Langmuir

Article

pubs.acs.org/Langmuir

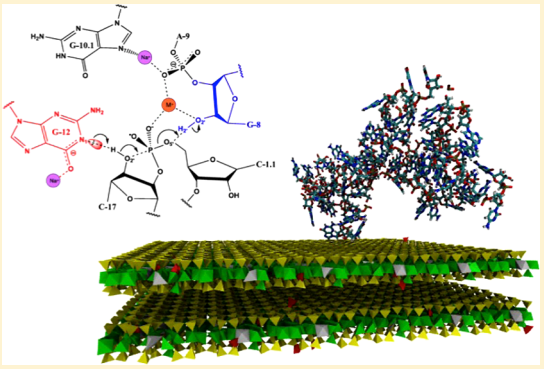
# Structure, Dynamics, and Function of the Hammerhead Ribozyme in Bulk Water and at a Clay Mineral Surface from Replica Exchange Molecular Dynamics

Jacob B. Swadling, David W. Wright,† James L. Suter, and Peter V. Coveney*

Centre for Computational Science, Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, United Kingdom

Supporting Information

**ABSTRACT:** Compared with proteins, the relationship between structure, dynamics, and function of RNA enzymes (known as ribozymes) is far less well understood, despite the fact that ribozymes are found in many organisms and are often conceived as "molecular fossils" of the first self-replicating molecules to have arisen on Earth. To investigate how ribozymal function is governed by structure and dynamics, we study the full hammerhead ribozyme in bulk water and in an aqueous clay mineral environment by computer simulation using replica-exchange molecular dynamics. Through extensive sampling of the major conformational states of the hammerhead ribozyme, we are able to show that the hammerhead manifests a free-energy landscape reminiscent of that which is well known in proteins, exhibiting a "funnel" topology that guides the ribozyme into its globally most stable conformation. The active-site geometry is found to be closely correlated to the tertiary structure of the ribozyme, thereby reconciling conflicts between previously proposed mechanisms for the self-scission of the hammerhead. The conformational analysis also accounts for the differences reported experimentally in the catalytic activity of the hammerhead ribozyme, which is reduced when interacting with clay minerals as compared with bulk water.





## INTRODUCTION

Ribonucleic acid (RNA) was first suggested to be capable of exhibiting enzymatic activity by Woese, Crick, and Orgel in 1967 after the discovery that RNA could fold to form secondary structures in an analogous way to protein enzymes.[1−3] It was not until the 1980s that Cech and Altman independently discovered the first ribozymes, for which they were awarded the 1989 Nobel Prize in Chemistry.[4] It is possible that ribozymes played a key role in the origin of life on Earth had the critical event been the emergence of RNA molecules capable of replicating a primordial RNA "genome". A world based entirely on RNA would have required the presence of an environmental chaperone in which RNA could polymerize and become biologically active. 'Indeed, the synthesis of precursors of nucleotides in the presence of clay minerals[5] and their polymerization into oligonucleotides up to the length of a small ribozyme, such as the hammerhead ribozyme, have been observed.[6] RNA monomers and polymers adsorbed and tethered on clay mineral surfaces are able to persist in the presence of degrading agents and extreme temperature and pressure conditions to interact with surrounding molecules enabling the copying of genetic information.[7]

Simple ribozymes can help us understand concepts in modern biology; for example, the ribosome translates RNA into proteins in the cell via a conserved catalytic center composed entirely of RNA. The hammerhead ribozyme is a small self-cleaving RNA that consists of a core of invariant residues. For optimal activity it requires the presence of the full tertiary structure, as shown by the reduced activity of the minimal hammerhead ribozyme.[8] Because of its size and the amount of experimental effort dedicated to it, the hammerhead ribozyme thus makes an excellent prototype for understanding the role of structure and dynamics as a function of the wider class of ribozymes, of which we know very little compared with the analogous protein structure−dynamics−function relationship.[9,10]

The hammerhead crystal structure, determined by Martick et al. and shown in Figure 1, was derived under nonphysiological conditions but provides a structural basis that lends itself well to further theoretical modeling of the structure and mechanism.[11] The hammerhead ribozyme catalyzes an RNA self-cleavage phosphodiester isomerization reaction that involves nucleophilic attack of a cytosine 2′ oxygen upon the adjacent scissile phosphate, yielding two RNA product strands.

Here we use fully atomistic replica-exchange molecular dynamics (REMD) computer simulation to exhibit the delicate relationship between tertiary structure and catalytic activity in ribozyme molecules. REMD is well-suited for the study of

**Received:** September 15, 2014
**Revised:** January 10, 2015
**Published:** February 3, 2015

DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501

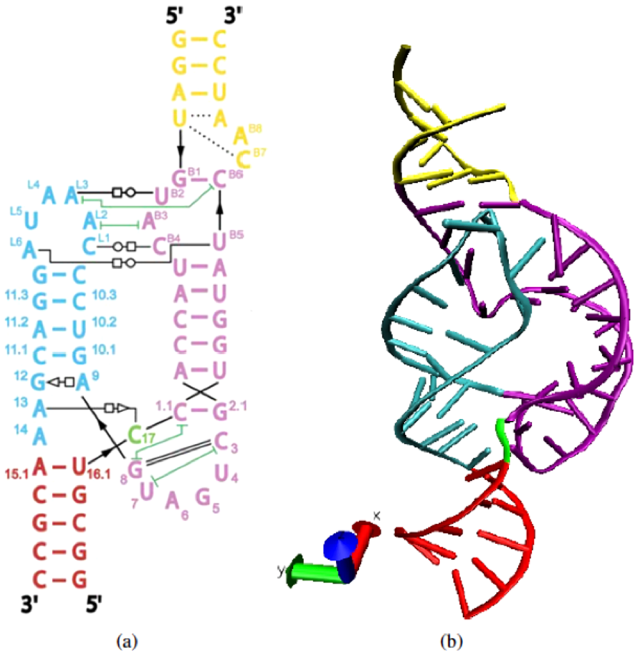



**Figure 1.** (a) Schematic representation of the secondary structure of the full-length hammerhead ribozyme and (b) the 3D representation taken from the crystal structure. Annotations in panel (a) follow the scheme developed by Yang et al.[42] Canonical base pairing is shown as colored lines and hydrogen base pairing is shown as black lines. Open circle next to open square, Watson–Crick/Hoogsteen; open square next to open triangle, Hoogsteen/sugar edge; dashed line, H bond; green line, nonadjacent base stacking. The thick black line shows the direction and continuity of the backbone. The color scheme depicts Stem I in yellow and purple, Stem II in blue, Stem III in red, and the scissile phosphate group in green.[29]

complex biomolecules, like RNA and proteins, because of the greatly enhanced sampling of conformational space it delivers over single trajectory molecular dynamics (MD). As we show here, REMD allows the identification of the major conformational states of the hammerhead ribozyme and the native state (and near-native folding intermediates). We use the atomic resolution gained through simulation to resolve conflicting theories about how metal cations participate in the active site conformation. We study the hammerhead both in bulk water and at an aqueous clay mineral surface. The latter is known to perturb the rate of catalytic activity of the hammerhead,[12] which allows us to answer questions about how the local environment and ribozyme conformation alter its function. The clay-mineral-bound ribozyme is also implicated in various origins of life scenarios according to which clay minerals are thought to have concentrated and protected the first catalytic RNA molecules at the inception of the "RNA world".[13,14]

