

---

# An active site rearrangement within the *Tetrahymena* group I ribozyme releases nonproductive interactions and allows formation of catalytic interactions

---

RAGHUVIR N. SENGUPTA,<sup>1</sup> SABINE N.S. VAN SCHIE,<sup>1,2</sup> GEORGE GIAMBAŞU,<sup>3</sup> QING DAI,<sup>4,5</sup>  
JOSEPH D. YESSELMAN,<sup>1</sup> DARRIN YORK,<sup>3</sup> JOSEPH A. PICCIRILLI,<sup>4,5</sup> and DANIEL HERSCHLAG<sup>1,6,7,8</sup>

<sup>1</sup>Department of Biochemistry, Stanford University, Stanford, California 94305, USA

<sup>2</sup>Leiden Institute of Chemistry, Leiden University, Leiden, 2333 CC, The Netherlands

<sup>3</sup>Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854, USA

<sup>4</sup>Department of Chemistry, <sup>5</sup>Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois 60637, USA

<sup>6</sup>Department of Chemical Engineering, <sup>7</sup>Department of Chemistry, <sup>8</sup>Department of Chemistry, Engineering, and Medicine for Human Health (ChEM-H), Stanford University, Stanford, California 94305, USA

## ABSTRACT

Biological catalysis hinges on the precise structural integrity of an active site that binds and transforms its substrates and meeting this requirement presents a unique challenge for RNA enzymes. Functional RNAs, including ribozymes, fold into their active conformations within rugged energy landscapes that often contain misfolded conformers. Here we uncover and characterize one such “off-pathway” species within an active site after overall folding of the ribozyme is complete. The *Tetrahymena* group I ribozyme (E) catalyzes cleavage of an oligonucleotide substrate (S) by an exogenous guanosine (G) cofactor. We tested whether specific catalytic interactions with G are present in the preceding E•S•G and E•G ground-state complexes. We monitored interactions with G via the effects of 2'- and 3'-deoxy (-H) and -amino (-NH<sub>2</sub>) substitutions on G binding. These and prior results reveal that G is bound in an inactive configuration within E•G, with the nucleophilic 3'-OH making a nonproductive interaction with an active site metal ion termed M<sub>A</sub> and with the adjacent 2'-OH making no interaction. Upon S binding, a rearrangement occurs that allows both -OH groups to contact a different active site metal ion, termed M<sub>C</sub>, to make what are likely to be their catalytic interactions. The reactive phosphoryl group on S promotes this change, presumably by repositioning the metal ions with respect to G. This conformational transition demonstrates local rearrangements within an otherwise folded RNA, underscoring RNA's difficulty in specifying a unique conformation and highlighting Nature's potential to use local transitions of RNA in complex function.

**Keywords:** active site; conformational change; metal ion; noncoding RNA; RNA catalysis

## INTRODUCTION

A requirement for robust biological catalysis is the assembly of an active site that binds substrates and positions reactive groups and active site residues with respect to one another (Fersht 1999). For protein enzymes, X-ray structures with and without bound ligands suggest that protein active sites are largely organized upon folding but typically undergo some level of conformational rearrangement upon substrate binding, such as flap closure or domain rotation (e.g., Bennett and Steitz 1978; Lesk and Chothia 1984; Lolis and Petsko 1990; Bystroff and Kraut 1991; Sawaya et al. 1997). In contrast, examples of proteins with resting unfolded states are less common (Schulenburg and Hilvert 2013).

Fulfilling the requirement for positioning poses a special challenge for catalytic RNAs (ribozymes), as RNA's negative-

ly charged backbone and limited side-chain diversity are expected to hinder its ability to form compact and precisely positioned structures (Herschlag 1995; Narlikar and Herschlag 1997). Early X-ray structures of ribozymes implied that RNA could form protein-like structures containing active sites that are preorganized for catalysis (Ferré-D'Amaré et al. 1998; Golden et al. 1998; Herschlag 1998) but subsequent results and reflection led to the recognition of differences between RNA and protein enzymes that may be of biological importance, as described below.

While proteins harness extensive noncovalent interactions to form a densely packed core (Dill 1990), RNAs, which exist as rigid helices connected by loops and junctions, are brought

© 2015 Sengupta et al. This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

---

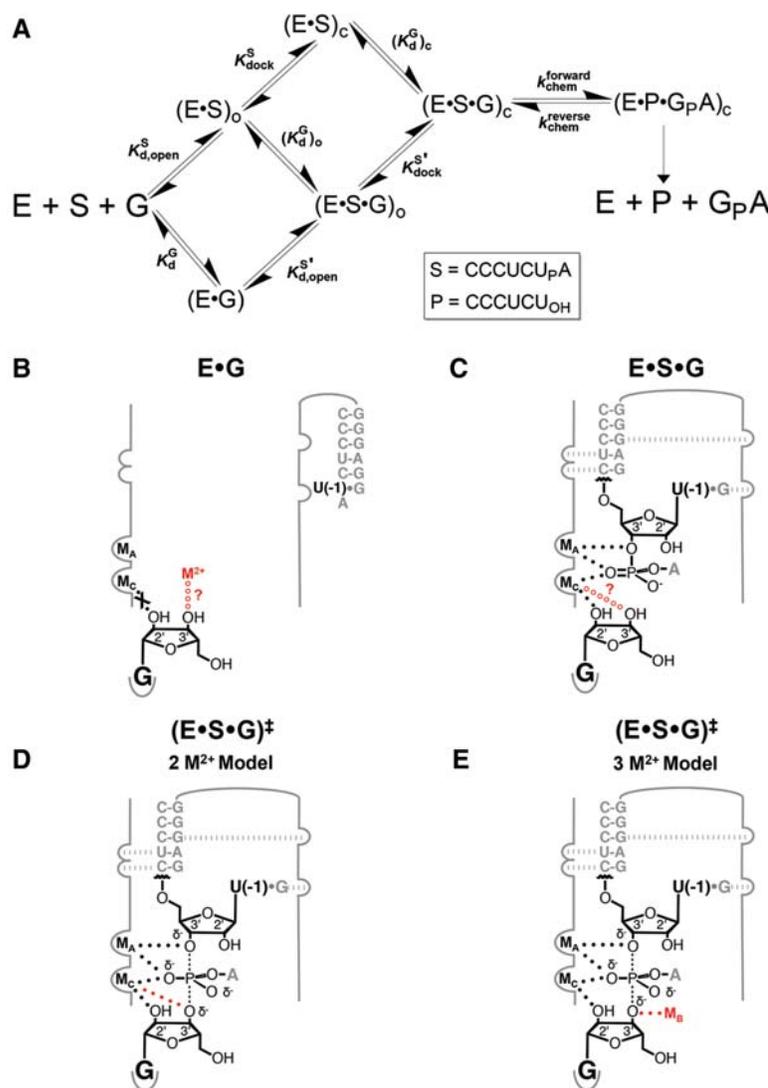
Corresponding author: [herschla@stanford.edu](mailto:herschla@stanford.edu)

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.053710.115>.

together through a handful of tertiary contacts and thus have more limited structural interconnections (Sigler 1975). The absence of extensive close packing, along with RNA's inherent ability and tendency to form stable alternative structures (Herschlag 1995; Russell 2008), may result in rugged folding energy landscapes for RNAs (e.g., Chen and Dill 2000; Treiber and Williamson 2001; Bokinsky and Zhuang 2005; Shcherbakova et al. 2008). Within these energy landscapes are "off-pathway" species that form nonnative interactions and require additional rearrangements to attain functional states. Indeed, conformational transitions have been inferred within the active sites of ribozymes (e.g., Chanfreau and Jacquier 1996; Wang et al. 1999; Zhuang et al. 2002a; Schmeing et al. 2005; Houglund et al. 2006; Martick and Scott 2006; Hsieh and Fierke 2009; Marcia and Pyle 2012; Sripathi et al. 2014) and in the binding sites of in vitro-selected aptamers (e.g., Jucker et al. 2003; Flinders et al. 2004; Duchardt-Ferner et al. 2010), perhaps reflecting RNA's difficulty in specifying a unique structure. Nature appears to have harnessed this feature to regulate gene expression via conformational transitions of riboswitches (e.g., Wickiser et al. 2005; Gilbert et al. 2006; Noeske et al. 2007; Ottink et al. 2007; Montange and Batey 2008) and coordinate complex multistep processes such as pre-mRNA splicing and protein synthesis (Staley and Guthrie 1998; Guo et al. 2009; Frank and Gonzalez 2010; Voorhees and Ramakrishnan 2013; Chen and Moore 2014). Given the widespread occurrence of RNA conformational changes, understanding such changes at an atomic level and how they impact assembly of RNA active sites is an important endeavor that can be pursued by studying model RNAs.

Here we investigate a structural rearrangement within the active site of the *Tetrahymena* group I ribozyme (E), which catalyzes the cleavage of an oligonucleotide substrate (S) by an exogenous guanosine (G) cofactor (Fig. 1A; see Houglund et al. 2006 and references therein). A transition state model derived from biochemical and crystallographic data implicates contacts between reactive groups on S and G and two distinct metal ions,  $M_A$  and  $M_C$  (Fig. 1D), and possibly a third metal ion ( $M_B$ , Fig. 1E; Piccirilli

et al. 1993; Sjögren et al. 1997; Weinstein et al. 1997; Shan et al. 1999a, 2001; Shan and Herschlag 1999; Yoshida et al. 1999; Stahley and Strobel 2005; Lipchock and Strobel 2008). G binding to the ribozyme is several orders of magnitude slower than diffusion, suggestive of a required rearrangement of the G binding site to allow binding to occur (Karbstein and Herschlag 2003). Once G is bound, an



**FIGURE 1.** The *Tetrahymena* group I ribozyme reaction. (A) Framework showing individual steps of the forward reaction (Houglund et al. 2005). The subscripts "O" and "C" refer to complexes in the open and closed states, respectively. The subscript, "chem," refers to the chemical step of the reaction. The subscript "P" on CCCUCU<sub>p</sub>A refers to the reactive phosphoryl group. The framework for the reverse reaction (Karbstein et al. 2002) is not shown for simplicity. (B–E) Models for active site metal ion interactions within E·G (B), E·S·G (C), and (E·S·G)<sup>‡</sup> (D,E). Closed circles and hatched lines represent metal ion interactions and hydrogen bonds, respectively. Partial-negative charges in (E·S·G)<sup>‡</sup> (D,E) are represented by "δ<sup>-</sup>." The absence of an interaction between  $M_C$  and the G 2'-oxygen atom in E·G is denoted by "X." In this work, we test whether the 3'-oxygen atom of G contacts a metal ion ( $M^{2+}$ ) in E·G (B) and  $M_C$ , as suggested by crystallographic data (Stahley and Strobel 2005), in E·S·G (C) (represented by open circles and "?"). In (E·S·G)<sup>‡</sup>, structural data (Stahley and Strobel 2005) suggest that the G 3'-oxygen atom contacts  $M_C$  (D) and biochemical data (Shan et al. 1999a) suggest that this atom interacts with a metal ion distinct from  $M_C$ , termed  $M_B$  (E).

additional rearrangement is required to assemble the active site metal ion interactions required for catalysis. The 2'-OH group of G, which contacts  $M_C$  in the transition state (Fig. 1D,E), does not form this interaction until S is present in the active site (cf. Fig. 1B,C; Shan and Herschlag 1999), and these and other data (Bevilacqua et al. 1993; McConnell et al. 1993; Shan and Herschlag 1999, 2002; Forconi et al. 2010, 2011) provide strong evidence for a conformational change within the active site from  $E \cdot G$  to  $E \cdot S \cdot G$ .

To further interrogate this proposed structural rearrangement, we tested for potential interactions made with the nucleophilic 3'-OH group of G in the absence and presence of bound S. Our results reveal an inactive, off-pathway complex formed within  $E \cdot G$  and a rearrangement to a catalytically competent configuration upon binding of S. Such non-stepwise assembly of active site interactions highlights RNA's tendency to form alternative interactions, even within the active site of a folded RNA, and underscores the importance of local rearrangements in forming a functional RNA.

## RESULTS

### Use of 2'- and 3'-deoxyguanosine to probe interactions within the $(E \cdot S \cdot G)_O$ and $(E \cdot S \cdot G)_C$ complexes

The *Tetrahymena* ribozyme (E) catalyzes cleavage of an oligonucleotide substrate (S) by an exogenous guanosine (G) cofactor. Binding of S occurs in two steps (Fig. 1A), first by base-pairing with E to form the  $(E \cdot S)_O$  complex (where "O" stands for open) and then by docking of the substrate-containing duplex, referred to as P1, into the active site via tertiary interactions to form the  $(E \cdot S)_C$  complex (where "C" stands for closed) (Bevilacqua et al. 1992, 1994; Herschlag 1992; Li et al. 1995; Narlikar and Herschlag 1996). Prior results strongly suggest that the interactions made with G are the same within the  $E \cdot G$  and  $(E \cdot S \cdot G)_O$  complexes because S is absent from the active site in both complexes (Wang et al. 1993; Narlikar and Herschlag 1996), but are different within the  $(E \cdot S \cdot G)_C$  complex (McConnell et al. 1993; Wang et al. 1993; Shan and Herschlag 1999; Karbstein et al. 2002; Forconi et al. 2010).<sup>9</sup>

Here we focus on interactions with the 2'- and 3'-OH groups of G within the  $(E \cdot S \cdot G)_O$  and  $(E \cdot S \cdot G)_C$  complexes. The G 2'-OH group contacts an active site metal ion only when S is present in the active site, i.e., in the closed complex (Fig. 1C; Shan and Herschlag 1999), but it is not known whether the adjacent 3'-OH forms metal ion interactions in the open and/or closed complexes. Below, we describe experiments designed to probe interactions with the 3'-OH and, for comparison, the 2'-OH groups of G in the  $(E \cdot S \cdot G)_O$  and  $(E \cdot S \cdot G)_C$  complexes.

If bound G is stabilized by interactions made with a hydroxyl group, then hydrogen (-H) substitution would ablate

those interactions and be expected to reduce the binding affinity. Alternatively, the absence of a deleterious effect on G binding would suggest, most simply, that no interaction is made with the hydroxyl group. This modification has been used to probe the functional relevance of -OH groups in essentially every catalytic RNA (e.g., Bass and Cech 1984; Abramovitz et al. 1996; Ortoleva-Donnelly et al. 1998; Weinger et al. 2004; Nelson and Uhlenbeck 2008). We thus introduced single -H substitutions at the 2'- and 3'-positions and carried out binding assays without or with the substrate-containing P1 helix docked into the catalytic core [the  $(E \cdot S)_O$  and  $(E \cdot S)_C$  complexes, respectively; Materials and Methods; Fig. 1A].

