Chemistry & Biology
Article

# Structural Investigation of the GlmS Ribozyme Bound to Its Catalytic Cofactor

Jesse C. Cochrane,[1] Sarah V. Lipchock,[2] and Scott A. Strobel[1,2,*]
[1] Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06520, USA
[2] Department of Chemistry, Yale University, 260 Whitney Avenue, New Haven, CT 06520, USA
*Correspondence: scott.strobel@yale.edu
DOI 10.1016/j.chembiol.2006.12.005

## SUMMARY

**The GlmS riboswitch is located in the 5′-untranslated region of the gene encoding glucosamine-6-phosphate (GlcN6P) synthetase. The GlmS riboswitch is a ribozyme with activity triggered by binding of the metabolite GlcN6P. Presented here is the structure of the GlmS ribozyme (2.5 Å resolution) with GlcN6P bound in the active site. The GlmS ribozyme adopts a compact double pseudoknot tertiary structure, with two closely packed helical stacks. Recognition of GlcN6P is achieved through coordination of the phosphate moiety by two hydrated magnesium ions as well as specific nucleobase contacts to the GlcN6P sugar ring. Comparison of this activator bound and the previously published apoenzyme complex supports a model in which GlcN6P does not induce a conformational change in the RNA, as is typical of other riboswitches, but instead functions as a catalytic cofactor for the reaction. This demonstrates that RNA, like protein enzymes, can employ the chemical diversity of small molecules to promote catalytic activity.**

## INTRODUCTION

Riboswitches are *cis*-acting structural RNA molecules that directly control gene expression [1]. These biosensors are sensitive to the concentration of a specific small molecule and control expression through a feedback loop mechanism. Ligand binding causes a structural rearrangement in the RNA that generally leads to the downregulation of the gene through transcriptional termination or inefficient translational initiation. In many cases, genes controlled by the riboswitch are part of the biosynthetic pathway of the small molecule recognized by the RNA [2]. Riboswitches are commonly found in Gram-positive bacteria and control over 2% of genes in *Bacillus subtillis* in response to a variety of cofactors, including guanine, adenine, co-enzyme B12, lysine, and even glycine [1].

The GlmS riboswitch is located in the 5′-untranslated region (5′-UTR) of the gene for glucosamine-6-phosphate synthetase [3], which converts glutamine and fructose-6-phosphate to glucosamine-6-phosphate (GlcN6P) [4]. Like many other riboswitches that control downstream gene expression, the GlmS riboswitch negatively regulates GlcN6P synthetase production [3]. High GlcN6P concentrations activate the riboswitch, resulting in reduced synthetase expression. However, unlike every other riboswitch identified to date, the GlmS riboswitch does not appear to undergo any structural rearrangement upon effector binding, suggesting that the binding pocket is preformed in the absence of ligand [3, 5]. Instead, GlcN6P binding results in the specific cleavage of the GlmS mRNA at a single site 5′ of the riboswitch sequence. The GlmS riboswitch is a ribozyme and, like other autolytic ribozymes, the cleavage reaction involves nucleophilic attack of a vicinal 2′-OH on the scissile phosphate to produce products with a 2′-3′ cyclic phosphate and a 5′-OH [3, 6].

Although no large-scale conformational changes were detected biochemically upon GlcN6P binding, the effector is absolutely required for GlmS ribozyme activation [3]. GlcN6P addition increases the cleavage rate 100,000-fold over background hydrolysis, with the amino group being the principal determinant of activity. Glucose-6-phosphate (Glc6P), which contains a hydroxyl in place of the GlcN6P amine, is a competitive inhibitor of the reaction [7]. Glucosamine (GlcN), which lacks the phosphate but retains the primary amine, is able to promote the cleavage reaction, albeit 30-fold less efficiently than GlcN6P [3, 7]. Activity has even been reported for other small molecules such as Tris and serinol, although they are poor activators, presumably due to weak RNA binding [7]. However, all activators share the common feature of an ethanolamine moiety, a primary amine, and a vicinal hydroxyl group. Additional evidence for the catalytic importance of the amine in the phosphoryltransfer reaction came from biochemical experiments that demonstrated a pH dependence of the reaction as a function of the $pK_a$ of the effector [7]. For GlcN6P, GlcN, and serinol, the reaction rate went up with increasing pH, revealing an apparent $pK_a$ that approximately matched the $pK_a$ of the amine for each ligand. This suggested that the amino rather than the ammonium form of the metabolite is active in the cleavage reaction [7].