## MATERIALS AND METHODS

The starting structure of the hammerhead ribozyme was modified from the 2.2 Å resolution crystal structure of the full-length catalytically active hammerhead ribozyme (PDB 2GOZ) by adding hydrogen atoms and replacing $Mn^{2+}$ sites with $Mg^{2+}$; see Figure 1. The full-length hammerhead ribozyme consists of single-stranded RNA sequences 43 and 20 nucleotides in length. The chains form three base-paired helix-like stems. 20 Å of SPC/E water and 50 charge-balancing $Na^{+}$ ions were placed around the ribozyme using the *solvate* function in AmberTools. This is the same procedure that we used in ref 15. There we showed that the nature of the bulk counterions can play a very important role in determining aspects of the interaction between RNA and clay surface; in particular, there is a substantial difference depending on whether the bulk cation is monovalent (as here) or divalent. Crystal water molecules were retained along with the divalent metal ion sites that were populated with 5 $Mg^{2+}$ ions. The complete model consisted of 54 10 atoms with unit cell dimensions of $83.6 \times 73.8 \times 95.4$ Å$^3$.

The clay mineral used in this study was derived from the chemical formula $Na_3[Al_{14}Mg_2][Si_{31}Al]O_{80}OH_{16}$, consistent with naturally occurring Wyoming montmorillonite. Partial substitutions were created in both tetrahedral and octahedral sheets. Montmorillonite was chosen because it is widely used in experimental work.[6,7,12]

The bonding interactions and partial charges for the RNA ribozyme molecule were described using the Amber ff99 force-field parameters. The Barcelona bsc0 refinement was used to furnish an improved description of the alpha\gamma concerted rotations within the RNA molecule.[16] The ClayFF force field was used to provide parameters for atoms belonging to the inorganic montmorillonite clay.[17] ClayFF produces very good agreement with experiment in terms of lattice parameters, water diffusion coefficients, far-infrared spectra, and elastic properties. Indeed, we have used this parametrization successfully to describe clay–nucleic acid systems at higher temperatures and pressures in the past,[14] so we have not needed to constrain the clay surface in any way. We have previously demonstrated that the combined organic-clay force field, which contains cross-terms determined using Lorentz–Berthelot rules, provides an accurate description of clay–nucleic acid systems.[18] The SHAKE algorithm was employed to constrain bonds and angles within hydroxyl groups and water.[19] This allowed the use of a 2 fs time step within the MD algorithm.

All simulations reported here were run using the large atomistic/molecular massively parallel simulator (LAMMPS) code developed by Plimpton.[20] The code exhibits a near-linear scaling relationship between the number of cores used and the speed-up in wall-clock time. In our case, we used 768 cores per replica and a total of 100 replicas.

Prior to commencing production simulations, the starting structure was energy-minimized using the steepest descent method, then thermalized with MD from 0 K, 0 atm to 300 K, 1 atm. MD was used to perform time integration on Nosé–Hoover-style non-Hamiltonian equations of motion, designed to generate positions and velocities sampled from the isothermal–isobaric ensemble.

The high level of conformational sampling in this study was achieved using the REMD algorithm.[21] The REMD method has been widely used in the computational study of protein folding[22] and also the analogous process of RNA folding.[23−26] REMD consists of $M$ noninteracting copies (or, replicas) of the original system in the canonical ensemble at $M$ different temperatures $T_m$ ($m = 0, ..., M - 1$). The replicas are arranged so that there is always one replica at each temperature. The trajectory of each independent replica is computed using MD. Adjacent replicas (replicas $i$ and $i + 1$) exchange temperatures according to a Boltzmann probability distribution. REMD essentially runs $N$ copies of the system, randomly initialized, at different temperatures. Then, on the basis of the Metropolis criterion, configurations are exchanged at different temperatures.[21] The idea of this method is to make configurations at higher temperatures available to the simulations at lower temperatures and vice versa. This results in a very robust ensemble that is able to sample both low- and high-energy configurations. REMD produces enhanced sampling over single trajectory MD because fixed-temperature conformations are much more easily trapped in local energy minima.

We have selected a temperature spacing so that $P_{ex}$ (the probability of exchange) is ∼20–30% for all replicas.[27] (See Figure S1 in the Supporting Information for a plot of the energy overlap.) REMD enhances the sampling of configurational space and thus provides an effective way of obtaining equilibrium and dynamical properties of a system described by a complicated energy landscape.[21]

We analyzed replicas of the hammerhead ribozyme system from temperature replicas under ambient conditions: 300.00, 301.96, 303.93, 305.91, and 307.90 K. Each replica was simulated for 50 ns, which equates to 5 $\mu$s of fully atomistic MD simulation for the entire




DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501

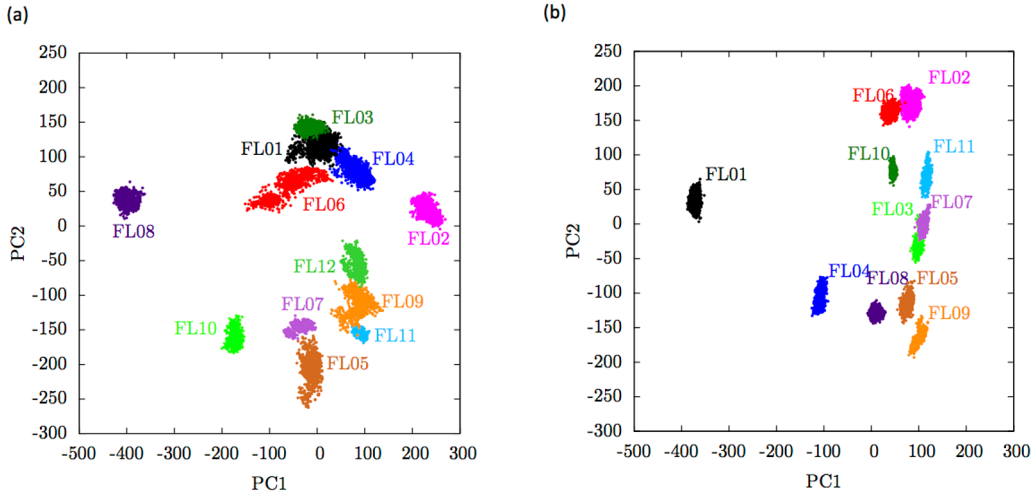



**Figure 2.** Principal component analysis showing the clustering of conformations as a function of the two most dominant modes of motion for the hammerhead ribozyme in bulk water (a) and while interacting with a mineral surface (b). The first dominant motion (PC1) corresponds to a helical junction bend while the second (PC2) corresponds to a spring-type helical compression. Representative structures for each cluster island can be viewed in Figures S4 and S6 of the Supporting Information for the ribozyme in bulk water and for the clay bound ribozyme, respectively.

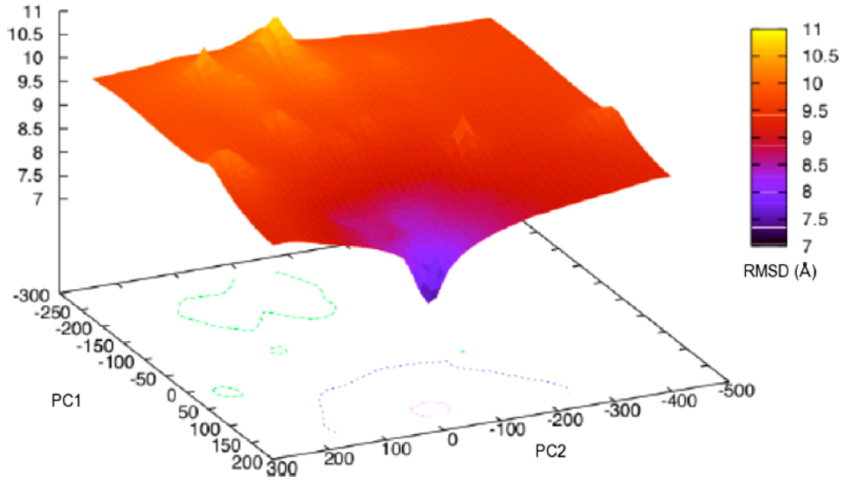



**Figure 3.** First two principal components for the hammerhead ribozyme are used to show the distinct cluster islands plotted against an RMSD surface relative to the crystal structure in (a) bulk water and (b) the clay mineral environment. The surface shows one distinct global minimum of which the structure with the lowest RMSD from the crystal structure is positioned 1.33 Å from the center of cluster FL03. It can be deduced from this that cluster FL03 is representative of the crystal structure.

ensemble of replicas following the equilibration and convergence of the model systems. The equations of motion were integrated with a time step of 2 fs, and periodic boundaries were imposed in all three spatial directions. Electrostatic interactions were computed using the particle−particle/particle−mesh ($P^3M$) method. Convergence was monitored by computing block averages in the change of free energy and is shown in Figure S2 in Supporting Information as well as monitoring the round-trip times of replicas, that is, the average time it takes for a replica to sample the highest temperature and return to the original temperature. A clustering process was applied to the accumulated REMD trajectory data to group mutual similarity and reduce the complexity of the data and to identify a set of conformationally distinct clusters in terms of which the REMD sampling can be understood (Figure 2). We employed *k*-means clustering based on the Cartesian coordinate root-mean-squared deviation (RMSD) between structures simulated from 300.00 to 307.90 K with a fixed radius of 4 Å. (See the section entitled "Structural Clustering Algorithm" in the Supporting Information for further details on the clustering process.)

## RESULTS

**Global Structural Differences Observed in Aqueous and Clay Environments.** Through REMD we are able to sample the major conformational states of the hammerhead ribozyme both in bulk water and while interacting with clay mineral surfaces. Figure 2 depicts the two most dominant modes of motion of the hammerhead ribozyme in both clay mineral and aqueous systems, contributing 60% of the total dynamics of the molecule determined from matrix eigenvalues. In both bulk water and clay mineral systems the first principal component of motion corresponds to a helical junction bend and the second corresponds to a spring-type helical compression.[28] Although the principal motions of the hammerhead ribozyme are the same in both bulk water and in a confined mineral environment, the conformational landscape and thus the populations of occupied conformational states are significantly perturbed in the latter. The differences between the two systems (bulk water versus clay mineral) in terms of nucleic acid structural metrics can be seen by comparing Tables





DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501

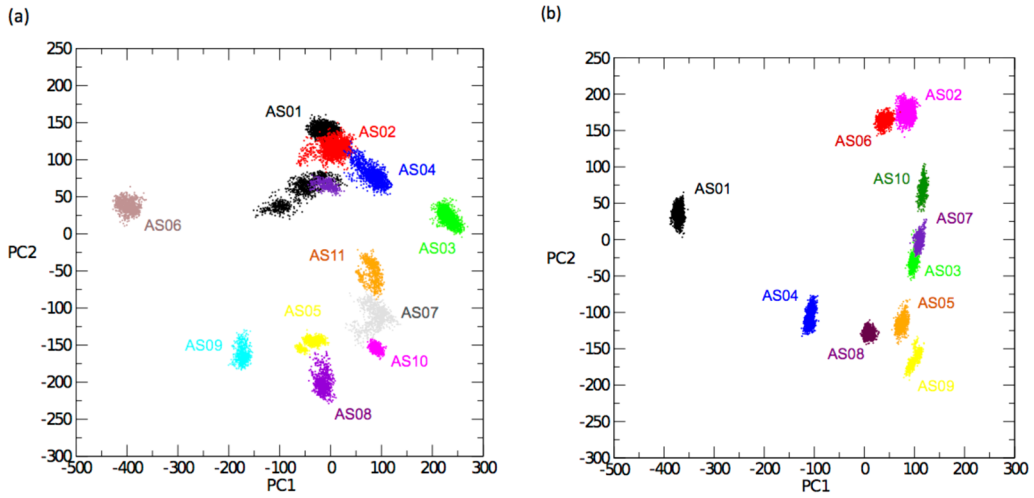



**Figure 4.** Clustering of the active-site conformations as a function of principal components 1 and 2 of the full hammerhead ribozyme structure (a) in bulk water and (b) with a clay mineral surface. Each structure is colored according to the clustering of the active site (not the full tertiary structure). The plot shows an excellent correlation between the principal motions of the full tertiary structure and the local conformation of the hammerhead ribozyme active site that is composed of six residues (compare with Figure 2). This suggests that the active site structure and therefore the ribozyme mechanism is dependent on the full tertiary structure.