We determined binding affinities for three G analogs: AUCG, AUCG(2'H), and AUCG(3'H) (Table 1). The common 5'-AUC extension enhances binding of the 3'-terminus G, G(2'H), and G(3'H) through base-pairing and stacking interactions (making the so-called "P9.0" interactions; (Michel et al. 1989; Burke et al. 1990; Michel and Westhof 1990; Moran et al. 1993; Russell and Herschlag 1999)) and thereby overcomes assay limitations imposed by the limited solubility of G and allows us to extend prior results that gave binding limits and provide side-by-side comparisons (Moran et al. 1993; Profenno et al. 1997; Karbstein et al. 2002).

The equilibrium constants for binding of AUCG to  $(E \cdot S)_O$  and  $(E \cdot S)_C$  were determined by measuring the rate constant for cleavage of <sup>32</sup>P-labeled S as a function of AUCG concentration in single-turnover assays, with versions of S designed to ensure that the desired  $E \cdot S$  complex was present (Materials and Methods; Table 2; Herschlag et al. 1993; Wang et al. 1993; Narlikar and Herschlag 1996; Bartley et al. 2003). However, AUCG(2'H) and AUCG(3'H) are essentially non-reactive (Bass and Cech 1986; Moran et al. 1993; Shan and Herschlag 1999), as G(3'H) has no lone pair of electrons to react with the scissile phosphoryl group, and G(2'H) reduces the rate of reaction by  $\geq 10^6$ -fold (Shan and Herschlag 1999). Therefore, we measured binding of these analogs through competitive inhibition of reactions with subsaturating G under single-turnover conditions so that observed inhibition constants represent true dissociation constants (see Materials

TABLE 1. Guanosine (G) analogs used in this work

Abbreviation	R	X	Y
G	H	OH	OH
G(2'NH <sub>2</sub> )	H	OH	NH <sub>2</sub>
G(3'NH <sub>2</sub> )	H	NH <sub>2</sub>	OH
AUCG	AUC	OH	OH
AUCG(2'H)	AUC	OH	H
AUCG(3'H)	AUC	H	OH
CUCG <sub>p</sub> A	CUC	pA	OH

p, reactive phosphoryl group.

<sup>9</sup>The  $E \cdot G$  and  $(E \cdot S \cdot G)_O$  complexes are used interchangeably as both complexes have only G in the active site (Fig. 1B). For simplicity we also use  $E \cdot S \cdot G$  and  $(E \cdot S \cdot G)_C$  to indicate that both G and S are present in the active site (Fig. 1C).

**TABLE 2.** Oligonucleotide substrates (S) and products (P) used in this work

Abbreviation	Position:	Residue							Attributes
		−6	−5	−4	−3	−2	−1	+1	
−1d,rSA		C	C	C	U	C	dU <sub>P</sub>	A	Docked, slow chemical step, forward reaction
−3m,rS <sub>Me</sub>		C	C	C	mU	C	U <sub>P</sub>	Me	Undocked, forward reaction
−1d,rP		C	C	C	U	C	dU		Docked, slow chemical step, reverse reaction

d, 2'-H; m, 2'-OCH<sub>3</sub>; <sub>P</sub>, reactive phosphoryl group; residues without a suffix have ribose sugars (2'-OH). −3m,rS<sub>Me</sub> bears a methyl group (Me) at the +1 position instead of an A residue, which eliminates a metal ion interaction with A(+1) and simplifies quantitative metal ion rescue experiments described herein (Shan and Herschlag 2000).

and Methods; Shan and Herschlag 1999; Karbstein et al. 2002; Hougland et al. 2008).

Table 3 reports the equilibrium dissociation constants obtained, and the effects of -H substitutions on AUCG binding to (E•S)<sub>O</sub> and (E•S)<sub>C</sub> are shown in Figure 2A and B, respectively (and see Supplemental Fig. S1 for raw data). AUCG(2'H) binds 60-fold weaker than AUCG to (E•S)<sub>C</sub>, but binds to (E•S)<sub>O</sub> with nearly the same affinity as AUCG. These binding effects are consistent with predictions from prior results with 2'-aminoguanosine [G(2'NH<sub>2</sub>), Table 1], which provided evidence for a stabilizing G 2'-OH interaction in (E•S•G)<sub>C</sub>, but not (E•S•G)<sub>O</sub> (Shan and Herschlag 1999). In contrast, the 3'-H substitution weakens binding to both (E•S)<sub>O</sub> and (E•S)<sub>C</sub> (Fig. 2A,B), suggesting that interactions to the 3'-OH group are present in both the open and closed complexes.

### Use of 2'- and 3'-amino substitutions to test for metal ion interactions with guanosine in the (E•S•G)<sub>C</sub> complex

Given the prior functional and crystallographic evidence for metal ion interactions with G hydroxyl groups in the group I ribozyme active site (Fig. 1B–E; Piccirilli et al. 1993; Sjögren et al. 1997; Weinstein et al. 1997; Shan et al. 1999a, 2001; Shan and Herschlag 1999; Yoshida et al. 1999; Stahley and Strobel 2005; Lipchock and Strobel 2008), we used metal ion rescue and thermodynamic fingerprint analysis (TFA) to delineate metal ion interactions. TFA provides a readout of the affinity for the metal ion responsible for rescuing a deleterious effect from a single atom or functional group substitution. As metal ions bound at different sites are expected, most simply, to have different affinities, the apparent affinity of a rescuing metal ion provides a functional thermodynamic “fingerprint.” By comparing the affinities of metal ions providing rescue of substitutions at different positions, one can determine whether these interactions are mediated by the same metal ion or by distinct metal ions (Shan et al. 1999a; Wang et al. 1999;

Christian 2005; Hougland et al. 2006; Frederiksen and Piccirilli 2009). Such analyses and follow-up experiments can lead to detailed maps of functional interactions (Shan et al. 1999a; Wang et al. 1999; Christian et al. 2000; Gordon et al. 2000; Gordon and Piccirilli 2001; Hougland et al. 2005; Forconi et al. 2008; Ward and Derose 2012; Fica et al. 2013).

Prior TFA experiments for the *Tetrahymena* ribozyme reaction provided evidence for transition state interactions between two metal ions, M<sub>C</sub> and M<sub>B</sub>, and the 2'-OH and 3'-OH groups of G, respectively (Fig. 1E; Shan et al. 1999a). The most complete and highest resolution crystal structures of the group I ribozyme, from *Azoarcus*, showed no evidence of M<sub>B</sub> and instead provided evidence for both the 2'-OH and 3'-OH groups of G interacting with M<sub>C</sub> (Fig. 1D; Stahley and Strobel 2005; Lipchock and Strobel 2008). Models for this discrepancy have been described (Discussion; see Hougland et al. 2006) and we note that the remaining interactions inferred from the functional and structural data are in excellent agreement (Hougland et al. 2006; Forconi et al. 2008, 2009). In this and the next section we test whether the 3'-OH group of G interacts with a metal ion in the (E•S•G)<sub>O</sub> and (E•S•G)<sub>C</sub> complexes and, if so, which one (M<sub>A</sub>, M<sub>B</sub>, or M<sub>C</sub>).

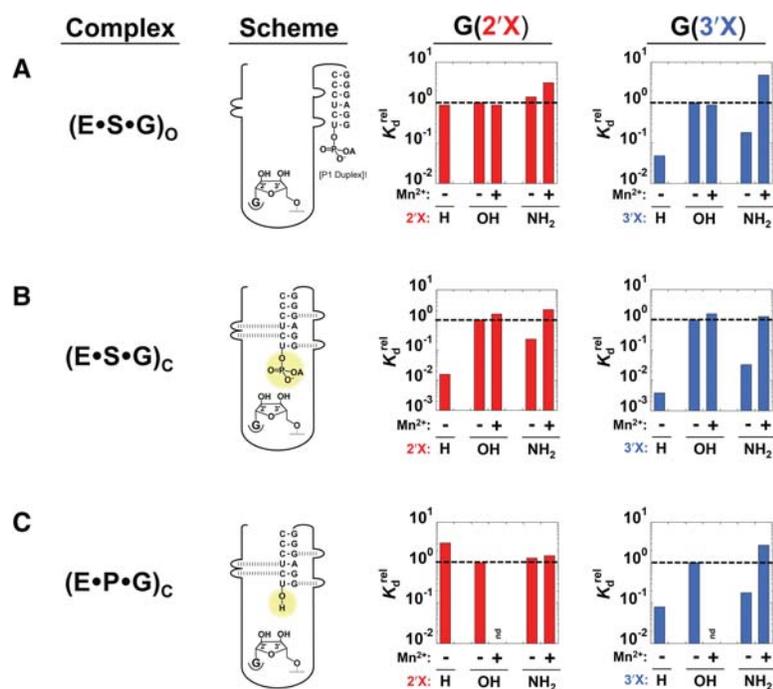
To test for a metal ion interaction, we replaced the G–OH group with an amino (–NH<sub>2</sub>) group. Nitrogen, compared to oxygen, generally interacts more favorably with Mn<sup>2+</sup> than Mg<sup>2+</sup>, and prior results show stronger binding of G(2'-NH<sub>2</sub>) (Table 1) to (E•S)<sub>C</sub> in the presence of Mn<sup>2+</sup> (Shan and Herschlag 1999). In addition, the Mn<sup>2+</sup>

**TABLE 3.** Effects of deoxyribose (–H) substitution on AUCG binding to (E•S)<sub>O</sub>, (E•S)<sub>C</sub>, and (E•P)<sub>C</sub>

AUCG analog	(E•S) <sub>O</sub>		(E•S) <sub>C</sub>		(E•P) <sub>C</sub>	
	K <sub>d</sub> (μM)	K <sub>d</sub> <sup>rel</sup>	K <sub>d</sub> (μM)	K <sub>d</sub> <sup>rel</sup>	K <sub>d</sub> (μM)	K <sub>d</sub> <sup>rel</sup>
AUCG	2.9 ± 0.5	(1.0)	0.32 ± 0.05	(1.0)	2.2 ± 0.2	(1.0)
AUCG(2'H)	3.3 ± 0.4	0.88	20 ± 3	0.016	0.70 ± 0.1	3.1
AUCG(3'H)	60 ± 4	0.048	82 ± 11	0.0039	27 ± 4	0.081

$$K_d^{\text{rel}} = K_d^{\text{AUCG}} / K_d^{\text{AUCG analog}}$$

Values from data in Supplemental Figures S1 and S2 following approaches described in Materials and Methods.



**FIGURE 2.** Effects of 2'- and 3'-modifications on G binding to (E•S)<sub>O</sub>, (E•S)<sub>C</sub>, and (E•P)<sub>C</sub>. A–C correspond to the (E•S•G)<sub>O</sub>, (E•S•G)<sub>C</sub>, and (E•P•G)<sub>C</sub> complexes respectively. The hatched lines correspond to tertiary interactions made between the ribozyme and the P1 duplex with S or P docked into the active site. The scissile phosphoryl group within the (E•S•G)<sub>C</sub> complex is shaded yellow as is the 3'-hydroxyl group in the (E•P•G)<sub>C</sub> complex.  $K_d^{\text{rel}} (= K_d^{\text{G}}/K_d^{\text{G analog}})$ , i.e.,  $K_d$  for G or AUCG relative to  $K_d$  for the G or AUCG analog, was obtained from the data shown in Tables 2 and 3. The dashed line corresponds to  $K_d^{\text{rel}} = 1$ .

concentration dependence of G(3'NH<sub>2</sub>) binding to (E•S)<sub>C</sub> provide evidence for an interaction with M<sub>C</sub> (Shan and Herschlag 1999). If M<sub>C</sub> contacts the adjacent 3'-OH group, the simplest expectation is that binding of 3'-aminoguanosine [G(3'NH<sub>2</sub>), (Table 1)] to (E•S)<sub>C</sub> would similarly be preferentially stabilized by Mn<sup>2+</sup> compared with Mg<sup>2+</sup>, with the affinity of the stabilizing Mn<sup>2+</sup> ion matching that for Mn<sup>2+</sup> binding to site C.

We measured binding of G(3'NH<sub>2</sub>) to (E•S)<sub>C</sub> through competitive inhibition of the reaction with G under conditions designed to ensure that we were monitoring binding of G(3'NH<sub>2</sub>) and not G(3'NH<sub>3</sub><sup>+</sup>) (Materials and Methods). Our results are summarized in Table 4 and the relative effects of –NH<sub>2</sub> substitution on G binding to (E•S)<sub>C</sub> are shown in Figure 2B (see Supplemental Figs. S3B and S4–S6 for raw data). In the presence of Mg<sup>2+</sup>, the 3'-NH<sub>2</sub> substitution weakens binding of G by 30-fold (Fig. 2B). However, binding is strengthened to approximately the level of G itself upon the addition of Mn<sup>2+</sup> (Fig. 2B). These results provide evidence for a metal ion interaction with the 3'-group of G in the ground-state (E•S•G)<sub>C</sub> complex, analogous to the prior evidence for such an interaction with the 2'-group (Table 4; Fig. 2B; Shan and Herschlag 1999).

To determine which active site metal ion contacts G(3'NH<sub>2</sub>), we measured equilibrium constants for binding

of 3'-aminoguanosine [G(3'N)] and, for comparison, 2'-aminoguanosine [G(2'N)] as a function of Mn<sup>2+</sup> concentration (Fig. 3A; see Supplemental Fig. S11 for raw data).<sup>10</sup> The results of Figure 2 and prior data (Shan and Herschlag 1999) indicate that Mn<sup>2+</sup> has a negligible effect on G binding, so there is no need for ratiometric corrections in this case (Christian 2005; Houglund et al. 2006; Frederiksen and Piccirilli 2009). For each analog, the data were fit to a model (Fig. 3B) in which a single Mn<sup>2+</sup> ion stabilizes binding of G<sub>N</sub> (i.e., G(2'N) or G(3'N)) to the (E•S)<sub>C</sub> complex. This analysis reports  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{C}}}$ , the apparent<sup>11</sup> Mn<sup>2+</sup> affinity for (E•S)<sub>C</sub>, a parameter that is unaffected by the amino modifications since G<sub>N</sub> is not bound to the ribozyme (Shan et al. 1999a; Wang et al. 1999; Christian 2005; Houglund et al. 2006; Frederiksen and Piccirilli 2009). Identical values of  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{C}}}$  for G(2'N) and G(3'N) would suggest, most simply, that a single metal ion contacts the 2'- and 3'-moieties, whereas distinct  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{C}}}$  values would indicate that two different metal ions mediate these contacts.