To understand the structural basis of effector mediated riboswitch activation, we determined the crystal structure of the *Bacillus anthracis* GlmS ribozyme bound to the effector GlcN6P. During the final stages of our structural refinement, Klein and Ferré-D'Amaré reported a series of structures of the *Thermoanerobacter tengcongenis* GlmS [8]. These included the 2′-deoxy and 2′-amino

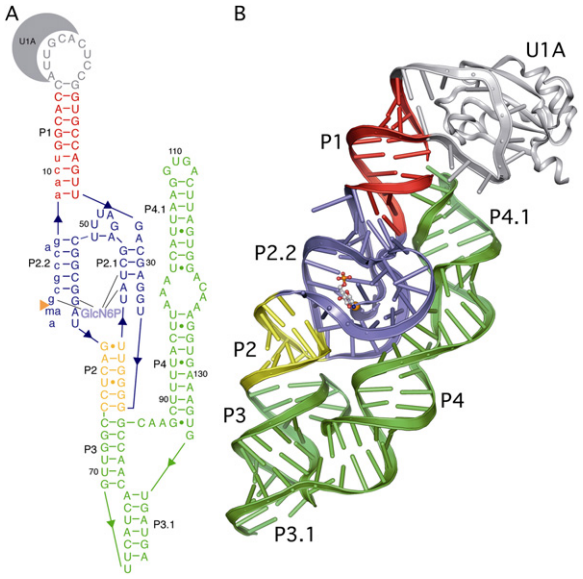



**Figure 1. Secondary and Tertiary Structures of the GlmS Ribozyme**

(A) Secondary structure of the *B. anthracis* GlmS ribozyme, with the transcript shown in uppercase letters and the synthetic substrate shown in lowercase letters.

(B) Tertiary structure depicted in ribbons format. The P1 helix is shown in red, P2 in yellow, and P3–P4 in green. The double pseudoknot, consisting of helices P2.1, P2.2, P2a, and related joiner regions is in blue. GlcN6P is bound in the active site and is shown in ball and stick representation primarily in gray. U1A and the U1A binding loop are in gray.

substrate form of the GlmS ribozyme with and without the competitive inhibitor, Glc6P, and the cleaved ribozyme product. However, they were unable to crystallize the GlcN6P bound complex [8]. Comparison of these reaction states with the GlcN6P bound complex reported here addresses how the RNA recognizes its ligand and provides insight into how GlcN6P, especially the amine and adjacent hydroxyl group, promotes catalysis.

## RESULTS AND DISCUSSION

### Crystallization and Activity of a Two-Part *Bacillus anthracis* GlmS Ribozyme

The crystallization construct was designed based on the GlmS ribozyme located in the 5′-UTR of the *Bacillus anthracis* glmS gene [9] (Figure 1A). The crystallization construct contained all of the nucleotides previously identified as critical for ribozyme activity [3]. Two modifications were made to the ribozyme to promote crystallization. First, the recognition sequence for the RNA binding domain of the U1A protein was appended to the top of helix P1, a loop of variable size in GlmS ribozymes [9], to promote cocrystallization with the U1A protein [10]. Second, the first eleven nucleotides of the GlmS ribozyme were removed from the transcribed RNA sequence and added as a substrate strand in *trans*. The synthetic oligonucleotide contained two nucleotides 5′ of the cleavage site and the first eleven nucleotides of the GlmS ribozyme. This made it possible to inhibit ribozyme activity by introducing a 2′-O-methyl substitution at the cleavage site. Diffraction quality crystals were obtained of a complex containing equimolar ratios of the GlmS ribozyme RNA, the synthetic oligonucleotide, and the U1A protein.