S1 and S2 in the Supporting Information. The differences observed in the major conformational states of the hammerhead in the two different environments provide us with an initial indication as to why the catalytic activity of the ribozyme is reduced when adsorbed on the clay surface compared with that in bulk water.

The application of the *k*-means clustering algorithm successfully identified structurally and conformationally distinct clusters within the principal component space. Figure 2 displays only those cluster islands that have a population over a given threshold (of 500 structures). A histogram of population clusters and the imposed cutoff is shown in the Supporting Information in Figure S3. A visualization of each cluster "best member" can be seen in Figures S4 and S6 of the Supporting Information. It is positioned in principal component space at a location corresponding to a significant bend, indicating a propensity for the ribozyme to adopt a more compact active site structure in the confined clay mineral environment compared with when in bulk aqueous solution. The most populated cluster in each system, given by the label FL01 in Figure 2, corresponds to the conformational state with the lowest free energy that is, by definition, the native fold. The free energy is related to the state population (cluster size) through the equation

$$\Delta G = -RT \ln(P_n/P_0)$$

where $P_n$ is the fraction of native structures, $P_0$ the fraction of populated state, $R$ is the gas constant, and $T$ is the temperature. In bulk water, the structure with the most "crystal-structure-like" conformation with a RMSD from it of 5.8 Å belongs to cluster FL03 at the bottom of the free-energy "funnel" in Figure 3. The native conformational state for the clay-adsorbed ribozyme has an RMSD of 11.3 Å from the crystal structure, indicating that the folding pathways funnel to a native state unlike the crystal structure.

Biondi et al. assessed the ability of the hammerhead ribozyme to catalyze reactions in a clay mineral environment by investigating its self-cleavage reaction.[12] Their experimental results indicate that the hammerhead ribozyme is still active when adsorbed on a clay surface, even though its efficiency is reduced to ∼20% of that in solution. Here we describe the differences in the structure, dynamics, and function of the hammerhead ribozyme in the two environments: in bulk water and at a clay surface.

**Analysis of Catalytic Competency.** The precise role, if any, that $Mg^{2+}$ ions and solvent components play in the catalytic phosphodiester bond cleavage reaction mechanism of the hammerhead ribozyme has hitherto remained unknown. The first reported crystal structure of the full-length hammerhead ribozyme[29] showed that the scissile phosphate is positioned 4.3 Å from the A9 phosphate. This led Martick and Scott to suggest that there is a likely electrostatic requirement for cation occupation to screen the repulsion between these phosphates even prior to the formation of the transition state.[29] Although no actual bound $Mg^{2+}$ ions were observed in the original full-length crystal structure, this structure was observed under nonphysiological high monovalent salt (1 M $NH_4^+$) conditions. Under physiological conditions, a divalent metal ion may be needed to bind and bridge closely approaching phosphates in the active site area. It is also possible that the divalent ion is not bound but rather opportunistically resides in the active site. The most recent crystal structure of the full-length hammerhead ribozyme, at 2.2 Å resolution, was reported by Martick et al.[11] using crystals prepared in the presence of 10 mM $Mn^{2+}$. Martick et al.'s work indicated that a divalent ion participates indirectly, or not at all, in the catalytic mechanism within the active site. The divalent ion most likely withdraws some of the accumulating negative charge of the scissile phosphate and may introduce order to the solvent molecules near the cleavage site. Martick et al.'s argument for this mechanism is that it does not require the repositioning of any ions or atoms in the crystal structure. Anderson et al. resolved the structure of the hammerhead ribozyme to 1.55 Å under conditions that permit observation of $Na^+$ positions.[30] The geometry led them to suggest that $Na^+$ directly substitutes for divalent cations in the active site.

The native structure for the hammerhead ribozyme in bulk water reveals an active site catalytic structure that coordinates around a single $Mg^{2+}$ ion. (See Figure S5 in the Supporting Information.) This structure supports the direct single metal


DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501

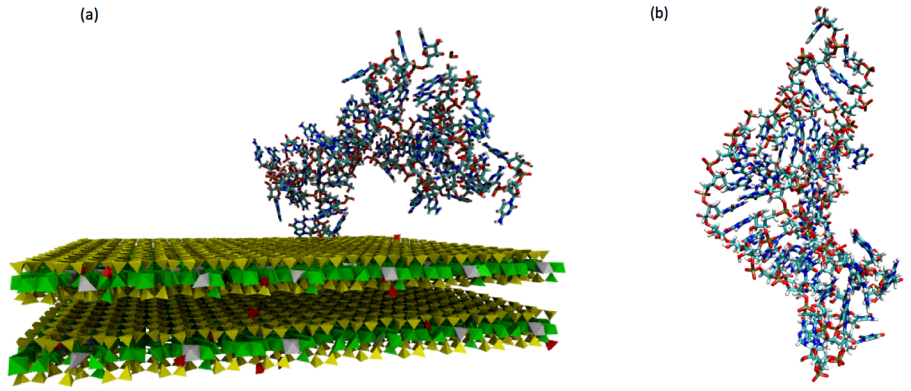



**Figure 5.** Simulation snapshot of the hammerhead ribozyme from 50 ns of replica exchange molecular dynamics (a) while interacting with a clay surface and (b) in bulk water. The surface-bound ribozyme shows a significantly altered conformation compared with the ribozyme in bulk water. The ribozyme is tethered to the mineral surface via a terminal planar base.