As shown in Figure 3, the Mn<sup>2+</sup> rescue profiles for the two analogs are indistinguishable, and fitting the data to a model for binding of a single Mn<sup>2+</sup> ion to (E•S)<sub>C</sub> yields a  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{C}}}$  of  $2.0 \pm 0.8$  and  $1.3 \pm 0.3$  mM for G(3'N) and G(2'N), respectively (Fig. 3B). These values are the same, within error, as the reported Mn<sup>2+</sup> affinity to site C of the (E•S)<sub>C</sub> complex ( $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{C}}} = 1.1$  mM) (Shan et al. 1999a; Shan and Herschlag 1999) and differ significantly from the affinities for sites A and B (4.1 and 13 mM, respectively<sup>12</sup>; (Shan et al. 1999a)). These TFA results suggest that G(2'NH<sub>2</sub>) and G(3'NH<sub>2</sub>) contact the same Mn<sup>2+</sup> ion, Mn<sub>C</sub>, within (E•S•G)<sub>C</sub> (Fig. 3C). This is consistent with X-ray structural data obtained with the *Azoarcus* group I ribozyme,

<sup>10</sup>We refer to 3'-aminoguanosine and 2'-aminoguanosine as G(3'N) and G(2'N), respectively, instead of G(3'NH<sub>2</sub>) and G(2'NH<sub>2</sub>), respectively. This nomenclature denotes that, under the conditions of our TFA experiments (50 mM Mg<sup>2+</sup> at pH 7.7), both the –NH<sub>2</sub> and –NH<sub>3</sub><sup>+</sup> forms of these analogs are present. Mn<sup>2+</sup> stabilizes binding of only the –NH<sub>2</sub> form of these G analogs (Supplemental Figs. S5–S10; Shan and Herschlag 1999; Shan et al. 1999b), allowing us to obtain apparent affinities for the Mn<sup>2+</sup> ion that stabilizes binding of G(3'NH<sub>2</sub>) and G(2'NH<sub>2</sub>).

<sup>11</sup>The Mn<sup>2+</sup> affinities are apparent and not absolute affinities as prior results indicate that Mn<sup>2+</sup> competes with Mg<sup>2+</sup> for binding to metal sites A and C and, at higher concentrations, for binding to site B (Shan et al. 1999a). Thus, these values hold for a particular background concentration of competing Mg<sup>2+</sup>.

<sup>12</sup>The apparent Mn<sup>2+</sup> affinities for sites A, B, and C (see footnote 11) were obtained for (E•S)<sub>O</sub> (Shan et al. 1999a). Prior data (Shan and Herschlag 1999) indicate that formation of (E•S)<sub>C</sub> is unaffected by Mn<sup>2+</sup> binding to sites A–C. Therefore, the Mn<sup>2+</sup> affinities for sites A, B, and C are the same for (E•S)<sub>O</sub> and (E•S)<sub>C</sub>. Independent experimental evidence for the absence of a change in the apparent Mn<sup>2+</sup> affinities to site C is presented in Shan and Herschlag (1999).

**TABLE 4.** Effects of amino (–NH<sub>2</sub>) substitution on G binding to (E•S)<sub>O</sub>, (E•S)<sub>C</sub>, and (E•P)<sub>C</sub>

G analog	Mn <sup>2+</sup>	(E•S) <sub>O</sub>		(E•S) <sub>C</sub>		(E•P) <sub>C</sub>	
		K <sub>d</sub> (μM)	K <sub>d</sub> <sup>rel</sup>	K <sub>d</sub> (μM)	K <sub>d</sub> <sup>rel</sup>	K <sub>d</sub> (μM)	K <sub>d</sub> <sup>rel</sup>
G	–	340 ± 20	(1.0)	100 ± 30	(1.0)	480 ± 100	(1.0)
	+	390 ± 40	0.87	63 ± 5	1.6	nd	–
G(2'NH <sub>2</sub> )	–	240 <sup>a</sup>	1.4	440 ± 20	0.23	360 <sup>a</sup>	1.3
	+	110 <sup>a</sup>	3.1	45 ± 8	2.2	320 <sup>a</sup>	1.5
G(3'NH <sub>2</sub> )	–	1850 ± 270	0.18	3030 ± 280	0.033	2700 ± 730	0.18
	+	71 ± 20	4.8	77 ± 20	1.3	130 ± 30	3.7

$K_d^{rel} = K_d^G / K_d^{G\ analog}$ . Unless otherwise stated, values for G, G(2'NH<sub>2</sub>), and G(3'NH<sub>2</sub>) obtained from data in Supplemental Figures S3–S10, respectively. Measurements made in the absence or presence of 10 mM Mn<sup>2+</sup> (see Materials and Methods). (nd) Not determined.

<sup>a</sup>Values from Shan and Herschlag (1999). Measurements were made in the absence or presence of 20 mM Mn<sup>2+</sup>.

which suggest that both –OH groups contact M<sub>C</sub> (Stahley and Strobel 2005; Lipchock and Strobel 2008).

### Use of 3'-amino substitution to test for a metal ion interaction with the guanosine nucleophile in the (E•S•G)<sub>O</sub> complex

Functional data indicate that the G 2'-OH group, which contacts M<sub>C</sub> in (E•S•G)<sub>C</sub>, makes no interactions in (E•S•G)<sub>O</sub> (Table 3; Shan and Herschlag 1999). In contrast, our results with AUCG(3'H) suggest that an interaction is made with the 3'-OH of G in the open complex (Table 3). To determine whether this hydroxyl group interacts with a metal ion and, if so, which metal ion, we again conducted rescue and TFA experiments with G(3'NH<sub>2</sub>).

Table 4 reports equilibrium dissociation constants for G(3'NH<sub>2</sub>) from the (E•S•G(3'NH<sub>2</sub>))<sub>O</sub> complex, and Figure 2A shows the relative effects of –NH<sub>2</sub> substitution on G binding to (E•S)<sub>O</sub> (see Supplemental Figs. S3A, S7, and S8 for raw data).

G(3'NH<sub>2</sub>) binds fivefold weaker than G, and Mn<sup>2+</sup> strengthens binding of G(3'NH<sub>2</sub>) to (E•S)<sub>O</sub> by 27-fold, suggesting that the 3'-NH<sub>2</sub> group interacts with a metal ion. In contrast, prior data show that replacing the 2'-OH with a –NH<sub>2</sub> group does not weaken binding of G to (E•S)<sub>O</sub>, and Mn<sup>2+</sup> has no significant stabilizing effect on binding (Shan and Herschlag 1999). These results indicate that the 2'-OH group does not interact with a metal ion in the (E•S•G)<sub>O</sub> complex.

To determine which active site metal ion (M<sub>A</sub>, M<sub>B</sub>, or M<sub>C</sub>) (Fig. 1E) contacts the 3'-moiety of G(3'NH<sub>2</sub>), we measured equilibrium constants for binding of G(3'N) to (E•S)<sub>O</sub> across a range of Mn<sup>2+</sup> concentrations (see Supplemental Fig. S12A for raw data). As binding of G is the same in the absence and presence of Mn<sup>2+</sup> (Fig. 2A), it was not necessary to correct for the effect of Mn<sup>2+</sup> on G binding (Christian 2005; Houglund et al. 2006; Frederiksen and Piccirilli 2009). The data were fit to a model (Fig. 4C) for binding of a single Mn<sup>2+</sup> ion (black line) to give  $K_{Mn,app}^{(E•S)_O}$ , the apparent<sup>3</sup> Mn<sup>2+</sup> affinity for (E•S)<sub>O</sub>.

For comparison, we include the previously determined  $K_{Mn,app}^{(E•S)_O}$  values for sites A (4.1 mM, blue dashed line), B (13 mM, green dashed line), and C (1.3 mM, red dashed line) (Shan et al. 1999a). As shown in Figure 4A, the data strongly suggest that Mn<sub>C</sub> is not responsible for rescuing G(3'NH<sub>2</sub>) binding but do not unambiguously distinguish between Mn<sub>A</sub> and Mn<sub>B</sub> rescue.

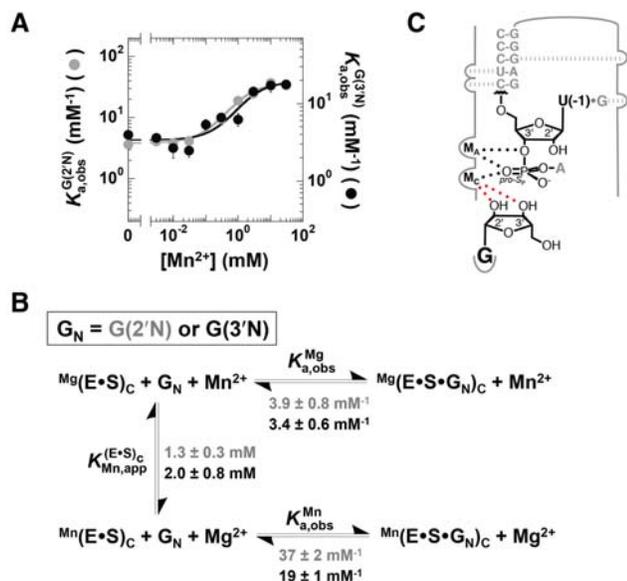
To determine whether Mn<sub>A</sub> or Mn<sub>B</sub> contacts G(3'NH<sub>2</sub>), we measured the Mn<sup>2+</sup> concentration dependence of G(3'N) binding to (E•S)<sub>O</sub> at a lower background concentration of Mg<sup>2+</sup> (10 mM, Fig. 4B). Prior work showed that Mg<sup>2+</sup>

competes for binding of Mn<sup>2+</sup> to site A but not site B (Shan et al. 1999a). Hence, if Mn<sub>A</sub> contact the 3'-NH<sub>2</sub> group,  $K_{Mn,app}^{(E•S)_O}$  is expected to decrease by a factor of 5 from 50 to 10 mM Mg<sup>2+</sup>, whereas little or no change would be expected for a Mn<sub>B</sub> contact (Shan et al. 1999a). The results indicate that lower Mn<sup>2+</sup> concentrations are required to stabilize G(3'N) binding in a 10 mM Mg<sup>2+</sup> background (cf. Fig. 4A,B), and the value of  $K_{Mn,app}^{(E•S)_O}$  quantitatively matches that predicted for Mn<sub>A</sub> but not Mn<sub>B</sub> (Fig. 4, blue versus green dashed lines). Thus, the 3'-moiety of G changes its interactions from (E•S•G)<sub>O</sub> (M<sub>A</sub>, Fig. 4D) to (E•S•G)<sub>C</sub> (M<sub>C</sub>, Fig. 3C).

### Is the reactive phosphoryl group on S responsible for the different interactions with G in the (E•S•G)<sub>O</sub> and (E•S•G)<sub>C</sub> complexes?

The different interactions made with G in the open (Fig. 4D) and closed (Fig. 3C) complexes could be explained by the reactive phosphoryl group (on S) reorienting the active site metal ions that contact G (Fig. 1C) or could occur even in the absence of this phosphoryl group due to indirect effects from the tertiary interactions made between the P1 duplex and ribozyme in the closed complex (Wang et al. 1993; Narlikar and Herschlag 1996, 1998; Narlikar et al. 1997). To distinguish these models, we compared binding of the G analogs to the (E•P)<sub>C</sub>, (E•S)<sub>C</sub>, and (E•S)<sub>O</sub> complexes. Both S and P (Table 2) utilize the same tertiary interactions to form the closed complex (Bevilacqua and Turner 1991; Pyle and Cech 1991; Pyle et al. 1992; Herschlag et al. 1993; Narlikar et al. 1997). However, the reactive phosphoryl group is not present in P (cf. schemes in Fig. 2B,C).

Binding of G analogs to the (E•P)<sub>C</sub> complex was measured through inhibition of the reverse reaction (see Materials and Methods). The results are summarized in Tables 3 and 4, and the relative effects of –H and –NH<sub>2</sub> substitution on G binding to (E•P)<sub>C</sub> are shown in Figure 2C (see Supplemental Figs. S2, S3C, S9, S10 for raw data). Replacing the 2'-OH group with a –H atom or a –NH<sub>2</sub> group does not weaken binding of G to (E•P)<sub>C</sub>. In contrast, 3'-H and 3'-NH<sub>2</sub> substitutions weaken



**FIGURE 3.**  $M_C$  stabilizes binding of G(2'N) and G(3'N) to the  $(E \cdot S)_C$  complex. (A) Effects of  $Mn^{2+}$  on the observed equilibrium association constant for binding of G(2'N) ( $K_{a,obs}^{G(2'N)}$ , ●) and G(3'N) ( $K_{a,obs}^{G(3'N)}$ , ●) to  $(E \cdot S)_C$  at pH 7.7. Data were collected as described in Materials and Methods and are shown in Supplemental Figure S11. The lines are fits of the  $Mn^{2+}$  concentration dependences of  $K_{a,obs}$  for G(2'N) and G(3'N) according to the model shown in B. The y-axes are on different scales to facilitate comparison between the  $Mn^{2+}$  concentration dependence of G(2'N) and G(3'N) binding. (B) Model for binding of the  $Mn^{2+}$  ion that stabilizes binding of G(2'N) and G(3'N) to  $(E \cdot S)_C$ , and  $K_{a,obs}^{Mn}$  report binding of G(2'N) and G(3'N) to  $(E \cdot S)_C$  in the absence and presence of saturating  $Mn^{2+}$ , respectively, and  $K_{Mn,app}^{(E \cdot S)_C}$  reports the apparent affinity of the  $Mn^{2+}$  ion that stabilizes binding of G(2'N) and G(3'N). (C) Model of interactions in  $(E \cdot S \cdot G)_C$  complex. Black circles represent interactions established previously (Shan et al. 1999a, 2001; Shan and Herschlag 1999; Stahley and Strobel 2005) and red circles represent interactions inferred from the data in A.

G binding to  $(E \cdot P)_C$  by 12-fold and sixfold, respectively (Fig. 2C), and  $Mn^{2+}$  stabilizes binding of G(3'NH<sub>2</sub>) to  $(E \cdot P)_C$  by 21-fold. These results are identical to those observed for binding of the G analogs to  $(E \cdot S)_O$  (cf. Fig. 2A,C) and indicate that, within  $(E \cdot P \cdot G)_C$ , the G 2'-OH group makes no interaction and the adjacent 3'-OH contacts a metal ion. As G 3'-OH contacts  $M_A$  in  $(E \cdot S \cdot G)_O$  (Fig. 4D), the simplest expectation is that this interaction is made in the  $(E \cdot P \cdot G)_C$  complex.

### An atomic model for the active site rearrangement in the G binding site upon S docking

Our data indicate that G is bound in a nonfunctional configuration within the  $(E \cdot S \cdot G)_O$  complex (Fig. 4D), as neither of its -OH groups form the metal ion interactions required for catalysis (Fig. 1D,E), and this difference is consistent with a model in which one or more of the active site metal ions are repositioned with respect to G.<sup>13</sup> Understanding this rearrangement at an atomic level requires structures of the  $(E \cdot S \cdot G)_O$  and  $(E \cdot S \cdot G)_C$  complexes, but high-resolution structures of a complete group I intron in the  $(E \cdot S \cdot G)_O$  complex

are not available.<sup>14,15</sup> We therefore utilized a computational approach to model this complex, applying distance constraints from the available functional data along with an energy minimization step to determine what changes may occur when S is not present in the active site.

