profiles were built using vFEP³⁸ and are presented in Figure 4. In order to estimate the statistical error of these calculations, 100 random subsamples of 10% of the data were collected. The 95% confidence intervals calculated from this distribution of subsamples are essentially the width of the lines in the plots presented in Figure 4.

Restrained MD exploring U-1 stacking states

During the equilibration stage (20 ns) restraints enforcing both the stacking state of U-1 and an "active" in-line conformation were maintained. The additional restraints supporting the "active" in-line conformation were slowly relaxed in the same fashion as described in the literature,³⁷ and then removed for the data production stage. Keeping the restraints on D_{stack} ensured that throughout the course of the simulations statistics could be gathered on a single stacking state of U-1, since the barriers are sufficiently small to allow for transitions between the states. The nucleophile attack angle (θ_{in}) defined as the angle between U-1:O2', A1:P, and A1:O5', and distance between the nucleophile and the phosphate (D_{in}).²¹ This data was collected every 10 ps over the course of each of seven 75 ns restrained simulations, and

clustered based on the hydrogen bonding patterning involving G33, the 2'-hydroxyl nucleophile, and the pro-R phosphate oxygen as seen in Figure 5 and Figure 6a.

Transition state mimic simulations

Initially, restraints enforcing the characteristic hydrogen bonding observed for each of the U-1 stacking states were applied. Once minimized, these structures were equilibrated, removing the restraints, using the same procedure as referenced previously and then simulated at constant temperature and pressure for 120 ns. Data was collected from the last 100 ns of the simulations.

Estimates of pH-activity profiles and apparent pK_a values

In order to determine the pH dependent probabilities of each of the four microstates in Figure 8, a series of thermodynamic integration calculations were completed. First, as a reference, the free energy of deprotonating the N3 position of adenosine and the N1 position of guanosine was calculated and assigned pK_a values of 1.5 ± 0.3^{22} and 9.4 ± 0.2^{39} respectively. The model nucleotides were simulated as methyl capped trinucleotides, including the residues flanking the proposed general acid and base in twister (e.g. A1 was modeled by the trinucleotide: $\text{CH}_3\text{-pU-pA-CH}_3$). Each model was simulated for a total of 5 ns (1 ns per value of λ), with the data being collected over the final 750 ps of each trajectory. The free energy for each leg of the thermodynamic cycle presented as Figure 8 was calculated for the full ribozyme in the U-1 "Stacked" conformation, from 500 ps simulations at each of nine values for λ . A set of restraints reinforcing the hydrogen bonding interactions between the A2-G33 base pair and from residues A1 to C16:OP1 and C17:OP2 were applied to maintain the integrity of the active site throughout the simulation, in addition to restraining D_{stack} to between 3.5 Å and 7.0 Å. These restraints allowed for a more stable calculation by restricting the sampling to just the conformational basin (U-1 Stacked) predicted to be catalytically active. The difference in free energy between the models and the ribozyme microstates was then used to calculate the shifted (at 300 K) microscopic pK_a values relative to the experimentally determined reference values for the nucleotides in solution:

$$pK_{a,\text{ribo}} = pK_{a,\text{model}} + (\Delta G_{\text{ribo}} - \Delta G_{\text{model}}) / (RT * \ln(10)) \quad (5)$$

From this data, the macroscopic "apparent pK_a " values could then be estimated, by fitting a two state apparent pK_a model²⁷ to the active fraction, $AH^+E_B^-$, as described in the text.

It should be noted that the thermodynamic cycle imposes the constraint that the legs of the cycle sum to zero, however, the free energy calculations performed here were done so independent of one another. As a result the calculated cycle did not close and instead summed to -0.53 ± 0.19 kcal/mol. In order to close the cycle, the residual energy (-0.53 kcal/mol) was distributed evenly across the four microstates and the apparent pK_a model was fit to the active fraction generated from the "closed cycle" pK_a values. The data for these free energy calculations and the resulting pK_a values used in the fitting procedure are presented in Table S1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Dr. David M.J. Lilley and Dr. Timothy J. Wilson for discussion. Additionally, the authors are grateful for financial support provided by the National Institutes of Health (GM62248 to DY). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575 (project number TG-MCB110101 to DY), and also supported in part by the Blue Waters sustained petascale computing project (NSF OCI 07-25070 and PRAC OCI-1036208).