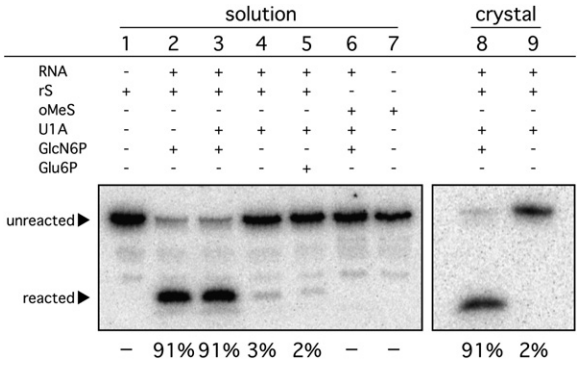



**Figure 2. Reactivity of the *B. anthracis* GlmS Ribozyme Construct under Crystallization Conditions**

Activity assays conditions were similar to those used for the refolding of the GlmS RNA with a radiolabeled synthetic oligonucleotide substrate. The substrate rS has a ribose at position A-1; oMeS has a 2′-O-methyl at A-1. Reactions were performed either in solution or in crystals as indicated. The reagents included in each reaction are as indicated above each lane. The mobility of the substrate and product are marked with arrows. The percent cleavage for each reaction after a 2 hr incubation is listed below the autoradiogram.

The GlcN6P-dependent reactivity of this GlmS ribozyme construct was confirmed in solution and within the crystals. In solution, the GlmS RNA reacted with an all-ribose substrate at the rate of 6 $min^{-1}$, and the rate was unaffected by U1A addition (Figure 2, lanes 2 and 3). Substantially less activity was observed in the absence of GlcN6P (0.0006 $min^{-1}$) or upon addition of Glc6P instead of GlcN6P (0.0006 $min^{-1}$) (Figure 2, lanes 4 and 5). The 2′-O-methyl substitution at the cleavage site resulted in complete inactivation of ribozyme reaction even in the presence of GlcN6P (Figure 2, lane 6).

To determine the activity of the crystallized RNA, the ribozyme was crystallized with an all ribose substrate radiolabeled at its 5′-end in the absence of GlcN6P. The radioactive crystals were harvested, washed, and soaked in buffer either with or without 2 mM GlcN6P. RNA in crystals incubated with GlcN6P reacted to completion within 2 hr, while those soaked in buffer alone showed no activity (Figure 2, lanes 8 and 9). No noticeable change in crystal morphology was evident upon the addition of GlcN6P. These results suggest that the RNA is in a catalytically relevant conformation in the crystalline state.

### Overview of the *B. anthracis* GlmS Ribozyme

The structure of the GlmS ribozyme in a state prior to cleavage (Figure 1B) was determined by multiple isomorphous replacement to 2.5 Å resolution (Supplemental Table S1). There are four molecules in the asymmetric unit with intermolecular contacts dominated by the U1A protein, which forms a dimer of dimers. The tertiary structure