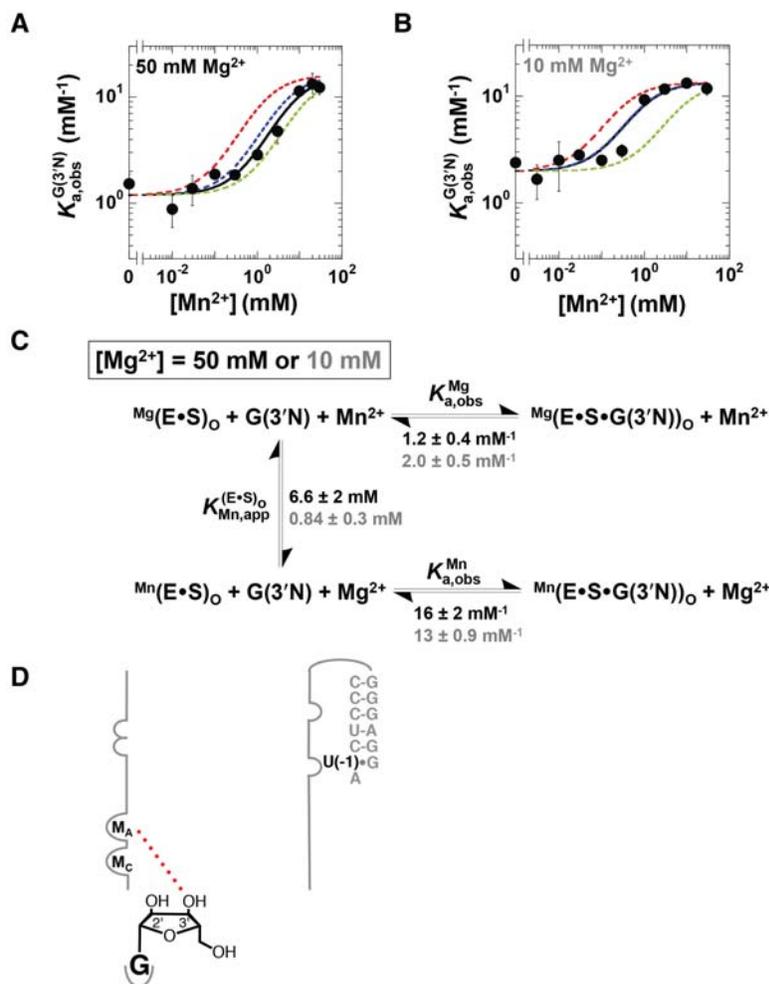
We began with the *Azoarcus*  $(E \cdot P \cdot G)_C$  structure (Lipchock and Strobel 2008) and removed the P1 helix, which contains P, from the active site, so that only G is bound to the ribozyme. We also altered the conformation of the A261 ribose ring to ensure that its 2'-OH group is not donating a hydrogen bond to the G exocyclic amino group, as prior biochemical data provided evidence for the absence of this interaction (Forconi et al. 2010). We then applied harmonic distance restraints based on the available functional data (Supplemental Table S1) to ensure an interaction between  $M_A$  and G 3'-OH (Fig. 4D) and the absence of an interaction with the G 2'-OH group (Fig. 2A; Shan and Herschlag 1999). The resulting model of the active site within  $(E \cdot S \cdot G)_O$  is shown in Figure 5A. For comparison, we also generated a model of the  $(E \cdot S \cdot G)_C$  complex (Fig. 5B) following the same procedures but with both G -OH groups contacting  $M_C$  (Fig. 3C) and a hydrogen bond between the G exocyclic amino group and the A261 2'-OH group (Materials and Methods; Supplemental Table S1; Forconi et al. 2010).

Overlaying the models for  $(E \cdot S \cdot G)_O$  and  $(E \cdot S \cdot G)_C$  allows us to suggest changes that may occur upon S docking into the active site. While we describe these differences below, we also emphasize that the differences are derived from a computational model and do not represent direct experimental observations. The structural overlay (Fig. 5C) reveals that  $M_C$  has moved away from G in the  $(E \cdot S \cdot G)_O$  model and does not form the inner-sphere contacts ( $\leq 2.5 \text{ \AA}$ ) (Schnabl et al. 2012) made with the -OH groups in the  $(E \cdot S \cdot G)_C$  complex. The models also suggest a modest ( $\sim 0.95 \text{ \AA}$ ) change in the position of  $M_A$ . Repositioning of  $M_C$  may arise from the absence of the hydrogen bond between the G exocyclic amino group and the A261 2'-OH group (cf. Fig. 5A,B). Functional data (Forconi et al. 2010) suggest that this hydrogen bond is made in  $(E \cdot S \cdot G)_C$  but not  $(E \cdot S \cdot G)_O$  and we speculate that the absence of this contact releases the adjacent C262 residue, a ligand for  $M_C$  (Houglund et al. 2005). Such a change might allow  $M_C$  to sample alternative positions (Fig. 5C, arrows). This model is consistent with an x-ray structure of

<sup>13</sup>An alternative model in which  $M^{2+}$  binds weakly to site C of  $(E \cdot S \cdot G)_O$  is ruled out by prior work showing that  $Mn^{2+}$  can saturate this metal site within this complex (Shan and Herschlag 1999).

<sup>14</sup>A structure of a truncated form of the *Tetrahymena* group I intron with only G bound in the active site is available (Guo et al. 2004). We did not use this structure to model  $(E \cdot S \cdot G)_O$  because one metal ion, presumably  $M_C$ , is observed in the active site. We speculate that metal ion binding to site A may have been disrupted from mutating A304, a  $M_A$  ligand (Forconi et al. 2008), to G or from removing helices P1, P2, P9.1, and P9.2 from the intron.

<sup>15</sup>Structures of  $(E \cdot P \cdot G)_C$  from *Azoarcus* (Lipchock and Strobel 2008) and Twort (Golden et al. 2005) show both G -OH groups within coordination distance to  $M_C$ , which differs from the interactions we biochemically infer in the *Tetrahymena*  $(E \cdot P \cdot G)_C$  and  $(E \cdot S \cdot G)_O$  complexes. Preliminary data suggest that this difference between *Tetrahymena* and *Azoarcus*, and possibly Twort, may be due to differences in active site preorganization between the ribozymes (S van Schie, R Sengupta, and D Herschlag, unpubl.).



**FIGURE 4.**  $M_A$  stabilizes binding of G(3'N) to the  $(\text{E}\cdot\text{S})_{\text{O}}$  complex. Effects of  $\text{Mn}^{2+}$  on the observed equilibrium association constant for binding of G(3'N) ( $K_{a,obs}^{G(3'N)}$ ) to  $(\text{E}\cdot\text{S})_{\text{O}}$  at pH 7.7 in a background of 50 mM  $\text{Mg}^{2+}$  (A) or 10 mM  $\text{Mg}^{2+}$  (B). The solid lines are fits of the  $\text{Mn}^{2+}$  concentration dependences of  $K_{a,obs}^{G(3'N)}$  according to the model shown in C. The dashed lines correspond to fits of the data with  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{O}}}$  set to the reported  $M_{\text{NA}}$  (blue),  $M_{\text{NB}}$  (green), and  $M_{\text{NC}}$  (red) affinities to  $(\text{E}\cdot\text{S})_{\text{O}}$  (Shan et al. 1999a, 2001). Values of  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{O}}}$  for sites A, B, and C are 4.1, 13, and 1.3 mM, respectively, in the presence of 50 mM  $\text{Mg}^{2+}$  and 0.82, 7, and 1.3 mM, respectively, in the presence of 10 mM  $\text{Mg}^{2+}$  (Shan et al. 1999a). (C) Model for binding of the  $\text{Mn}^{2+}$  ion that stabilizes binding of G(3'N) to  $(\text{E}\cdot\text{S})_{\text{O}}$ . Values obtained in a 50 mM (black) and 10 mM (gray)  $\text{Mg}^{2+}$  background were obtained from fits of the data from A and B, respectively.  $K_{a,obs}^{\text{Mg}}$  and  $K_{a,obs}^{\text{Mn}}$  report binding of G(3'N) to  $(\text{E}\cdot\text{S})_{\text{O}}$  in the absence and presence of saturating  $\text{Mn}^{2+}$ , respectively, and  $K_{\text{Mn,app}}^{(\text{E}\cdot\text{S})_{\text{O}}}$  reports the apparent affinity of the  $\text{Mn}^{2+}$  ion that stabilizes binding of G(3'N) to  $(\text{E}\cdot\text{S})_{\text{O}}$ . (D) Model of interactions in  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{O}}$  complex. Red circles represent interactions inferred from the data in A.

a truncated *Tetrahymena* ribozyme that may mimic the  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{O}}$  complex<sup>6</sup>; in this structure, the positions of C262 and  $M_C$  vary in each of the four molecules in the asymmetric unit, perhaps because the hydrogen bond between A261 and G is not formed (Supplemental Fig. S13; Forconi et al. 2010).

The models for the open and closed complexes also suggest a conformational change of the G ribose ring (Fig. 5C). In  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{O}}$ , the G ribose ring adopts a C3'-endo configuration and in  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$ , a C1'-exo configuration. The C3'-endo configuration is typical for RNA sugars (Saenger 1983), and a

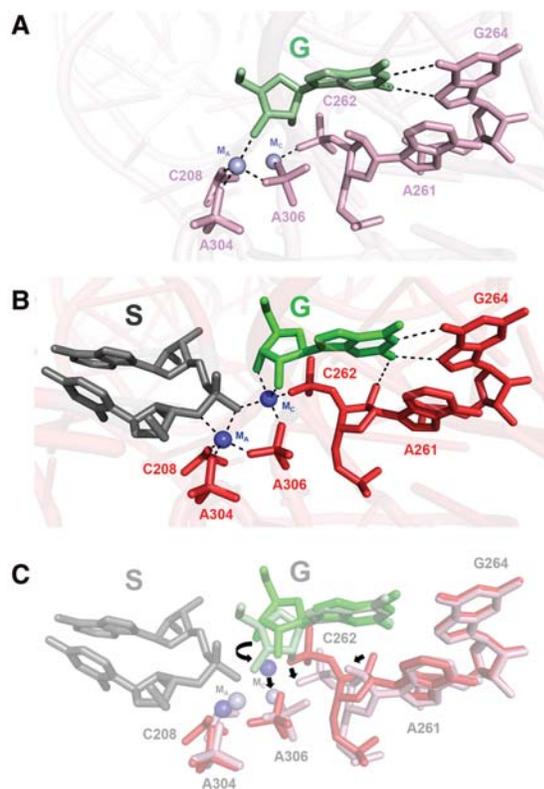
survey of phase angles for all nonhelical G residues within the nonredundant set of all RNA structures in the PDB (Supplemental Fig. S14; Leontis and Zirbel 2012) suggests that the C1'-exo configuration is less prevalent. This survey reveals two peaks, major (~80%) and minor (~20%) corresponding primarily to the C3'-endo (0–36°) and C2'-endo (144–180°) configurations, respectively (Supplemental Fig. S14), with the C1'-exo conformation observed in the  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$  model in the tail end of the minor C2'-endo peak. It may be that the transition to the C1'-exo or a similar C2'-endo conformation is required from the structural constraints of needing the adjacent hydroxyl groups to coordinate a single metal ion within the active site (Supplemental Fig. S15). Future studies that utilize NMR spectroscopy on isotopically labeled G analogs may allow tests of this model and reveal information on the role of sugar pucker in recognition and positioning of G for catalysis.

The distinct active site interactions with G in  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{O}}$  and  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$  arise from the presence of the reactive phosphoryl group (see above and Fig. 2), which contacts  $M_A$  and  $M_C$  in  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$  (Fig. 5B). In the absence of S, these metal ions are not bridged by the *pro*- $S_P$  oxygen atom of the reactive phosphoryl group and are ~5 Å apart in the proposed model (Fig. 5A). When S is present in the active site, the  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$  model and related X-ray structures of the *Azoarcus* group I ribozyme suggest that  $M_A$  and  $M_C$  are ~3.8–4.4 Å apart, the distance observed between active site metal ions in several protein enzymes that catalyze phosphoryl transfer (Steitz and Steitz 1993; Strater et al. 1996). From this distance, both metal ions can contact a nonbridging oxygen atom on the reactive phosphoryl group

and make additional catalytic interactions with the nucleophile and leaving group (Freemont et al. 1988; Kim and Wyckoff 1991; Steitz 1998; Stahley and Strobel 2005; Yang et al. 2006; Toor et al. 2008).

## DISCUSSION

The startling discovery of RNA catalysts in the early 1980s (Kruger et al. 1982; Guerrier-Takada et al. 1983) demonstrated that RNA could be more than just a passive carrier



**FIGURE 5.** Models for active site interactions within  $(E\cdot S\cdot G)_O$  and  $(E\cdot S\cdot G)_C$ . Interactions within  $(E\cdot S\cdot G)_O$  and  $(E\cdot S\cdot G)_C$  are shown in A and B, respectively, and an overlay of the two complexes is shown in C. The structures in C are partially transparent to facilitate comparison. The black arrows in C highlight changes in the positions of active site residues in going from  $(E\cdot S\cdot G)_C$  to  $(E\cdot S\cdot G)_O$ . The models were obtained through molecular modeling using constraints from the functional data obtained here and in prior work (Materials and Methods; Supplemental Table S1).

of genetic information. From a mechanistic point of view, the discovery of ribozymes was surprising because chemical and structural properties of RNA were expected to hinder its ability to serve as an efficient catalyst (Herschlag 1995; Narlikar and Herschlag 1997; Russell 2008). Nevertheless, it appears likely that RNA played an extensive catalytic role early in evolution (Gilbert 1986; Robertson and Joyce 2012).

Extant catalytic RNAs in Nature, including the ribosome and spliceosome, seem to utilize conformational changes in function (e.g., Chanfreau and Jacquier 1996; Staley and Guthrie 1998; Wang et al. 1999; Zhuang et al. 2002b; Schmeing et al. 2005; Hougland et al. 2006; Martick and Scott 2006; Guo et al. 2009; Hsieh and Fierke 2009; Marcia and Pyle 2012; Voorhees and Ramakrishnan 2013; Chen and Moore 2014; Sripathi et al. 2014), and an intrinsically rugged conformational landscape (e.g., Chen and Dill 2000; Treiber and Williamson 2001; Bokinsky and Zhuang 2005; Shcherbakova et al. 2008) may predispose RNA to be able to toggle between conformational states and thereby carry out such complex functions (Herschlag 1995; Munro et al. 2009; Frank and Gonzalez 2010). To learn more about the

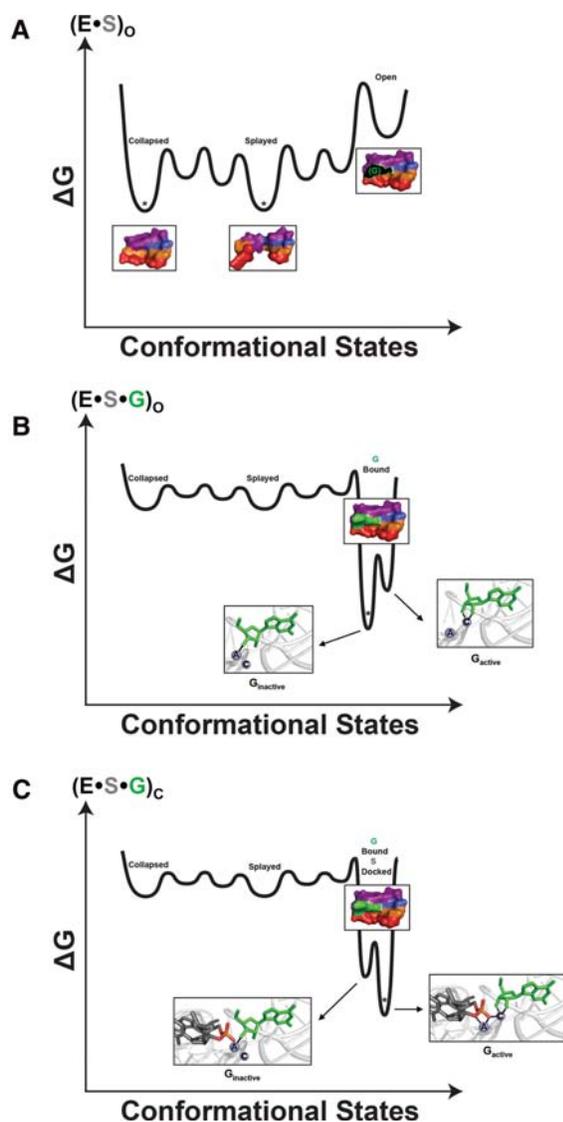
properties of conformational states and conformational changes within otherwise folded RNAs, we investigated the G binding site of the *Tetrahymena* group I ribozyme, a system that has been a workhorse for uncovering RNA structure–function relationships (Hougland et al. 2006).