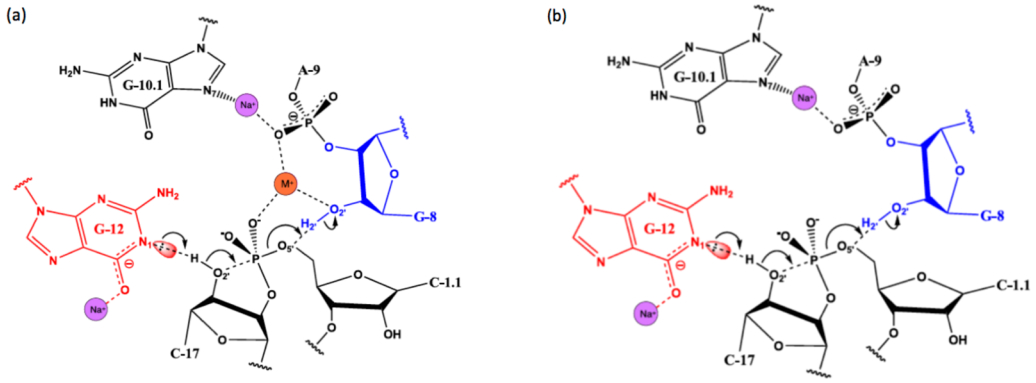



**Figure 6.** Reaction structure and mechanism schematics of the catalytically active center of the hammerhead ribozyme in bulk water (a) and when the ribozyme is adsorbed at the clay mineral surface (b) based on the lowest free energy "native" structures taken from the REMD simulations reported here.

catalytic reaction mechanism proposed as an alternative to Scott's indirect/no metal ion mechanism. In the direct metal ion mechanism observed in our simulations, an $Mg^{2+}$ ion migrates from the crystallographic site to bridge the A9 and scissile phosphate to screen the negative electrostatic charge that accumulates as the transition state is formed. This mechanism was previously suggested by Wang et al.[31] in their experimental studies of the hammerhead using $Cd^{2+}$ ions. Wang et al. proposed that the hammerhead undergoes a substantial conformational rearrangement in aqueous solution to attain its catalytic conformation. Such rearrangements appear to be general features of small functional RNAs.[32] The single-trajectory MD simulations performed on the hammerhead ribozyme crystal structure by Martick and Scott[29] employed a harmonic B-factor restraint on the RNA and bound metal ions; it is therefore hardly surprising that they do not observe a deviation from the crystal structure including the migration of divalent metal ions. By contrast, our REMD simulations start from the crystal structure and bound ion positions but do not restrain any heavy atoms, so we are able to observe transitions in metal ion sites.

Figure 3 shows the first two principal components plotted against the RMSD relative to the crystal structure for the ribozyme in bulk water and in a clay mineral environment. The surfaces both exhibit single distinct minima of which the structure with the lowest RMSD from the crystal structure is located. It can be deduced from this that cluster FL03 in bulk water and FL01 in the clay mineral case are representative of the crystal structure. Both of the surfaces show a funnelling-type landscape that ends at the native structure, a feature often described in protein folding[9,10] but rarely reported in the analogous RNA folding schemes.

From the clustering of the active site structure, we were able to identify a one-to-one mapping between the global tertiary structure and the six residue active site structure. (See Figure 4.) Through our sampling of the conformational landscape of the ribozyme molecule, we were able to identify the native structure, which is the conformation with the lowest free energy. Figure 4 displays principal components 1 and 2 of the full hammerhead ribozyme structure but where each structure is colored according to the clustering of the six residues that compose the active site (not the full tertiary structure). Comparison of this plot and Figure 2 shows an excellent correlation between the principal motions of the full tertiary structure and the local conformation of the hammerhead ribozyme active site, as the principal components are taken from the full structure while cluster analysis was performed on the six residues belonging to the active site. This suggests that the active-site structure, and therefore ribozyme mechanism, is dependent on the full tertiary structure for both the bulk water and clay-mineral-bound ribozyme.

A snapshot from the simulation of the hammerhead ribozyme at an aqueous clay mineral surface is shown in Figure 5. The influence of this montmorillonite surface and resulting perturbation on the conformational landscape has a significant effect on the overall ribozyme structure and the active-site





DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501

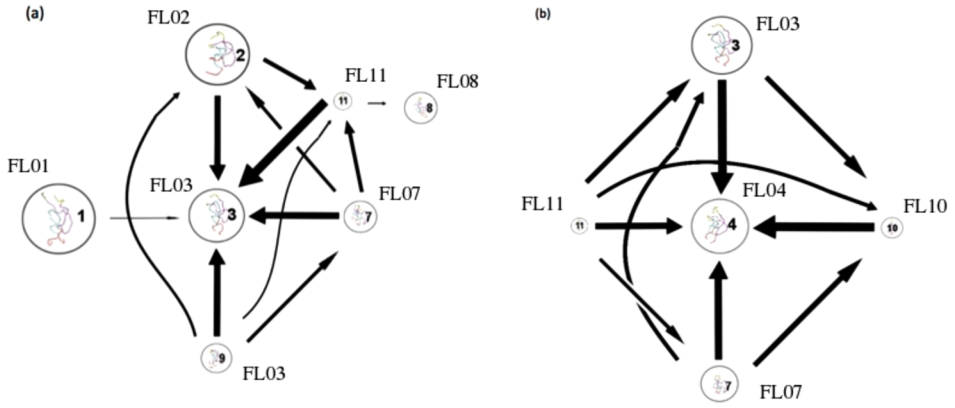



**Figure 7.** Superposition of the largest folding fluxes between Markov states for the aqueous-based ribozyme (a) and the clay-based ribozyme (b). Each state is scaled in terms of its relative free energy, and each arrow is scaled proportional to the interstate flux. These fluxes account for 40% of the total flux. The macrostates chosen for the coarse grained Markov state model were selected to correspond to the structural cluster islands given in Figure 2. Cluster labels are given next to each macrostate. For a fine-grained interpretation of the corresponding structures, see Figures S4 and S6 in the Supporting Information for "best member" structures taken from the two systems.

**Table 1. Six Most Prevalent Hydrogen Bonds for the Hammerhead Ribozyme in Bulk Water**[a]

| donor | acceptor H | acceptor | % occupied | distance (Å) | angle (deg) | lifetime (ps) |
|---|---|---|---|---|---|---|
| C3:N3 | G19:H1 | A9:N1 | 97.47 | 2.874(0.09) | 15.97(8.84) | 69.8(160.2) |
| G19:06 | C3:H41 | C3:N4 | 97.25 | 2.856(0.14) | 16.96(10.22) | 60.4(107.0) |
| C3:02 | G19:H21 | A9:N2 | 97.21 | 2.799(0.12) | 16.80(9.56) | 61.0(106.8) |
| G2.1:06 | C1.1:H41 | C1.1:N4 | 95.31 | 2.876(0.16) | 17.97(10.78) | 29.9(40.7) |
| Cl.l:02 | G2.1:H21 | G2.1:N2 | 91.62 | 2.805(0.13) | 15.62(9.04) | 18.9(44.6) |
| C1.1:N3 | G2.1:H1 | G2.1:N1 | 91.29 | 2.888(0.10) | 17.06(10.05) | 18.1(39.3) |