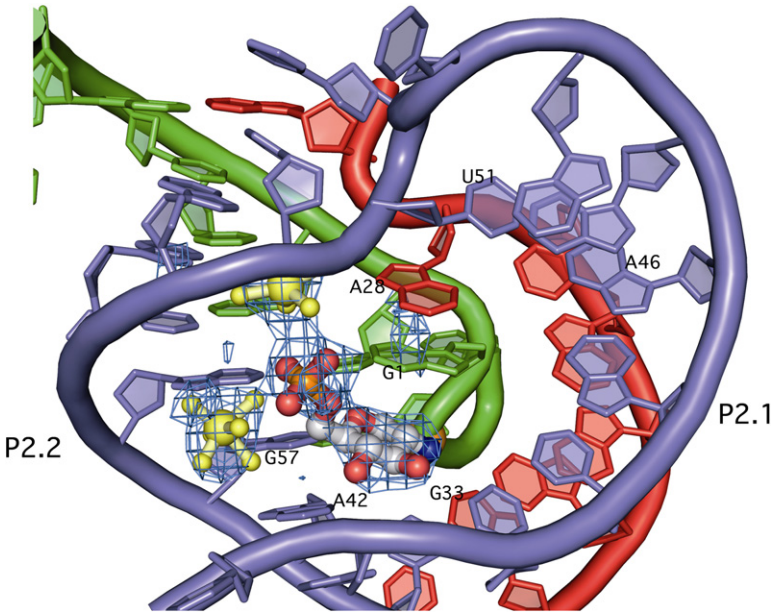



**Figure 3. Active Site of the GlmS Ribozyme**

The double pseudoknotted active site of the GlmS ribozyme in the same orientation as in Figure 1b. The color scheme has been altered to emphasize the contribution of each strand in the active site. The substrate strand (nucleotides -2 to 9) are shown in green. Nucleotides 27-34 are in red, and nucleotides 41-59, which comprise the P2 loop, are in blue. This color scheme is also used in Figures 4, 5 and 7. The P2.1 pseudoknot is formed from the red and blue strands; the P2.2 pseudoknot is formed from the green and blue strands. GlcN6P is shown primarily in gray and the two hydrated magnesium ions are in yellow. The scissile phosphate, barely visible behind GlcN6P, is shown as an orange sphere. $F_o$-$F_c$ density is shown in blue mesh, contoured at 2.5σ, calculated before addition of GlcN6P, metal ions or water molecules to the model.

of the GlmS ribozyme is highly compact: P1, P2, P3, and P3.1 form one helical stack that abuts a second helical stack formed from P4 and P4.1 (Figure 1B). In all four molecules, there is unambiguous density for the entire phosphate backbone, and a molecular model for all nucleotides was built and refined to an $R_{free}$ of 26.9%. Clear density was observed in the active site for glucosamine-6-phosphate (Figure 3). This GlmS ribozyme structure includes the required cofactor for the reaction.

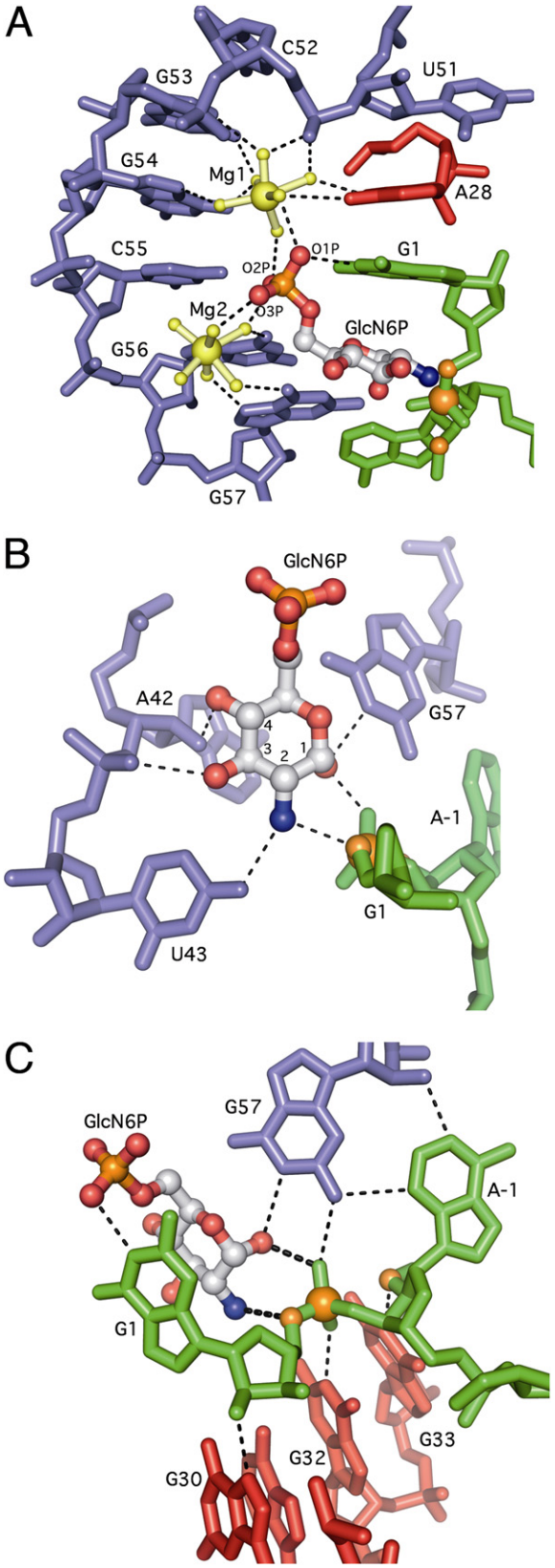
The structures of the *B. anthracis* and the *T tengcongenis* GlmS ribozymes reveal a complex secondary structural fold that differs significantly from that predicted for the GlmS ribozyme in the core region of the molecule [3, 8]. The critical difference is a double pseudoknot formed from nucleotides in the P2 loop (Figure 3). Although the topology differs, double pseudoknots have also been observed in the structures of the HDV and Diels-Alderase ribozymes [11, 12]. Like the HDV ribozyme, the double pseudoknot forms an active site cleft, with the scissile phosphate located at the bottom [11]. One pseudoknot, P2.1, is formed from the joiner region between P1 and P2 and the 5′ end of the P2 loop. The second pseudoknot, P2.2, is formed from a base-pairing interaction between the 3′ side of the P2 loop and the 5′ end of the ribozyme. These pseudoknots form the ribozyme core, including the GlcN6P binding pocket. The bottom of the active site is organized through the nonsequential stacking of unpaired nucleotides. The P2.1 and P2.2 pseudoknots form the sides of an active site that is capped by a single base pair between A46 and U51 that closes the P2 loop. This single base pair establishes the relative orientations of the P2.1 and P2.2 helices. This constrained geometry forces the unpaired A28 in the joiner region between P1 and P2.1 to stack directly on top of G1, the nucleotide 3′ of the scissile phosphate. G1, in turn, stacks on top of the GlcN6P.