### Molecular recognition of the G nucleophile by the *Tetrahymena* ribozyme

Although catalysis hinges upon the establishment of a precise network of interacting groups, our results indicate that alternative interactions are made within the active site of the *Tetrahymena* ribozyme and that structural rearrangements are necessary to form a functional catalytic core. These results highlight RNA's tendency to form alternative interactions, even within the catalytic core of the folded ribozyme. Below, we provide a model of interactions made in and around the G binding site of the G-free form of the *Tetrahymena* ribozyme,  $(E\cdot S)_O$ , and the  $(E\cdot S\cdot G)_O$  and  $(E\cdot S\cdot G)_C$  complexes. This model highlights that native interactions with G are formed in the  $(E\cdot S\cdot G)_C$  complex only after releasing nonproductive contacts that are present in the  $(E\cdot S\cdot G)_O$  complex.

The rate constant for G association ( $\sim 10^3 \text{ M}^{-1} \text{ sec}^{-1}$ ) is several orders of magnitude slower than diffusion (Karbstein and Herschlag 2003), consistent with a model in which the G binding site exists in a configurations that cannot accommodate G and an alternative “open” configuration that can bind G (Hougland et al. 2006; Benz-Moy and Herschlag 2011). Multiple structures of group I introns with bound G are available (Adams et al. 2004; Guo et al. 2004; Golden et al. 2005; Stahley and Strobel 2005; Lipchock and Strobel 2008), and all show G forming a base triple interaction with a C–G pair within a “sandwich” of base triples (Supplemental Fig. S16A). In contrast, structures of the G-free form of the intron have not been solved, perhaps because the greater dynamics of the G binding site renders crystallization of the RNA more difficult. Below, we present and evaluate models for the G binding site upon removal of G (Fig. 6A; Supplemental Fig. S16).

By definition, there must be a G-accessible “open” state that allows G binding. We model this state by simply removing G from its binding site of the *Tetrahymena* group I intron (Fig. 6A; Supplemental Fig. S16B). However, slow binding of G to the ribozyme implies the prevalence of one or more alternative, G-inaccessible states, and we present two models for such states. Prior work modeled a “collapsed” form of the G binding site by manipulating the position of the top base triple within the G binding site to prevent accommodation of G (Fig. 6A; see also Supplemental Fig. S16C; Benz-Moy and Herschlag 2011). Whereas this model posits the simplest motion upon removal of bound G, we know of no example in the PDB of an RNA structure in which a base “angles” or collapses in this way. Such a structure would suffer from limited stacking and van der Waals overlap and poor hydrogen bonds.



**FIGURE 6.** Multiple conformational states within the G binding site. Qualitative free energy landscapes of the G binding site within  $(E\cdot S)_O$  (A),  $(E\cdot S\cdot G)_O$  (B), and  $(E\cdot S\cdot G)_C$  (C). (A) There is strong evidence that the open state is not the preferred state at equilibrium (Karbstein and Herschlag 2003) but we do not know whether the splayed and/or collapsed states or an alternative state or ensemble of states is favored (see text). For simplicity, the splayed and collapsed states are reported as the most stable species and are equal in free energy; these states are noted with asterisks. (B) Binding of G (green) stabilizes the open state. G is bound in two distinct conformations, referred to as  $G_{active}$ , which forms the catalytically relevant interactions, and  $G_{inactive}$ , an inactive form. Structures of these states, obtained from Figure 5, are shown, with G and the active site metal ions colored light blue.  $G_{inactive}$  is the preferred state in  $(E\cdot S\cdot G)_O$  and is thus lower in free energy (denoted by asterisk). (C) Within  $(E\cdot S\cdot G)_C$ , S (gray, with the reactive phosphoryl in orange and red) reorganizes the active site so that  $G_{active}$  is lower in free energy and thus the preferred state (denoted by asterisk).

We also consider a “splayed” form of the G binding site, in which none of the base triples are formed, as was obtained from a recent structure of the lariat capping (LC) ribozyme, an RNA that contains a G binding site with a secondary struc-

ture that matches that of the group I intron (Fig. 6A; Supplemental Fig. S16D; Meyer et al. 2014). However, metal ions analogous to  $M_A$  and  $M_C$  were not observed in the LC ribozyme, and the putative ligands of these metal ions were not positioned to ligate a common metal ion. Functional experiments with the *Tetrahymena* ribozyme strongly suggest that these metal ions are stably bound in the free ribozyme (Shan et al. 1999a; Shan and Herschlag 1999), consistent with a model in which the conformational rearrangement for this ribozyme is less dramatic than that seen for the LC ribozyme. Conformational changes upon ligand binding that are less dramatic than that inferred for G binding to the LC ribozyme but nevertheless involve base flipping and alternative stacking interactions have been observed for other RNAs including Spinach, an in vitro-selected RNA aptamer (Huang et al. 2014), and the purine riboswitch aptamer domain (Gilbert et al. 2006).

Binding of G stabilizes the open form of the G binding site (Fig. 6B). Nevertheless, our data indicate that there are (at least) two distinct G-bound conformers.  $G_{active}$  and  $G_{inactive}$  are represented by two wells in the free energy landscapes shown in Figure 6B and C and correspond to the preferred states of G within  $(E\cdot S\cdot G)_O$  and  $(E\cdot S\cdot G)_C$ . Within  $(E\cdot S\cdot G)_O$ , the G 3'-OH group contacts  $M_A$ , whereas the adjacent 2'-OH makes no interaction; these interactions differ from those made in the transition state. We therefore refer to this state, the most stable form of the  $(E\cdot S\cdot G)_O$  complex, as  $G_{inactive}$  (Fig. 6B). Our model of  $(E\cdot S\cdot G)_O$  (Fig. 5) suggests that this alternate binding mode arises from the absence of a hydrogen bond interaction between the G exocyclic amino group and the A261 2'-OH group, which releases the adjacent C262 residue, a ligand for  $M_C$ , and allows  $M_C$  to move away from G. In addition, a small motion of  $M_A$  allows it to interact with the 3'-OH group of G while the ribose ring of G is in its preferred sugar pucker (Saenger 1983).

S binding promotes active site rearrangements that favor a catalytically competent configuration (i.e.,  $G_{active}$ , Fig. 6B,C). Our models (Fig. 5) allow us to propose what changes may occur within the active site.  $M_A$  and  $M_C$  move to interact with the reactive phosphoryl group of S. The repositioning of  $M_C$  is accompanied by changes in the position of one of its ligand, the *pro*- $S_P$  phosphoryl oxygen of C262. This rearrangement, in turn, allows the 2'-OH of the adjacent A261 to make its interaction with the exocyclic amino group of G (Fig. 5). Finally, there is a change in the conformation of the G ribose ring so that its 2'- and 3'-OH groups contact  $M_C$ .

### Interactions formed in the transition state

Once S and G are bound to the ribozyme, the G 3'-OH group is deprotonated and reacts with the reactive phosphoryl group on U(-1) (Herschlag 1995; Karbstein et al. 2002). In the transition state, the U(-1) and G 3'-oxygen atoms form partial bonds with the reactive phosphoryl group and are stabilized through metal ion interactions with  $M_A$  and  $M_C$  (Piccirilli

et al. 1993; Weinstein et al. 1997; Shan et al. 1999a). As shown in Figure 1D,  $M_A$  contacts the U(-1) 3'-oxygen atom,  $M_C$  contacts the G 2'- and 3'-oxygen atoms, and both metals contact the *pro-S\_P* phosphoryl oxygen. In addition, TFA experiments suggested that a third metal ion,  $M_B$ , supplants the role of  $M_C$  in contacting the 3'-oxygen atom of G (Fig. 1E), but this metal is not observed in any of the group I structures (Shan et al. 1999a; Vicens and Cech 2006; Lipchock and Strobel 2008). One possible explanation for this difference is that  $M_C$  contacts the 3'-oxygen atom in the transition state for the normal reaction but not in TFA experiments that substitute this oxygen atom with a sulfur atom. The bulkier sulfur atom needed for the TFA experiments may prevent simultaneous formation of the sterically constrained contacts to the 2'- and 3'- groups of G. This may result in recruitment of a third metal ion,  $M_B$ , that normally does not take part in the reaction, to make the interaction with the 3'-oxygen of G (Houglund et al. 2006).

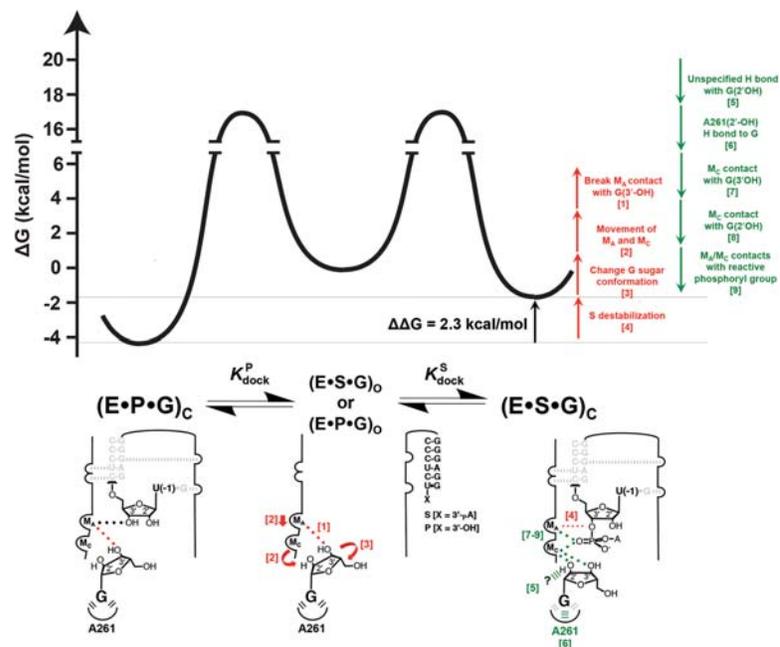
The results described here demonstrate that  $M_C$  can contact the G 3'-oxygen atom within the  $(E \cdot S \cdot G)_C$  complex, a result which is consistent with and most simply interpreted in terms of the recruitment model, although direct experimental data distinguishing these models are still needed. Finally, while it is possible that no interaction is made with the G 3'-OH group within  $(E \cdot S \cdot G)_O$  and that rescue experiments with G(3'-NH<sub>2</sub>) recruit  $M_A$  to contact the 3'-NH<sub>2</sub> group, this alternative model is unlikely given that the 3'-H substitution destabilizes G binding to  $(E \cdot S)_O$  (Fig. 2).

### Energetics of active site interactions: substrate destabilization revisited

Our results provide new information on "substrate destabilization" in the *Tetrahymena* ribozyme reaction. Both the substrate and product oligonucleotides (S and P, respectively, Fig. 1A) hybridize to the same internal guide sequence on the ribozyme and utilize identical tertiary interactions to dock into the active site, but additional interactions are made between the reactive phosphoryl group present on S and the catalytic metal ions (cf. schemes in Fig. 7). Despite this difference, docking of P into the active site is 2.3 kcal mol<sup>-1</sup> more stable than S (Fig. 7; Narlikar et al. 1995). Our results indicate that the reactive phosphoryl group on S promotes a rearrangement within the active site of  $(E \cdot S \cdot G)_C$  and provide new perspective about this paradoxical observation.

The weaker binding of S relative to P was suggested to result from an unfavorable interaction between the electron-deficient U(-1) 3'-bridging phosphoryl oxygen and the electropositive Mg<sup>2+</sup> ion at site A, relative to the interaction of that Mg<sup>2+</sup> ion with the 3'-OH of P (Narlikar et al. 1995). Lower electron density on the U(-1) 3'-oxygen of S relative to P might be expected because of the electron-withdrawing phosphoryl group that is attached in S (Fig. 7, [4]). This model is consistent with recent metal ion rescue experiments that provided evidence for negative charge stabilization of the U(-1) 3'-atom by  $M_A$  (Sengupta et al. 2012).

Our data reveal additional penalties upon S docking: the nonproductive contact between  $M_A$  and the G 3'-OH group within  $(E \cdot S \cdot G)_O$  (Fig. 7, [1]) is broken to allow formation of the catalytic interactions with G (Fig. 7, [7,8]); the G ribose ring changes to a less common conformation (Fig. 7, [3]) to allow both its 2'- and 3'-OH groups to contact  $M_C$ , and the repositioned active site metal ions are likely to experience electrostatic repulsion from the shortened distance between them (Fig. 7, [2]).



**FIGURE 7.** Model for differential stabilization of S and P docking. Free energy profile for docking of P and S into the active site. Docking of P is 2.3 kcal/mol more stable than docking of S (Narlikar et al. 1995). A scheme for interactions made within each complex is shown below the free energy profile, with dots and hatched lines representing metal ion interactions and hydrogen bonds, respectively. Unfavorable (1–4) and favorable (5–9) interactions made in  $(E \cdot S \cdot G)_C$  are represented by the upward red arrows and downward green arrows, respectively. The energetic contribution of each of these interactions could not be determined but, for simplicity, the height of each of these arrows is presented as the same. Equilibrium constants for docking ( $K_{\text{dock}}^P = 600$  and  $K_{\text{dock}}^S = 16$ ) were determined from prior work (Narlikar et al. 1995), the rate constants for docking of S and P was assumed to be the same ( $\sim 17 \text{ sec}^{-1}$ ) (Narlikar et al. 1999), the rate constants for undocking were determined from the ratio  $k_{\text{dock}}/K_{\text{dock}}$ , and free energy differences were obtained from the rate ( $k$ , in units of  $\text{sec}^{-1}$ ) and equilibrium constants ( $K_{\text{dock}}$ , unitless) using the standard conversions:  $\Delta G^\ddagger = -RT \ln(kh/k_B T)$  and  $\Delta G = -RT \ln K_{\text{dock}}$ , where  $R = 1.987 \text{ cal}/(\text{mol} \cdot \text{K})$ ,  $T = 323 \text{ K}$  (50°C),  $h = 1.58 \times 10^{-34} \text{ cal s}$ , and  $k_B = 3.30 \times 10^{-24} \text{ cal K}^{-1}$ .