References

- (1). Lilley DMJ. *Curr. Opin. Struct. Biol.* 2005; 15:313–323. [PubMed: 15919196]
- (2). Lönnberg T, Lönnberg H. *Current Opinion in Chemical Biology.* 2005; 9:665–673. [PubMed: 16233986]
- (3). Bevilacqua PC, Yajima R. *Curr. Opin. Chem. Biol.* 2006; 10:455–464. [PubMed: 16935552]
- (4). Strobel SA, Cochrane JC. *Curr. Opin. Chem. Biol.* 2007; 11:636–643. [PubMed: 17981494]
- (5). Fedor MJ. *Annu. Rev. Biophys.* 2009; 38:271–299. [PubMed: 19416070]
- (6). Ward WL, Plakos K, DeRose VJ. *Chem. Rev.* 2014; 114:4318–4342. [PubMed: 24730975]
- (7). Link KH, Breaker RR. *Gene Therapy.* 2009; 16:1189–1201. [PubMed: 19587710]
- (8). Chen X, Ellington AD. *PLoS Comput Biol.* 2009; 5:1000620.
- (9). Fastrez J. *ChemBioChem.* 2009; 10:2824–2835. [PubMed: 19937897]
- (10). Suess B, Weigand JE. *RNA Biol.* 2008; 5:1–6.
- (11). Ferré-D'Amaré AR. *Phil. Trans. R. Soc. B.* 2011; 366:2942–2948. [PubMed: 21930586]
- (12). Penchovsky R. *BiotechAdv.* 2014; 32:1015–1027.
- (13). Wilson TJ, Lilley DMJ. *Science.* 2009; 323:1436–1438. [PubMed: 19286542]
- (14). Chen X, Li N, Ellington AD. *Chem. Biodivers.* 2007; 4:633–655. [PubMed: 17443876]
- (15). Liu Y, Wilson TJ, McPhee SA, Lilley DMJ. *Nature Chem. Biol.* 2014; 10:739–744. [PubMed: 25038788]
- (16). Ren A, Košutić M, Rajashankar KR, Frener M, Santner T, Westhof E, Micura R, Patel DJ. *Nat. Commun.* 2014; 5:5534–5544. [PubMed: 25410397]
- (17). Eiler D, Wang J, Steitz TA. *Proc. Natl. Acad. Sci. USA.* 2014; 111:13028–13033. [PubMed: 25157168]
- (18). Roth A, Weinberg Z, Chen AG, Kim PB, Ames TD, Breaker RR. *Nature Chem. Biol.* 2014; 10:56–62. [PubMed: 24240507]
- (19). Doherty EA, Doudna JA. *Annu. Rev. Biophys. Biomol. Struct.* 2001; 30:457–475. [PubMed: 11441810]
- (20). Panteva, MT.; Dissanayake, T.; Chen, H.; Radak, BK.; Kuechler, ER.; Giambalvo, GM.; Lee, T-S.; York, DM. *Multiscale Methods for Computational RNA Enzymology.* Chen, S-J.; Burke-Aguero, DH., editors. Elsevier; 2015. Chapter 14
- (21). Soukup GA, Breaker RR. *RNA.* 1999; 5:1308–1325. [PubMed: 10573122]
- (22). Kapinos LE, Operschall BP, Larsen E, Sigel H. *Chem. Eur. J.* 2011; 17:8156–8164. [PubMed: 21626581]
- (23). Wilcox JL, Bevilacqua PC. *J. Am. Chem. Soc.* 2013; 135:7390–7393. [PubMed: 23432144]
- (24). Viladoms J, Scott LG, Fedor MJ. *J. Am. Chem. Soc.* 2011; 133:18388–18396. [PubMed: 21936556]
- (25). Goh GB, Knight JL, C. L. B. *J. Chem. Theory Comput.* 2013; 9:935–943. [PubMed: 23525495]
- (26). Dissanayake T, Swails JM, Harris ME, Roitberg AE, York DM. *Biochemistry.* 2015; 54:1307–1313. [PubMed: 25615525]
- (27). Bevilacqua PC. *Biochemistry.* 2003; 42:2259–2265. [PubMed: 12600192]

- (28). York DM, Yang W, Lee H, Darden T, Pedersen LG. J. Am. Chem. Soc. 1995; 117:5001–5002.
- (29). Martick M, Lee T-S, York DM, Scott WG. Chem. Biol. 2008; 15:332–342. [PubMed: 18420140]
- (30). Janowski PA, Cerutti DS, Holton J, Case DA. J. Am. Chem. Soc. 2013; 135:7938–7948. [PubMed: 23631449]
- (31). Heldenbrand H, Janowski PA, Giamba u G, Giese TJ, Wedekind JE, York DM. J. Am. Chem. Soc. 2014; 136:7789–7792. [PubMed: 24842535]
- (32). Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T. J. Chem. Phys. 2004; 120:9665–9678. [PubMed: 15267980]
- (33). Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. J. Chem. Theory Comput. 2013; 9:3878–3888. [PubMed: 26592383]
- (34). Case, DA.; Babin, V.; Berryman, JT.; Betz, RM.; Cai, Q.; Cerutti, D.; Cheatham, TE., III; Darden, TA.; Duke, RE.; Gohlke, H.; Götz, AW.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T.; Le Grand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, KM.; Paesani, F.; Roe, DR.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, CL.; Smith, W.; Swails, J.; Walker, RC.; Wang, J.; Wolf, RM.; Wu, X.; Kollman, PA. AMBER 14. University of California, San Francisco; San Francisco, CA: 2014.
- (35). Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. J. Chem. Theory Comput. 2015; 11:3696–3713. [PubMed: 26574453]
- (36). Joung IS, Cheatham TE III. J. Phys. Chem. B. 2008; 112:9020–9041. [PubMed: 18593145]
- (37). Lee T-S, Giamba u GM, Sosa CP, Martick M, Scott WG, York DM. J. Mol. Biol. 2009; 388:195–206. [PubMed: 19265710]
- (38). Lee T-S, Radak BK, Pabis A, York DM. J. Chem. Theory Comput. 2013; 9:153–164. [PubMed: 23457427]
- (39). Verdolino V, Cammi R, Munk BH, Schlegel HB. J. Phys. Chem. B. 2008; 112:16860–16873. [PubMed: 19049279]

Figure 1.