[a]Each hydrogen bond is given in terms of the donor and acceptor atoms, percentage of the trajectory that the bond is occupied for, distance, angle, and average lifetime of the hydrogen bond in picoseconds. Criteria for hydrogen bonding is a distance of 3.5 Å between donor and acceptor and an angle of 120° between the donor and acceptor hydrogen and heavy atom.

structure and as such the catalytic rate of bond cleavage. (See Figure S8 of the Supporting Information.) Biondi et al.'s studies showed that the hammerhead ribozyme was still active when adsorbed on the clay surface, even though its efficiency was reduced to ∼20% of that in bulk aqueous solution. The native structure observed in our clay-surface−ribozyme simulations manifests a more compact structure than that in bulk water; as a result, the active site undergoes a significant structural rearrangement, whereby the central coordinating $Mg^{2+}$ ion vacates the active site pocket, as shown in Figure 6. The highly ionic atmosphere associated with the charged mineral surface effectively screens the electrostatic repulsion arising from the A9 and scissile phosphate groups of the ribozyme, yielding a more compact active-site structure indicative of the "no metal ion participation" mechanism described by Martick and Scott.[29]

**Modeling Long-Time Scale Kinetics.** To describe the kinetic and mechanistic aspects of conversion between conformational states, we employ Markov state modeling (MSM). MSM works by identifying a set of kinetically metastable states from generalized ensemble simulations as well as other simulation datasets and efficiently sampling transitions between these states. This allows modeling of long-time-scale kinetics from simulations of much shorter duration.[33−37]

The REMD hammerhead conformational data is clustered based on an all-atom RMSD metric that employs hybrid k-centers and k-medoids. K-centers clustering is applied iteratively until the intercluster distance reaches 4 Å. At that point, 100 iterations of hybrid k-medoids are performed to refine those clusters. We calculated the relaxation time scales for a sequence of lagtimes, which indicated that relaxation time scales begin to flatten out at ∼300 ps. (See Figure S2 in Supporting Information.) Thus, a Markov state model was created using a lagtime of 300 ps. Spectral clustering methods such as robust Perron cluster analysis (PCCA+) can be used to construct metastable models with a minimal number of states. First, we construct a microstate model with a short lagtime. The short lagtime is necessary because PCCA+ tries to create macrostates that are long-lived, termed metastable. At long lagtimes, states become less and less metastable. Detailed examination of the relaxation time scales indicated that there were 12 kinetic processes, so we chose to build a model with 12 macroscopic-states.

To compute transition pathways, we selected macrostates identified during the MSM building method as subsets of state space. The pathways with the highest flux were calculated using transition path theory (TPT) and are visualized in Figure 7.[38] The major states that we sampled using REMD span a large domain within conformational space, and as such there are many transitions between the vast number of conformations. In an analogous way to complex protein folding schemes, we find that the hammerhead ribozyme exhibits so-called "downhill" folding, or "funneling".[9]

The increased rate of folding we see between conformational states of the hammerhead ribozyme when it is in the clay mineral environment is analogous to that we found previously for smaller RNA strands (of 25 nucleotides).[15] The ionic atmosphere associated with the charged clay surface screens the electrostatic repulsion felt between phosphate groups and encourages a hydrophobic collapse where base groups become




DOI: 10.1021/la503685t
Langmuir 2015, 31, 2493−2501

Table 2. Six Most Prevalent Hydrogen Bonds Taken from the Hammerhead Ribozyme at an Aqueous Montmorillonite Surface

| donor | acceptor H | acceptor | % occupied | distance (Å) | angle (deg) | lifetime (ps) |
|---|---|---|---|---|---|---|
| AI3:05′ | AI4:H0′2 | AI4:0′2 | 68.23 | 2.932(0.20) | 38.16(13.44) | 3.7(3.6) |
| AI2:05′ | CI1:H0′2 | Cll:0′2 | 61.23 | 2.871(0.19) | 31.30(14.08) | 3.7(3.7) |
| G46:02P | C39:H61 | C39:H6 | 52.60 | 2.858(0.18) | 18.25(11.23) | 3.3(3.5) |
| U50:O2 | A6:H21 | A6:N2 | 49.93 | 2.850(0.18) | 26.13(13.60) | 3.2(3.3) |
| C17:02P | C1.1:H42 | C1.1:N4 | 48.28 | 2.788(0.15) | 15.17(8.52) | 3.2(3.2) |
| C17:05′ | U16.1:H0′2 | U16.1:02′ | 43.10 | 2.879(0.21) | 30.56(14.91) | 2.5(2.4) |

buried within hydrophobic pockets. Figure 7 displays scaled arrows that correspond to the relative fluxes between the conformational states. It follows that many of the pathways between native and non-native states pass through a small number of intermediate folds. (Both plots show that state FL11 is an intermediate along the folding pathway to the native conformation.)

**Attachment of RNA at the Clay Surface.** Our previous results showed that RNA adsorbs to cationic clay mineral surfaces through nucleic acid base groups that bind face down, parallel to the surface for which we reproduce some snapshots in Figure S7 in the Supporting Information.[15,39] The simulations reported here show that this interaction is similarly present for larger RNA molecules. The hammerhead ribozyme binds to the clay mineral surface through a planar terminal base group, leaving the catalytic center intact and exposed to the aqueous region (Figure 5). The ribozyme was found to bind only through base groups that belong to the chain that is cleaved into two daughter strands during the self-cleavage reaction.

The folding pathways of the clay-adsorbed ribozyme funnel to a native state unlike the crystal structure. There is a severe disruption of the hydrogen-bonding network when the hammerhead ribozyme is adsorbed on the clay surface, shown in Tables 1 and 2. Although the folding landscape has undergone a significant change, the robust hammerhead ribozyme still retains active site contacts that facilitate phosphodiester bond cleavage, as has been shown experimentally.[12]