Despite these unfavorable energetic effects, changes occur upon S (but not P) docking. This observation indicates that there are favorable interactions that allow the active site within  $(E \cdot S \cdot G)_C$  to prefer the “active” state. For example, the active state can be stabilized with docked S by interactions that lead to catalysis (Fig. 7, [5–9])—i.e., interactions of the *pro*- $S_P$  phosphoryl oxygen atom with both  $M_A$  and  $M_C$  (Shan et al. 2001); interactions of the 2′- and 3′-OH groups of G with  $M_C$ ; an interaction of the 2′-OH of A261 with the exocyclic amino group of the guanine base of G that is allowed because of the repositioning of  $M_C$  (Forconi et al. 2010); and an unspecified hydrogen bond donated by the 2′-OH group of G (Hougland et al. 2008). In addition, the inactive state may be destabilized by the reactive phosphoryl group because of a steric clash with the 3′-OH of G (Fig. 5C). This model highlights the complexity of active site energetics, especially when binding events are coupled to conformational rearrangements.

### General implications for RNA in biology

Our results highlight RNA’s tendency to form alternative structures, even within the active site of an RNA that has completed overall folding. Despite the apparent difficulties that RNA faces in specifying a unique structure that would be optimal for catalysis, RNA enzymes have persisted to the modern day, and several of them, including the spliceosome and ribosome, coordinate complex multistep processes. Perhaps “problems” RNA has in folding have been harnessed by Nature to evolve dynamic ribonucleoproteins such as the ribosome and spliceosome, which undergo local and global conformational changes during their functional cycles.

## MATERIALS AND METHODS

### Materials

L-21 ScaI ribozyme (E) was transcribed and gel purified according to reported procedures (Zaug et al. 1988). Phosphoramidites for guanosine (G), 2′-deoxyguanosine [G(2′H)], and 3′-deoxyguanosine [G(3′H)] were purchased from Glen Research and incorporated into the oligonucleotide AUCG<sub>X</sub> [where G<sub>X</sub> denotes G, G(2′H), or G(3′H)] (Table 1) through solid phase synthesis at the Protein and Nucleic Acid Facility at Stanford University. AUCG, AUCG(2′H), and AUCG(3′H) were subsequently HPLC-purified, as described below. G (≥98%) was purchased from Sigma-Aldrich. 2′-aminoguanosine [G(2′N)] and 3′-aminoguanosine [G(2′N)] were purchased from Santa Cruz Biotechnology and were of the highest purity commercially available (≥98%). All other oligonucleotides were purchased from Dharmacon Inc. or Integrated DNA Technologies and purified through HPLC or gel purification, as described below.

### HPLC purification of oligonucleotides

AUCG, AUCG(2′H), and AUCG(3′H) were HPLC-purified on an Lichrosorb RP-18 Column (Alltech-Applied Science) using a linear

gradient of acetonitrile in 0.1 M triethylammonium acetate buffer (pH 6.0). Triethylammonium acetate was removed from HPLC-purified samples by repeated drying and aqueous resuspension cycles in a Savant Speedvac Concentrator (Thermo Scientific). After resuspension in water, the concentrations of the AUCG<sub>X</sub> analogs were determined by measuring the UV absorbance at 260 nm with an extinction coefficient of 41,900 M<sup>-1</sup> cm<sup>-1</sup>. The oligonucleotide -1d,rP (Table 1) was purified on a DNAPac PA-100 column (Dionex) using a linear gradient of 1 M NaCl in 20 mM Tris-HCl buffer (pH 7.6), and subsequently desalted using a Sep-Pak (Waters). After drying and resuspension in water, the concentration of -1d,rP was determined by measuring the UV absorbance at 260 nm with an extinction coefficient of 46,500 M<sup>-1</sup> cm<sup>-1</sup>.

### 5′-<sup>32</sup>P end labeling of oligonucleotides

Oligonucleotide substrates for kinetic experiments were 5′-<sup>32</sup>P-radiolabeled using [ $\gamma$ -<sup>32</sup>P]ATP (MP Biomedicals) and T4 polynucleotide kinase (New England Biolabs), according to the manufacturer’s protocol. Labeled substrates were subsequently purified through nondenaturing polyacrylamide gel electrophoresis, as previously described (Herschlag et al. 1993).

### General reaction conditions

All reactions were single-turnover, with ribozyme in excess of labeled substrate, and were measured at 30°C in the presence of 50 mM MgCl<sub>2</sub> and 50 mM buffer. The buffers used were the following: sodium acetate (pH 5.0–5.5), NaMES (pH 6.1–6.7), NaMOPS (pH 7.1), NaEPPS (pH 7.7–8.2), and NaCHES (pH 8.7–9.7).

Reactions were carried out and analyzed according to reported procedures (Herschlag and Cech 1990; Shan and Herschlag 1999). Ribozyme was allowed to fold in 10 mM MgCl<sub>2</sub> and 50 mM buffer at 50°C for 30 min and subsequently cooled to room temperature. For reactions carried out above pH 8.0, the folding step was performed in 25 mM NaMES, pH 6.7 to avoid degradation of the ribozyme. Following the folding step, ribozyme was diluted 20-fold into reaction tubes containing the desired concentrations of divalent metal ion (MgCl<sub>2</sub> and MnCl<sub>2</sub>), buffer, and G analog. After a 5-min incubation at 30°C, reactions were initiated by the addition of labeled substrate (<0.1 nM). At specified times, six 2  $\mu$ L aliquots of the reaction mixture were removed from the 20  $\mu$ L reaction and added to a 4  $\mu$ L quench solution containing 90% formamide, 50 mM EDTA, 0.01% bromophenol blue, and 0.01% xylene cyanol. Substrate and product were separated by electrophoresis on a 20% polyacrylamide gel containing 7 M urea, 100 mM Tris, 83 mM boric acid, and 1 mM EDTA. The ratio of substrate to product was quantitated through Phosphorimager analysis (GE Healthcare) with TotalLab (TotalLab Ltd).

Reactions were followed for  $\geq 3t_{1/2}$  except for very slow reactions. First-order fits ( $R^2 > 0.95$ ) to the data points, with end points  $\geq 90\%$  were obtained (KaleidaGraph, Synergy Software). The slow reactions were typically linear for up to 20 h, and an end point of 95% was assumed to obtain observed rate constants from the initial rates

### Periodate treatment of deoxyguanosine analogs

Prior to initiation of the reactions, AUCG(2′H) and AUCG(3′H) were pretreated with 12.5  $\mu$ M sodium periodate for 1 h at room

temperature to convert the *cis*-diol of any guanosine contaminant to an unreactive dialdehyde form. Subsequently, the excess sodium periodate was inactivated by adding 100  $\mu\text{M}$  ethylene glycol for a 1 h at room temperature, and the resultant mixture was added to reaction tubes for kinetic experiments. Control experiments demonstrated that the sodium periodate/ethylene glycol treatment did not affect ribozyme-mediated cleavage activity (data not shown).

### Measurement of G, AUCG, and G(2'N) affinities to (E•S)<sub>O</sub>

Binding of G, AUCG and G(2'N) to the (E•S)<sub>O</sub> complex ( $(K_d^{G_x})_O$ , Fig. 1A) was determined by measuring the observed rate of cleavage ( $k_{\text{obs}}$ ) of 5'-end labeled -3m,rS<sub>Me</sub> (Table 1) at different G analog ( $G_x$ ) concentrations (0–2.1 mM G or G(2'N) and 0–30  $\mu\text{M}$  AUCG), under conditions where E is saturating with respect to S ([E] = 50 nM;  $K_d^S \sim 1$  nM (Shan and Herschlag 2000; Karbstein et al. 2002)). The -3m,rS<sub>Me</sub> substrate contains a 2'-OCH<sub>3</sub> substitution at the U(-3) position that disrupts a tertiary interaction and results in a stable open complex (Narlikar and Herschlag 1996). In addition, the A(+1) residue is replaced with an OCH<sub>3</sub> group to ablate a stabilizing interaction between A(+1) and a Mn<sup>2+</sup> ion (Shan and Herschlag 2000; Karbstein et al. 2002), ensuring that we monitor binding of  $G_x$  to (E•S)<sub>O</sub> and not (E•S)<sub>C</sub> in experiments where we vary the Mn<sup>2+</sup> concentration. In all experiments,  $k_{\text{obs}}$  was plotted as a function of [ $G_x$ ] and fit by equation 1:

$$k_{\text{obs}} = \frac{k_{\text{max}}[G_x]}{[G_x] + K_{1/2}^{G_x}} \quad (1)$$

For  $K_{1/2}^{G_x}$ , the concentration of  $G_x$  that provides half the maximal rate ( $k_{\text{max}}$ ), to be equal to  $(K_d^{G_x})_O$ . The chemical step must be rate-limiting. To determine whether this condition is met, we measured the pH-dependence of  $k_{\text{obs}}$  and observed a log-linear pH-rate dependence with a slope of 1 up to pH 8.2 (data not shown). This pH-dependence is consistent with prior work that the chemical step is the rate-limiting step (Herschlag and Khosla 1994).

### Measurement of G, AUCG, and G(2'N) affinities to (E•S)<sub>C</sub>

The binding affinities of G, AUCG, and G(2'N) to the (E•S)<sub>C</sub> complex ( $(K_d^{G_x})_C$ , Fig. 1A) were determined by measuring the observed rate of cleavage ( $k_{\text{obs}}$ ) of 5'-end labeled -1d,rSA (Table 1) at different G analog ( $G_x$ ) concentrations (see above for range of concentrations), under conditions where E is saturating with respect to S ([E] = 50 nM;  $K_d^S \sim 1$  nM) (Karbstein et al. 2002).  $k_{\text{obs}}$  was plotted as a function of [ $G_x$ ] and the data were fit to equation 1 to obtain  $K_{1/2}^{G_x}$ . The -1d,rSA substrate, which favors formation of the (E•S)<sub>C</sub> complex (Bartley et al. 2003), contains a 2'-H substitution at the U(-1) position that renders the chemical step ( $k_c$ , Fig. 1A) rate-limiting, ensuring that  $K_{1/2}^{G_x}$  is equal to  $(K_d^{G_x})_C$  (McConnell et al. 1993; Herschlag and Khosla 1994; Knitt and Herschlag 1996).

### Measurement of AUCG(2'H), AUCG(3'H) and G(3'N) affinities to (E•S)<sub>O</sub>

We did not observe any detectable cleavage activity with AUCG (2'H), AUCG(3'H) and G(3'N) (data not shown) and thus measured binding of these analogs through competitive inhibition of

the reaction, (E•S)<sub>O</sub> + G → Products. Experiments were carried out using 5'-end labeled -3m,rS<sub>Me</sub> (Table 1) with E saturating with respect to S ([E] = 50 nM;  $K_d^S = 13$  nM) and with subsaturating G ([G] = 30  $\mu\text{M}$ ;  $(K_d^G)_O \geq 440$   $\mu\text{M}$ ). The concentration of inhibitor,  $G_x$ , was varied (0–220  $\mu\text{M}$  AUCG(2'H), 0–365  $\mu\text{M}$  AUCG(3'H), and 0–3.5 mM G(3'N)) and the inhibition constant,  $K_i$ , which reports the affinity of  $G_x$  to (E•S)<sub>O</sub> ( $(K_d^{G_x})_O$ ), was determined through Equation 2:

$$k_{\text{obs}} = \frac{k_{\text{max}}K_i}{K_i + [G_x]} \quad (2)$$

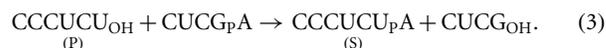
where  $k_{\text{obs}}$  is the observed rate of cleavage of -3m,rS<sub>Me</sub> and  $k_{\text{max}}$  is  $k_{\text{obs}}$  in the absence of inhibitor.

### Measurement of AUCG(2'H), AUCG(3'H), and G(3'N) affinities to (E•S)<sub>C</sub>

The binding affinities of AUCG(2'H), AUCG(3'H), and G(3'N) to the (E•S)<sub>C</sub> complex ( $(K_d^{G_x})_C$ , Fig. 1A) were determined through competitive inhibition of the reaction: (E•S)<sub>C</sub> + G → Products. Measurements were made using 5'-end labeled -1d,rSA, with ribozyme saturating with respect to S ([E] = 50 nM;  $K_d^S < 1$  nM) and with subsaturating G ([G] = 10  $\mu\text{M}$ ;  $(K_d^G)_C \geq 63$   $\mu\text{M}$ ) so that the observed  $K_i$  (Equation 3) is equal to  $(K_d^{G_x})_C$ .

### Measurement of G analog affinities to (E•P)<sub>C</sub>

The binding affinities of G, G(2'N), G(3'N), AUCG, AUCG(2'H), and AUCG(3'H) to (E•P)<sub>C</sub> were determined through competitive inhibition of the reverse reaction: (E•P)<sub>C</sub> + CUCGA → Products. The reverse reaction involves cleavage of a specific phosphodiester bond on CUCGA by P (Table 1), as shown in Equation 3:



As G and CUCGA occupy the same binding site on the ribozyme (Been and Perrotta 1991), the G analogs used in this study can bind to (E•P)<sub>C</sub> and thus prevent binding of CUCGA. With E•P subsaturating with respect to 5'-end labeled CUCGA (<0.1 nM CUCGA, 0.2  $\mu\text{M}$  E•P;  $K_d^{\text{CUCGA}} \geq 1$   $\mu\text{M}$ , data not shown), the concentration of G analog,  $G_x$ , was varied (see above for range of concentrations) and an inhibition constant,  $K_i$ , which reports the affinity of  $G_x$  to (E•P)<sub>C</sub>, was determined using Equation 2.

### Following binding of G(2'NH<sub>2</sub>) and G(3'NH<sub>2</sub>) to the ribozyme

The 2'-NH<sub>2</sub> and 3'-NH<sub>2</sub> groups of G(2'N) and G(3'N), respectively, can ionize to form the corresponding -NH<sub>3</sub><sup>+</sup> species (Aurup et al. 1994; Dai et al. 2007). To conduct metal ion rescue experiments with G(2'NH<sub>2</sub>) and G(3'NH<sub>2</sub>) (and not G(2'NH<sub>3</sub><sup>+</sup>) and G(3'NH<sub>3</sub><sup>+</sup>)), we established conditions under which we were able to obtain affinities for the -NH<sub>2</sub> forms of G(2'N) and G(3'N) to (E•S)<sub>O</sub>, (E•S)<sub>C</sub>, (E•P)<sub>C</sub>, as described below.