Overlay of experimental (PDB: 4OJI, grey) and average crystal simulation (colored) active sites. The positional, heavy atom RMSD is 0.78 Å for the full monomer and 0.44 Å for the active site (U-1, A1 and G33) alone.

Figure 2.

Comparison of simulation and experimental B-factors calculated for each residue (linear correlation coefficient: 0.90; standard error: 10.68 Å²). Experimental B-factors for residues 17 and 18 were not reported as this region was interpreted to be highly disordered. Residue numbers along the x-axis correspond to the order of residues found in the PDB (Figure S1a) and map to the general numbering scheme for twister (Figure S1b).

Figure 3.

Average structures from the restrained simulations exploring the conformational states of U-1 (light blue) are labeled as follows: (a) Extruded, (b) Stacked, and (c) Triple. The characteristic hydrogen bonding between G33 in its neutral, protonated form (G33N1H) and either the scissile phosphate or the nucleophile is shown in dark blue.

Figure 4.

The free energy profile along the U-1 stacking coordinate defined by the distance between U-1:N3 and A34:N7 (D_{stack}), for G33 both deprotonated at the N1 position (G33N1⁻ in black) and neutral (G33N1H in blue). The shaded region with $D_{stack} < 3.75$ Å corresponds to the "Triple" state, while the shaded region with $D_{stack} > 7$ Å defines the "Extruded" state. The "Stacked" state for U-1 is left unshaded with D_{stack} between 3.75 and 7.0 Å.

Figure 5.

First Row: Nucleophile attack angle versus nucleophile-phosphate distance with G33 neutral (G33N1H) and U-1 in each of the three conformational states (Extruded, Stacked, Triple). Second Row: Schematic highlighting hydrogen bonding used for clustering analysis, G33 deprotonated (G33N1⁻) with U-1 in the stacked state, and the average structure of the active site for the G33N1⁻ "Stacked" simulation. The black outlined region ($\theta_{inl} > 140^\circ$ and $D_{inl} < 3.5$ Å) indicates active "in-line" conformations that favor catalysis.

Figure 6.

Restrained simulation of presumed active state (a) nucleophile attack angle versus nucleophile-phosphate distance clustered based on hydrogen bonding interactions with G33. The black outlined region ($\theta_{inl} > 140^\circ$ and $D_{inl} < 3.5$ Å) indicates active "in-line" conformations that favor catalysis. (b) Distribution of distances between A1:N3 and A1:O5', (c) Average structure of the frames where the A1:N3 - A1:O5' distance is greater than 3.5 Å. (d) Average structure for the cluster where the presumptive general acid is directly hydrogen bonding to the leaving group. The two clusters were generated using the criteria that a heavy atom distance of than 3.5 Å is considered as the cutoff for a hydrogen bond.

Figure 7.

Average structure of the proposed dianionic phosphorane transition state mimic along with the characteristic hydrogen bonding network. With U-1 in the Stacked state, the presumptive base (G33) donates dual hydrogen bonds to the pro-R non-bridging oxygen, while N3 of A1 (dark blue) is seen hydrogen bonding with the leaving group.

Figure 8.

Microscopic general acid-base protonation state model:²⁶ The four microstates for the ribozyme (labeled E) are ${}_{AH^+}\mathbf{E}_{B^-}$ (the presumed active state), ${}_{AH^+}\mathbf{E}_{BH}$, ${}_A\mathbf{E}_{BH}$, and ${}_A\mathbf{E}_{B^-}$. The subscripts A and B indicate the pK_a was calculated for the acid or base, respectively, given the protonation state of the residue noted by the superscript.

Figure 9.

Comparison of the simulated and experimental pH-activity profiles and apparent pK_a values. The predicted pH-activity profile (red) is scaled such that the peak maximum corresponds to that of a two state apparent pK_a model fit to the experimental data¹⁵ (black). The derived apparent pK_a values for the acid and base are highlighted by the dotted lines for both the calculated (6.5 ± 0.4 and 9.0 ± 0.4) and experimentally derived (6.9 and 9.5) curves. The experimental rates shown here are the mean of at least three independent experiments, while the error bars are twice the standard deviation. For the calculated curve, the error associated with the predicted pK_a values (± 0.4 pK_a units) is derived primarily from the error in the experimental reference data used rather than from the free energy calculations or fitting procedure.