## DISCUSSION

The simulations reported here are of unprecedented scale, in terms of both size and duration as well as being novel in their application in understanding ribozyme function in inorganic settings. Because of the large-scale tightly coupled nature of REMD, it was only possible to run these simulations on peta-scale computing resources. REMD exchanges temperatures that allow us to traverse otherwise inaccessible areas of the conformational landscape, unlike regular uncoupled MD ensembles, such as Folding@Home,[40] that use many smaller single core resources but cannot exchange temperature information between replicas. For the simulations reported here, we used the JUGENE machine at Forschungszentrum, Jülich (Germany) and the Kraken machine at the National Institute for Computational Sciences in Tennessee (USA). Exploration of the relationship between structure, dynamics, and function of ribozyme molecules is well-suited to computationally expensive sampling techniques such as REMD because of its enhanced ability to identify complex free-energy landscapes such as those arising in the folding of RNAs and proteins. The way in which the structure and dynamics of the hammerhead permit the ribozyme to fold so as to function as a self-cleaving ribozyme, as well as if and how ions and water participate in the mechanism, gives us insight into to how the wider class of ribozyme molecules functions in nature. We have elucidated two very different active site conformations of the hammerhead ribozyme that have the potential to exhibit catalytic activity. This demonstrates the ability of ribozymes to retain catalytic capacity under differing environments and prevailing geochemistry. Our simulations show that direct participation of a magnesium ion in coordinating phosphate groups within the active site of the hammerhead ribozyme provides a mechanistic explanation for the enhanced catalytic activity of the ribozyme in bulk water over that in a confined clay mineral environment. The magnesium ion migrates from the crystallographically observed position to bridge the A9 and scissile phosphate, as shown in Figure 6a. This also allows the G8 base to act as a general acid in the reaction scheme.

The reduction in catalytic activity we see in the hammerhead ribozyme when adsorbed to the clay mineral surface can be attributed to the propensity of the ribozyme to then adopt a more compact conformation, concomitantly "closing" the catalytic pocket and thereby removing any participating cations. Despite the compact nature of the tertiary structure, the active-site geometry (shown in Figure 6b and Figure S9 in the SI) still supports the proposed reaction mechanism.

Like many ribozymes found in nature, the hammerhead ribozyme consists of a number of conserved residues that make up the active site. It follows that the ribozyme would have likely evolved through the addition of residues around the active site conferring an advantage over competing ribozymes by increasing the rate of catalysis. The minimal hammerhead is an example of an ancestral hammerhead ribozyme that has a reduced catalytic rate and is approximately half the size of the full-length hammerhead ribozyme studied here.[41] The clay mineral environment significantly disrupts the hydrogen-bonding network that scaffolds the full tertiary structure, leaving a minimal construct and reducing the catalytic advantage gained from the tertiary contacts.

Our analysis of folding rates between conformational states supports our previous findings that a clay mineral environment accelerates the rate of folding of RNA sequences over the same sequences in bulk water. The cationic atmosphere around the clay mineral surface that counters the net negative charge of the montmorillonite clay screens repulsive electrostatic charges between adjacent and nonadjacent phosphate groups along the backbone of RNA, reducing the energetic barrier between conformational folded states. Our original study showed that small (25 base) RNA sequences enjoy enhanced folding when interacting with a clay mineral surface.[15] The findings we report here show that larger, biologically functional RNAs also undergo enhanced folding in these environments along well-defined pathways to a native conformation.

The simulations reported here for the hammerhead ribozyme elucidate many facets of the relationship between the structure, dynamics, and function of naturally occurring ribozymes in





DOI: 10.1021/la503685t
Langmuir 2015, 31, 2493−2501

biologically and geochemically relevant settings. Our findings also resolve current conflicts regarding the mechanism of action of the hammerhead ribozyme and should thus serve to encourage structure-dynamics-function studies of the broader class of ribozymes found in nature.

## ■ ASSOCIATED CONTENT

### Supporting Information

Additional methods, simulation convergence verification, details of structural clustering algorithms used, computational resources, and additional results. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION


### Corresponding Author

*Tel: +44 (0)20 7679 4560. Fax: +44 (0)20 7679 1501. E-mail: p.v.coveney@ucl.ac.uk.

### Present Address

[†]D.W.W.: Department of Structural and Molecular Biology, University College London, London, WC1E 6BT, U.K.



### Funding

This research was supported by PRACE through JUGENE at Jülich Supercomputing Centre under project PRA044 and by the National Science Foundation through XSEDE resources provided by the National Institute for Computer Science within allocations under NRAC grant MCA04N014 and MRAC grant DMR070013N. We thank UCL Research Computing for the use of the Legion machine. J.B.S. was funded by an EPSRC Ph.D. studentship and the FP7VPH-SHARE (FP7-ICT 269978) project. We are grateful to EU FP7 CHAIN (HEALTH-2007-2.3.2-7) and the Qatar National Research Fund (09-26-01-048).


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS


We thank Dr. Benjamin Hall and Dr. Gregory Bowman for helpful discussions.


## ■ REFERENCES


(1) Woese, C. R. Evolution of the Genetic Code. *Naturwissenschaften* **1973**, *60*, 447−459.

(2) Crick, F. H. C. The Origin of the Genetic Code. *J. Mol. Biol.* **1968**, *38*, 367−379.

(3) Orgel, L. E. Evolution of the Genetic Apparatus. *J. Mol. Biol.* **1968**, *38*, 381−393.

(4) Cech, T. R. Nobel Lecture. Self-Splicing and Enzymatic Activity of an Intervening Sequence RNA from Tetrahymena. *Biosci. Rep.* **1990**, *10*, 239−261.

(5) Pitsch, S.; Eschenmoser, A.; Gedulin, B.; Hui, S.; Arrhenius, G. Mineral Induced Formation of Sugar Phosphates. *Origins Life Evol. Biospheres* **1995**, *25*, 297−334.

(6) Ferris, J. P. Montmorillonite Catalysis of 30-50 Mer Oligonucleotides: Laboratory Demonstration of Potential Steps in the Origin of the RNA World. *Origins Life Evol. Biospheres* **2002**, *32*, 311−332.

(7) Franchi, M.; Gallori, E. A Surface-Mediated Origin of the RNA World: Biogenic Activities of Clay-Adsorbed RNA Molecules. *Gene* **2005**, *346*, 205−214.

(8) De la Peña, M.; Gago, S.; Flores, R. Peripheral Regions of Natural Hammerhead Ribozymes Greatly Increase Their Self-Cleavage Activity. *EMBO J.* **2003**, *22*, 5561−5570.

(9) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884−890.

(10) Luk, L. Y. P.; Javier Ruiz-Pernía, J.; Dawson, W. M.; Roca, M.; Loveridge, E. J.; Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J.; Tuñón, I.; Moliner, V.; et al. Unraveling the Role of Protein Dynamics in Dihydrofolate Reductase Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16344−16349.

(11) Martick, M.; Lee, T.-S.; York, D. M.; Scott, W. G. Solvent Structure and Hammerhead Ribozyme Catalysis. *Chem. Biol.* **2008**, *15*, 332−342.