#### Following binding of G(2'NH<sub>2</sub>)

Prior work showed that G(2'NH<sub>3</sub><sup>+</sup>) deprotonates with a pK<sub>a</sub> of 6.1 in aqueous solution and with pK<sub>a</sub>s of 7.0 and 7.5 in the (E•G(2'N)) and (E•S•G(2'N))<sub>C</sub> complexes, respectively (10 mM MgCl<sub>2</sub>, 30°C) (Shan

et al. 1999b). In addition, the presence of 5 mM  $\text{MnCl}_2$  lowers the  $\text{pK}_a$  of G(2'N) in the  $(\text{E}\cdot\text{S}\cdot\text{G}(2'\text{N}))_{\text{C}}$  below 5. In our experiments (50 mM  $\text{MgCl}_2$ , 30°C), we observed that binding of G(2'N) to the ribozyme does not vary with pH above pH 8.2 (data not shown), suggesting that binding of the 2'- $\text{NH}_2$  form of G(2'N) was being followed. Therefore, to measure the affinity of G(2'NH<sub>2</sub>) to the ribozyme, we made our measurements with G(2'N) at pH 8.2 (Supplemental Fig. S4).

#### Following binding of G(3'NH<sub>2</sub>)

G(3'NH<sub>3</sub><sup>+</sup>) deprotonates with a  $\text{pK}_a$  of 7.0 in aqueous solution (Dai et al. 2007) and with  $\text{pK}_a$ s of 8.2 and 8.8 in the  $(\text{E}\cdot\text{S}\cdot\text{G}(3'\text{N}))_{\text{O}}$  and  $(\text{E}\cdot\text{S}\cdot\text{G}(3'\text{N}))_{\text{C}}$  complexes, respectively (Supplemental Figs. S5, S7; 50 mM  $\text{MgCl}_2$ , 30°C). The addition of 10 mM  $\text{MnCl}_2$  lowers the  $\text{pK}_a$  of G(3'N) in the  $(\text{E}\cdot\text{S}\cdot\text{G}(3'\text{N}))_{\text{O}}$  and  $(\text{E}\cdot\text{S}\cdot\text{G}(3'\text{N}))_{\text{C}}$  complexes to 7.2 and 7.3, respectively (Supplemental Figs S8, S6). Unlike G(2'N), the binding affinity of G(3'N) continues to decrease above pH 8.2 in  $\text{MgCl}_2$  alone, suggesting that  $K_{a,\text{obs}}^{\text{G3N}} (= 1/K_{d,\text{obs}}^{\text{G3N}})$ , the association constant for G(3'N) binding to ribozyme, reports binding of both G(3'NH<sub>2</sub>) and G(3'NH<sub>3</sub><sup>+</sup>) (Supplemental Figs S5, S7, S9). To determine  $K_a^{\text{G3NH}_2}$ , the association constant of G(3'NH<sub>2</sub>) to the ribozyme, we measured the pH-dependence of G(3'N) binding and fit the data according to model shown in Supplemental Figures S5–S10.

### Modeling the active site of the *Tetrahymena* ribozyme

We modeled the active site within the  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{O}}$  and  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$  complexes of the *Tetrahymena* ribozyme using functional data presented here and in previous work (Supplemental Table S1). Hydrogen bond and metal ion interactions inferred from functional data with the *Tetrahymena* ribozyme were translated into structural distance constraints by using harmonic potentials. For interactions not made within these complexes (Supplemental Table S1), we applied a repulsive potential to prevent these residues from coming within hydrogen bonding distance or inner-sphere coordination to each other. We applied the distance restraints to crystal structures of the *Azoarcus* group I intron. Only residues that were functional constrained were allowed to move; the remaining residues were kept at their initial positions from the X-ray structure. All calculations were carried out using the AMBER suite of programs (Case et al. 2014) and used the ff10 force field for nucleic acids (Wang et al. 2000; Pérez et al. 2007; Banáš et al. 2010; Zgarbová et al. 2011). Screening due to solvent and ions was treated using a variant of the Generalized Born model (Tsui and Case 2000). Refinement of the models using distance restraints was carried out by alternating conjugate gradient energy minimization and simulated annealing cycles. Conjugate gradient energy minimization was carried out for 3000 steps. During the simulated annealing stage, the temperature of the system was raised to 600K and slowly scaled to 0K for 20,000 steps of molecular dynamics.

Before applying the distance restraints, we introduced modifications to X-ray structures of the *Azoarcus* group I intron to represent different states of the catalytic cycle. For the  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{O}}$  complex, we removed the P1 domain from the ribozyme (PDB accession number 3B02) (Lipchok and Strobel 2008). In addition, to ensure that the A261 2'-OH group was not contacting the G exocyclic amino group (Forconi et al. 2010), we allowed the A261 ribose ring to adopt a configuration where its 2'-OH group was pointed away from G, as observed in the group I structure from Twort (Golden et al.

2005). For the  $(\text{E}\cdot\text{S}\cdot\text{G})_{\text{C}}$  complex, we transferred the reactive phosphoryl group from G206 to U(-1) within a structure mimicking the  $(\text{E}\cdot\text{P}\cdot\text{G}_{\text{pA}})_{\text{C}}$  complex (PDB accession number 3B02) (Stahley and Strobel 2005).

### Determining the prevalence of the G ribose ring conformations within an RNA structure database

To determine the distribution of G ribose ring conformations with the PDB (Supplemental Fig. S14), we downloaded the list of non-redundant set of RNA structures (Leontis and Zirbel 2012). The phase angle (Chen et al. 2010) for each G residue within this set was obtained using 3DNA (Lu and Olson 2003). We also used 3DNA to exclude G residues that are in Watson–Crick or G•U base pairs. Inclusion of helical G residues did not significantly affect the distribution of phase angles (data not shown).

### SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

### ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes of Health (NIH) to D.H. (GM049243) and a NIH training grant to R.N.S. (5 T32 GM007276). We thank Charles Lea for initial results, and Fang-Chieh Chou, Rhiju Das, Mark Krasnow, and members of the Herschlag laboratory for helpful discussions.

Received July 31, 2015; accepted September 23, 2015.

### REFERENCES

- Abramovitz DL, Friedman RA, Pyle AM. 1996. Catalytic role of 2'-hydroxyl groups within a group II intron active site. *Science* **271**: 1410–1413.
- Adams PL, Stahley MR, Kosek AB, Wang J, Strobel SA. 2004. Crystal structure of a self-splicing group I intron with both exons. *Nature* **430**: 45–50.
- Aurup H, Tuschl T, Benseler F, Ludwig J, Eckstein F. 1994. Oligonucleotide duplexes containing 2'-amino-2'-deoxycytidines: thermal stability and chemical reactivity. *Nucleic Acids Res* **22**: 20–24.
- Banáš P, Hollas D, Zgarbová M, Jurečka P, Orozco M, Cheatham TE, Šponer J, Otyepka M. 2010. Performance of molecular mechanics force fields for RNA simulations: stability of UUCG and GNRA hairpins. *J Chem Theory Comput* **6**: 3836–3849.
- Bartley LE, Zhuang X, Das R, Chu S, Herschlag D. 2003. Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J Mol Biol* **328**: 1011–1026.
- Bass BL, Cech TR. 1984. Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: implications for biological catalysis by RNA. *Nature* **308**: 820–826.
- Bass BL, Cech TR. 1986. Ribozyme inhibitors: deoxyguanosine and dideoxyguanosine are competitive inhibitors of self-splicing of the *Tetrahymena* ribosomal ribonucleic acid precursor. *Biochemistry* **25**: 4473–4477.
- Been MD, Perrotta AT. 1991. Group I intron self-splicing with adenosine: evidence for a single nucleoside-binding site. *Science* **252**: 434–437.
- Bennett WS Jr, Steitz TA. 1978. Glucose-induced conformational change in yeast hexokinase. *Proc Natl Acad Sci* **75**: 4848–4852.

- Benz-Moy TL, Herschlag D. 2011. Structure-function analysis from the outside in: long-range tertiary contacts in RNA exhibit distinct catalytic roles. *Biochemistry* **50**: 8733–8755.
- Bevilacqua PC, Turner DH. 1991. Comparison of binding of mixed ribose-deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: evidence for ribozyme specific interactions with 2' OH groups. *Biochemistry* **30**: 10632–10640.
- Bevilacqua PC, Kierzek R, Johnson KA, Turner DH. 1992. Dynamics of ribozyme binding of substrate revealed by fluorescence-detected stopped-flow methods. *Science* **258**: 1355–1358.
- Bevilacqua PC, Johnson KA, Turner DH. 1993. Cooperative and anti-cooperative binding to a ribozyme. *Proc Natl Acad Sci* **90**: 8357–8361.
- Bevilacqua PC, Li Y, Turner DH. 1994. Fluorescence-detected stopped flow with a pyrene labeled substrate reveals that guanosine facilitates docking of the 5' cleavage site into a high free energy binding mode in the *Tetrahymena* ribozyme. *Biochemistry* **33**: 11340–11348.
- Bokinsky G, Zhuang X. 2005. Single-molecule RNA folding. *Acc Chem Res* **38**: 566–573.
- Burke JM, Esherrick JS, Burfeind WR, King JL. 1990. A 3' splice site-binding sequence in the catalytic core of a group I intron. *Nature* **344**: 80–82.
- Bystroff C, Kraut J. 1991. Crystal structure of unliganded Escherichia coli dihydrofolate reductase. Ligand-induced conformational changes and cooperativity in binding. *Biochemistry* **30**: 2227–2239.
- Case DA, Darden TA, Cheatham TE. 2014. AMBER 14 (version 14). University of California, San Francisco.
- Chanfreau G, Jacquier A. 1996. An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J* **15**: 3466–3476.
- Chen SJ, Dill KA. 2000. RNA folding energy landscapes. *Proc Natl Acad Sci* **97**: 646–651.
- Chen W, Moore MJ. 2014. The spliceosome: disorder and dynamics defined. *Curr Opin Struct Biol* **24**: 141–149.
- Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**: 12–21.
- Christian EL. 2005. Identification and characterization of metal ion binding by thiophilic metal ion rescue. In *Handbook of RNA biochemistry*, Vol. 1, pp. 319–344. Wiley-VCH, Weinheim.
- Christian EL, Kaye NM, Harris ME. 2000. Helix P4 is a divalent metal ion binding site in the conserved core of the ribonuclease P ribozyme. *RNA* **6**: 511–519.
- Dai Q, Lea CR, Lu J, Piccirilli JA. 2007. Syntheses of (2')3'-15N-amino-(2')3'-deoxyguanosine and determination of their pKa values by 15N NMR spectroscopy. *Org Lett* **9**: 3057–3060.
- Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* **29**: 7133–7155.
- Duchardt-Ferner E, Weigand JE, Ohlenschläger O, Schmidtke SR, Suess B, Wöhnert J. 2010. Highly modular structure and ligand binding by conformational capture in a minimalistic riboswitch. *Angew Chem Int Ed Engl* **49**: 6216–6219.
- Ferré-D'Amaré AR, Zhou K, Doudna JA. 1998. Crystal structure of a hepatitis delta virus ribozyme. *Nature* **395**: 567–574.
- Fersht A. 1999. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. W.H. Freeman, New York.
- Fica SM, Tuttle N, Novak T, Li NS, Lu J, Koodathingal P, Dai Q, Staley JP, Piccirilli JA. 2013. RNA catalyzes nuclear pre-mRNA splicing. *Nature* **503**: 229–234.
- Flinders J, DeFina SC, Brackett DM, Baugh C, Wilson C, Dieckmann T. 2004. Recognition of planar and nonplanar ligands in the malachite green-RNA aptamer complex. *Chembiochem* **5**: 62–72.
- Forconi M, Lee J, Lee JK, Piccirilli JA, Herschlag D. 2008. Functional identification of ligands for a catalytic metal ion in group I introns. *Biochemistry* **47**: 6883–6894.
- Forconi M, Sengupta RN, Liu MC, Sartorelli AC, Piccirilli JA, Herschlag D. 2009. Structure and function converge to identify a hydrogen bond in a group I ribozyme active site. *Angew Chem Int Ed Engl* **48**: 7171–7175.
- Forconi M, Sengupta RN, Piccirilli JA, Herschlag D. 2010. A rearrangement of the guanosine-binding site establishes an extended network of functional interactions in the *Tetrahymena* group I ribozyme active site. *Biochemistry* **49**: 2753–2762.
- Forconi M, Porecha RH, Piccirilli JA, Herschlag D. 2011. Tightening of active site interactions en route to the transition state revealed by single-atom substitution in the guanosine-binding site of the *Tetrahymena* group I ribozyme. *J Am Chem Soc* **133**: 7791–7800.
- Frank J, Gonzalez RL Jr. 2010. Structure and dynamics of a processive Brownian motor: the translating ribosome. *Annu Rev Biochem* **79**: 381–412.
- Frederiksen JK, Piccirilli JA. 2009. Identification of catalytic metal ion ligands in ribozymes. *Methods* **49**: 148–166.
- Freemont PS, Friedman JM, Beese LS, Sanderson MR, Steitz TA. 1988. Cocrystal structure of an editing complex of Klenow fragment with DNA. *Proc Natl Acad Sci* **85**: 8924–8928.
- Gilbert W. 1986. Origin of life: the RNA world. *Nature* **319**: 618.
- Gilbert SD, Stoddard CD, Wise SJ, Batey RT. 2006. Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. *J Mol Biol* **359**: 754–768.
- Golden BL, Gooding AR, Podell ER, Cech TR. 1998. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science* **282**: 259–264.
- Golden BL, Kim H, Chase E. 2005. Crystal structure of a phage Twort group I ribozyme-product complex. *Nat Struct Mol Biol* **12**: 82–89.
- Gordon PM, Piccirilli JA. 2001. Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat Struct Mol Biol* **8**: 893–898.
- Gordon PM, Sontheimer EJ, Piccirilli JA. 2000. Kinetic characterization of the second step of group II intron splicing: role of metal ions and the cleavage site 2'-OH in catalysis. *Biochemistry* **39**: 12939–12952.
- Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35**: 849–857.
- Guo F, Gooding AR, Cech TR. 2004. Structure of the *Tetrahymena* ribozyme: base triple sandwich and metal ion at the active site. *Mol Cell* **16**: 351–362.
- Guo Z, Karunatilaka KS, Rueda D. 2009. Single-molecule analysis of protein-free U2-U6 snRNAs. *Nat Struct Mol Biol* **16**: 1154–1159.
- Herschlag D. 1992. Evidence for processivity and two-step binding of the RNA substrate from studies of J1/2 mutants of the *Tetrahymena* ribozyme. *Biochemistry* **31**: 1386–1399.
- Herschlag D. 1995. RNA chaperones and the RNA folding problem. *J Biol Chem* **270**: 20871–20874.
- Herschlag D. 1998. Ribozyme crevices and catalysis. *Nature* **395**: 548–549.
- Herschlag D, Cech TR. 1990. Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry* **29**: 10159–10171.
- Herschlag D, Khosla M. 1994. Comparison of pH dependencies of the *Tetrahymena* ribozyme reactions with RNA 2'-substituted and phosphorothioate substrates reveals a rate-limiting conformational step. *Biochemistry* **33**: 5291–5297.
- Herschlag D, Eckstein F, Cech TR. 1993. Contributions of 2'-hydroxyl groups of the RNA substrate to binding and catalysis by the *Tetrahymena* ribozyme. An energetic picture of an active-site composed of RNA. *Biochemistry* **32**: 8299–8311.
- Houglund JL, Kravchuk AV, Herschlag D, Piccirilli JA. 2005. Functional identification of catalytic metal ion binding sites within RNA. *PLoS Biol* **3**: e277.
- Houglund JL, Piccirilli JA, Forconi M, Lee J, Herschlag D. 2006. How the group I intron works: a case study of RNA structure and function. In *RNA World*, 3rd ed., vol. 43, pp. 133–205. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Houglund JL, Sengupta RN, Dai Q, Deb SK, Piccirilli JA. 2008. The 2'-hydroxyl group of the guanosine nucleophile donates a

- functionally important hydrogen bond in the *Tetrahymena* ribozyme reaction. *Biochemistry* **47**: 7684–7694.
- Hsieh J, Fierke CA. 2009. Conformational change in the *Bacillus subtilis* RNase P holoenzyme–pre-tRNA complex enhances substrate affinity and limits cleavage rate. *RNA* **15**: 1565–1577.
- Huang H, Suslov NB, Li NS, Shelke SA, Evans ME, Koldobskaya Y, Rice PA, Piccirilli JA. 2014. A G-quadruplex-containing RNA activates fluorescence in a GFP-like fluorophore. *Nat Chem Biol* **10**: 686–691.
- Jucker FM, Phillips RM, McCallum SA, Pardi A. 2003. Role of a heterogeneous free state in the formation of a specific RNA-theophylline complex. *Biochemistry* **42**: 2560–2567.
- Karbstein K, Herschlag D. 2003. Extraordinarily slow binding of guanosine to the *Tetrahymena* group I ribozyme: implications for RNA preorganization and function. *Proc Natl Acad Sci* **100**: 2300–2305.
- Karbstein K, Carroll KS, Herschlag D. 2002. Probing the *Tetrahymena* group I ribozyme reaction in both directions. *Biochemistry* **41**: 11171–11183.
- Kim EE, Wyckoff HW. 1991. Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. *J Mol Biol* **218**: 449–464.
- Knitt DS, Herschlag D. 1996. pH dependencies of the *Tetrahymena* ribozyme reveal an unconventional origin of an apparent pKa. *Biochemistry* **35**: 1560–1570.
- Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR. 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* **31**: 147–157.
- Leontis NB, Zirbel CL. 2012. Nonredundant 3D structure datasets for RNA knowledge extraction and benchmarking. In *Nucleic acids and molecular biology*, vol. 27 (ed. Leontis N, Westhof E), pp. 281–298. Springer, Berlin, Heidelberg.
- Lesk AM, Chothia C. 1984. Mechanisms of domain closure in proteins. *J Mol Biol* **174**: 175–191.
- Li Y, Bevilacqua PC, Mathews D, Turner DH. 1995. Thermodynamic and activation parameters for binding of a pyrene-labeled substrate by the *Tetrahymena* ribozyme: docking is not diffusion-controlled and is driven by a favorable entropy change. *Biochemistry* **34**: 14394–14399.
- Lipchock SV, Strobel SA. 2008. A relaxed active site after exon ligation by the group I intron. *Proc Natl Acad Sci* **105**: 5699–5704.
- Lolis E, Petsko GA. 1990. Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-Å resolution: implications for catalysis. *Biochemistry* **29**: 6619–6625.
- Lu XJ, Olson WK. 2003. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* **31**: 5108–5121.
- Marcia M, Pyle AM. 2012. Visualizing group II intron catalysis through the stages of splicing. *Cell* **151**: 497–507.
- Martick M, Scott WG. 2006. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* **126**: 309–320.
- McConnell TS, Cech TR, Herschlag D. 1993. Guanosine binding to the *Tetrahymena* ribozyme: thermodynamic coupling with oligonucleotide binding. *Proc Natl Acad Sci* **90**: 8362–8366.
- Meyer M, Nielsen H, Oliéric V, Roblin P, Johansen SD, Westhof E, Masquida B. 2014. Speciation of a group I intron into a lariat capping ribozyme. *Proc Natl Acad Sci* **111**: 7659–7664.
- Michel F, Westhof E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol* **216**: 585–610.
- Michel F, Hanna M, Green R, Bartel DP, Szostak JW. 1989. The guanosine binding site of the *Tetrahymena* ribozyme. *Nature* **342**: 391–395.
- Montange RK, Batey RT. 2008. Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys* **37**: 117–133.
- Moran S, Kierzek R, Turner DH. 1993. Binding of guanosine and 3' splice site analogues to a group I ribozyme: interactions with functional groups of guanosine and with additional nucleotides. *Biochemistry* **32**: 5247–5256.
- Munro JB, Sanbonmatsu KY, Spahn CM, Blanchard SC. 2009. Navigating the ribosome's metastable energy landscape. *Trends Biochem Sci* **34**: 390–400.
- Narlikar GJ, Herschlag D. 1996. Isolation of a local tertiary folding transition in the context of a globally folded RNA. *Nat Struct Biol* **3**: 701–710.
- Narlikar GJ, Herschlag D. 1997. Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. *Annu Rev Biochem* **66**: 19–59.
- Narlikar GJ, Herschlag D. 1998. Direct demonstration of the catalytic role of binding interactions in an enzymatic reaction. *Biochemistry* **37**: 9902–9911.
- Narlikar GJ, Gopalakrishnan V, McConnell TS, Usman N, Herschlag D. 1995. Use of binding energy by an RNA enzyme for catalysis by positioning and substrate destabilization. *Proc Natl Acad Sci* **92**: 3668–3672.
- Narlikar GJ, Khosla M, Usman N, Herschlag D. 1997. Quantitating tertiary binding energies of 2' OH groups on the P1 duplex of the *Tetrahymena* ribozyme: intrinsic binding energy in an RNA enzyme. *Biochemistry* **36**: 2465–2477.
- Narlikar GJ, Bartley LE, Khosla M, Herschlag D. 1999. Characterization of a local folding event of the *Tetrahymena* group I ribozyme: effects of oligonucleotide substrate length, pH, and temperature on the two substrate binding steps. *Biochemistry* **38**: 14192–14204.
- Nelson JA, Uhlenbeck OC. 2008. Hammerhead redux: does the new structure fit the old biochemical data? *RNA* **14**: 605–615.
- Noeske J, Buck J, Fürtig B, Nasiri HR, Schwalbe H, Wöhnert J. 2007. Interplay of 'induced fit' and preorganization in the ligand induced folding of the aptamer domain of the guanine binding riboswitch. *Nucleic Acids Res* **35**: 572–583.
- Ortoleva-Donnelly L, Szwczak AA, Gutell RR, Strobel SA. 1998. The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme. *RNA* **4**: 498–519.
- Ottink OM, Rampersad SM, Tessari M, Zaman GJ, Heus HA, Wijmenga SS. 2007. Ligand-induced folding of the guanine-sensing riboswitch is controlled by a combined predetermined induced fit mechanism. *RNA* **13**: 2202–2212.
- Pérez A, Marchán I, Svozil D, Sponer J, Cheatham TE III, Laughton CA, Orozco M. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of  $\alpha/\gamma$  conformers. *Biophys J* **92**: 3817–3829.
- Piccirilli JA, Vyle JS, Caruthers MH, Cech TR. 1993. Metal ion catalysis in the *Tetrahymena* ribozyme reaction. *Nature* **361**: 85–88.
- Profenno LA, Kierzek R, Testa SM, Turner DH. 1997. Guanosine binds to the *Tetrahymena* ribozyme in more than one step, and its 2'-OH and the nonbridging pro-SP phosphoryl oxygen at the cleavage site are required for productive docking. *Biochemistry* **36**: 12477–12485.
- Pyle AM, Cech TR. 1991. Ribozyme recognition of RNA by tertiary interactions with specific ribose 2'-OH groups. *Nature* **350**: 628–631.
- Pyle AM, Murphy FL, Cech TR. 1992. RNA substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. *Nature* **358**: 123–128.
- Robertson MP, Joyce GF. 2012. The origins of the RNA world. *Cold Spring Harb Perspect Biol* **4**: pii: a003608.
- Russell R. 2008. RNA misfolding and the action of chaperones. *Front Biosci* **13**: 1–20.
- Russell R, Herschlag D. 1999. Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme. *RNA* **5**: 158–166.
- Saenger W. 1983. *Principles of nucleic acid structure*. Springer-Verlag, Berlin.
- Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H. 1997. Crystal structures of human DNA polymerase  $\beta$  complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**: 11205–11215.
- Schmeing TM, Huang KS, Strobel SA, Steitz TA. 2005. An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature* **438**: 520–524.
- Schnabl J, Suter P, Sigel RKO. 2012. MINAS—a database of Metal Ions in Nucleic Acids. *Nucleic Acids Res* **40**: D434–D438.

- Schulenburg C, Hilvert D. 2013. Protein conformational disorder and enzyme catalysis. *Top Curr Chem* **337**: 41–67.
- Sengupta RN, Herschlag D, Piccirilli JA. 2012. Thermodynamic evidence for negative charge stabilization by a catalytic metal ion within an RNA active site. *ACS Chem Biol* **7**: 294–299.
- Shan SO, Herschlag D. 1999. Probing the role of metal ions in RNA catalysis: kinetic and thermodynamic characterization of a metal ion interaction with the 2'-moiety of the guanosine nucleophile in the *Tetrahymena* group I ribozyme. *Biochemistry* **38**: 10958–10975.
- Shan SO, Herschlag D. 2000. An unconventional origin of metal-ion rescue and inhibition in the *Tetrahymena* group I ribozyme reaction. *RNA* **6**: 795–813.
- Shan SO, Herschlag D. 2002. Dissection of a metal-ion-mediated conformational change in *Tetrahymena* ribozyme catalysis. *RNA* **8**: 861–872.
- Shan S, Yoshida A, Sun S, Piccirilli JA, Herschlag D. 1999a. Three metal ions at the active site of the *Tetrahymena* group I ribozyme. *Proc Natl Acad Sci* **96**: 12299–12304.
- Shan SO, Narlikar GJ, Herschlag D. 1999b. Protonated 2'-aminoguanosine as a probe of the electrostatic environment of the active site of the *Tetrahymena* group I ribozyme. *Biochemistry* **38**: 10976–10988.
- Shan S, Kravchuk AV, Piccirilli JA, Herschlag D. 2001. Defining the catalytic metal ion interactions in the *Tetrahymena* ribozyme reaction. *Biochemistry* **40**: 5161–5171.
- Shcherbakova I, Mitra S, Laederach A, Brenowitz M. 2008. Energy barriers, pathways, and dynamics during folding of large, multidomain RNAs. *Curr Opin Chem Biol* **12**: 655–666.
- Sigler PB. 1975. An analysis of the structure of tRNA. *Annu Rev Biophys Bio* **4**: 477–527.
- Sjögren AS, Pettersson E, Sjöberg BM, Strömberg R. 1997. Metal ion interaction with cosubstrate in self-splicing of group I introns. *Nucl Acids Res* **25**: 648–653.
- Sripathi KN, Tay WW, Banáš P, Otyepka M, Šponer J, Walter NG. 2014. Disparate HDV ribozyme crystal structures represent intermediates on a rugged free-energy landscape. *RNA* **20**: 1112–1128.
- Stahley MR, Strobel SA. 2005. Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science* **309**: 1587–1590.
- Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**: 315–326.
- Steitz TA. 1998. A mechanism for all polymerases. *Nature* **391**: 231–232.
- Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci* **90**: 6498–6502.
- Strater N, Lipscomb WN, Klabunde T, Krebs B. 1996. Two-metal ion catalysis in enzymatic acyl- and phosphoryl-transfer reactions. *Angew Chem Int Ed Engl* **35**: 2024–2055.
- Toor N, Rajashankar K, Keating KS, Pyle AM. 2008. Structural basis for exon recognition by a group II intron. *Nat Struct Mol Biol* **15**: 1221–1222.
- Treiber DK, Williamson JR. 2001. Beyond kinetic traps in RNA folding. *Curr Opin Struct Biol* **11**: 309–314.
- Tsui V, Case DA. 2000. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56**: 275–291.
- Vicens Q, Cech TR. 2006. Atomic level architecture of group I introns revealed. *Trends Biochem Sci* **31**: 41–51.
- Voorhees RM, Ramakrishnan V. 2013. Structural basis of the translational elongation cycle. *Annu Rev Biochem* **82**: 203–236.
- Wang JF, Downs WD, Cech TR. 1993. Movement of the guide sequence during RNA catalysis by a group I ribozyme. *Science* **260**: 504–508.
- Wang SL, Karbstein K, Peracchi A, Beigelman L, Herschlag D. 1999. Identification of the hammerhead ribozyme metal ion binding site responsible for rescue of the deleterious effect of a cleavage site phosphorothioate. *Biochemistry* **38**: 14363–14378.
- Wang JM, Cieplak P, Kollman PA. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* **21**: 1049–1074.
- Ward WL, Derose VJ. 2012. Ground-state coordination of a catalytic metal to the scissile phosphate of a tertiary-stabilized Hammerhead ribozyme. *RNA* **18**: 16–23.
- Weinger JS, Parnell KM, Dorner S, Green R, Strobel SA. 2004. Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat Struct Mol Biol* **11**: 1101–1106.
- Weinstein LB, Jones BC, Cosstick R, Cech TR. 1997. A second catalytic metal ion in group I ribozyme. *Nature* **388**: 805–808.
- Wickiser JK, Cheah MT, Breaker RR, Crothers DM. 2005. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **44**: 13404–13414.
- Yang W, Lee JY, Nowotny M. 2006. Making and breaking nucleic acids: two-Mg<sup>2+</sup>-ion catalysis and substrate specificity. *Mol Cell* **22**: 5–13.
- Yoshida A, Sun S, Piccirilli JA. 1999. A new metal ion interaction in the *Tetrahymena* ribozyme reaction revealed by double sulfur substitution. *Nat Struct Biol* **6**: 318–321.
- Zaug AJ, Grosshans CA, Cech TR. 1988. Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry* **27**: 8924–8931.
- Zgarbová M, Otyepka M, Šponer J, Mládek A, Banáš P, Cheatham TE III, Jurečka P. 2011. Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J Chem Theory Comput* **7**: 2886–2902.
- Zhuang X, Kim H, Pereira MJ, Babcock HP, Walter NG, Chu S. 2002a. Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**: 1473–1476.
- Zhuang XW, Kim H, Pereira MJB, Babcock HP, Walter NG, Chu S. 2002b. Correlating structural dynamics and function in single ribozyme molecules. *Science* **296**: 1473–1476.