(12) Biondi, E.; Branciamore, S.; Fusi, L.; Gago, S.; Gallori, E. Catalytic Activity of Hammerhead Ribozymes in a Clay Mineral Environment: Implications for the RNA World. *Gene* **2007**, *389*, 10−18.

(13) Coveney, P. V.; Swadling, J. B.; Wattis, J. A. D.; Greenwell, H. C. Theory, Modelling and Simulation in Origins of Life Studies. *Chem. Soc. Rev.* **2012**, *41*, 5430.

(14) Swadling, J. B.; Coveney, P. V.; Christopher Greenwell, H. Stability of Free and Mineral-Protected Nucleic Acids: Implications for the RNA World. *Geochim. Cosmochim. Acta* **2012**, *83*, 360−378.

(15) Swadling, J. B.; Coveney, P. V.; Greenwell, H. C. Clay Minerals Mediate Folding and Regioselective Interactions of RNA: A Large-Scale Atomistic Simulation Study. *J. Am. Chem. Soc.* **2010**, *132*, 13750−13764.

(16) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/gamma Conformers. *Biophys. J.* **2007**, *92*, 3817−3829.

(17) Cygan, R. T.; Liang, J.-J.; Kalinichev, A. G. Molecular Models of Hydroxide, Oxyhydroxide, and Clay Phases and the Development of a General Force Field. *J. Phys. Chem. B* **2004**, *108*, 1255−1266.

(18) Harding, J. H.; Duffy, D. M.; Sushko, M. L.; Rodger, P. M.; Quigley, D.; Elliott, J. A. Computational Techniques at the Organic-Inorganic Interface in Biomineralization. *Chem. Rev.* **2008**, *108*, 4823−4854.

(19) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(20) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(21) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(22) Pitera, J. W.; Swope, W. Understanding Folding and Design: Replica-Exchange Simulations of "Trp-Cage" Miniproteins. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7587−7592.

(23) Bowman, G. R.; Huang, X.; Yao, Y.; Sun, J.; Carlsson, G.; Guibas, L. J.; Pande, V. S. Structural Insight into RNA Hairpin Folding Intermediates. *J. Am. Chem. Soc.* **2008**, *130*, 9676−9678.

(24) Garcia, A. E.; Paschek, D. Simulation of the Pressure and Temperature Folding/unfolding Equilibrium of a Small RNA Hairpin. *J. Am. Chem. Soc.* **2008**, *130*, 815−817.

(25) Chen, A. A.; García, A. E. High-Resolution Reversible Folding of Hyperstable RNA Tetraloops Using Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16820−16825.

(26) Juneja, A.; Villa, A.; Nilsson, L. Elucidating the Relation between Internal Motions and Dihedral Angles in an RNA Hairpin Using Molecular Dynamics. *J. Chem. Theory Comput.* **2014**, 140627112247001.

(27) Johnson, R. R.; Kohlmeyer, A.; Johnson, A. T. C.; Klein, M. L. Free Energy Landscape of a DNA-Carbon Nanotube Hybrid Using Replica Exchange Molecular Dynamics. *Nano Lett.* **2009**, *9*, 537−541.

(28) *Computational Studies of RNA and DNA*; Šponer, J., Lankaš, F., Eds.; Springer: Dordrecht, The Netherlands, 2006.

(29) Martick, M.; Scott, W. G. Tertiary Contacts Distant from the Active Site Prime a Ribozyme for Catalysis. *Cell* **2006**, *126*, 309−320.

(30) Anderson, M.; Schultz, E. P.; Martick, M.; Scott, W. G. Active-Site Monovalent Cations Revealed in a 1.55-Å-Resolution Hammerhead Ribozyme Structure. *J. Mol. Biol.* **2013**, *425*, 3790−3798.

(31) Wang, S.; Karbstein, K.; Peracchi, A.; Beigelman, L.; Herschlag, D. Identification of the Hammerhead Ribozyme Metal Ion Binding






DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501

Site Responsible for Rescue of the Deleterious Effect of a Cleavage Site Phosphorothioate. *Biochemistry* **1999**, *38*, 14363−14378.
(32) Behrouzi, R.; Roh, J. H.; Kilburn, D.; Briber, R. M.; Woodson, S. A. Cooperative Tertiary Interaction Network Guides RNA Folding. *Cell* **2012**, *149*, 348−357.
(33) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic Discovery of Metastable States for the Construction of Markov Models of Macromolecular Conformational Dynamics. *J. Chem. Phys.* **2007**, *126*, 155101.
(34) Noé, F.; Fischer, S. Transition Networks for Modeling the Kinetics of Conformational Change in Macromolecules. *Curr. Opin. Struct. Biol.* **2008**, *18*, 154−162.
(35) Singhal, N.; Snow, C. D.; Pande, V. S. Using Path Sampling to Build Better Markovian State Models: Predicting the Folding Rate and Mechanism of a Tryptophan Zipper Beta Hairpin. *J. Chem. Phys.* **2004**, *121*, 415−425.
(36) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. Molecular Simulation of Ab Initio Protein Folding for a Millisecond Folder NTL9(1-39). *J. Am. Chem. Soc.* **2010**, *132*, 1526−1528.
(37) Bowman, G. R.; Huang, X.; Pande, V. S. Using Generalized Ensemble Simulations and Markov State Models to Identify Conformational States. *Methods* **2009**, *49*, 197−201.
(38) Noé, F.; Schütte, C.; Vanden-Eijnden, E.; Reich, L.; Weikl, T. R. Constructing the Equilibrium Ensemble of Folding Pathways from Short off-Equilibrium Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 19011−19016.
(39) Swadling, J.; Suter, J.; Greenwell, H.; Coveney, P. Influence of Surface Chemistry and Charge on Mineral−RNA Interactions. *Langmuir* **2013**, *29*, 1573−1583.
(40) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S. P.; Khaliq, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; et al. Atomistic Protein Folding Simulations on the Submillisecond Time Scale Using Worldwide Distributed Computing. *Biopolymers* **2003**, *68*, 91−109.
(41) Scott, W. G.; Finch, J. T.; Klug, A. The Crystal Structure of an AII-RNAhammerhead Ribozyme: A Proposed Mechanism for RNA Catalytic Cleavage. *Cell* **1995**, *81*, 991−1002.
(42) Yang, H.; Jossinet, F.; Leontis, N.; Chen, L.; Westbrook, J.; Berman, H.; Westhof, E. Tools for the Automatic Identification and Classification of RNA Base Pairs. *Nucleic Acids Res.* **2003**, *31*, 3450−3460.





DOI: 10.1021/la503685t
*Langmuir* 2015, 31, 2493−2501