The compact core of the ribozyme is buttressed by the functionally dispensable P3 and P4 helices [3]. About 1200 Å$^2$ of the ribozyme core surface area is buried by P4. Long-range tertiary interactions serve to further solidify this compact structure. The GNRA tetraloop at the top of P4.1 docks into the minor groove of P1, connecting these two disparate regions of the RNA structure. The cross-strand stack of As in the bulge between P4/P4.1 packs obliquely into the minor groove of P2.1. The predicted pseudoknot, P3.1, between the 3′ end of the ribozyme and the P3 loop completes the compact overall fold of this ribozyme [13].

## The GlcN6P Binding Pocket

Binding of GlcN6P by the GlmS ribozyme is achieved through recognition of both the phosphate and sugar moieties (Figure 4). Two fully hydrated $Mg^{2+}$ ions contact the phosphate group, while the glucosamine ring makes direct interactions with the RNA. This may be a common recognition mechanism for riboswitches that bind ligands containing phosphate. For example, the thiamine pyrophosphate riboswitch also binds the pyrophosphate using two solvated $Mg^{2+}$ ions while using nucleotide functional groups for direct interaction with the thiamine [14].

The phosphate binding edge of the GlcN6P pocket is solvent exposed. Magnesium ions and water molecules appear to screen the negative charge of the phosphate backbone from the phosphate group of GlcN6P. In the GlmS structure, each of the three GlcN6P nonbridging phosphate oxygens are coordinated to at least one of the two hydrated $Mg^{2+}$ ions (Mg1 and Mg2) (Figure 4A). The $Mg^{2+}$ ions are ~7.5 Å apart, with the phosphate of GlcN6P positioned equidistant from both. One phosphate oxygen, here designated O2P, hydrogen bonds to waters from both metals, while a second water from Mg2 interacts with the O3P oxygen. The O1P oxygen interacts




with a water from Mg1 and makes a direct contact to the N1 of G1.

Both metals are positioned in the major groove of the P2.2 helix in the RNA's effector binding pocket. The vast majority of the contacts are to nucleobase functional groups. An outer sphere interaction between Mg1 and the pro-$S_P$ nonbridging phosphate oxygen of C52 is the only contact to an RNA phosphate group. Mg1 makes additional outer sphere contacts with the N7 and O/N6 of A28, G53, and G54. Interference seen with 7-deazaadenosince at A28 demonstrates the importance of this functional group for activity [15]. Mg2 makes water-mediated contacts to the N7 and O6 of G56 and G57. In contrast to this structure, only one metal, Mg1, was observed in the *T. tengcongenis* structure bound to the inhibitor Glc6P [8]. In that case, Mg1 appears to make direct contact with one phosphate oxygen of Glc6P. This difference may result from alternate binding modes for Glc6P compared to GlcN6P or other aspects of the crystallization conditions.

The effector binding pocket in the RNA makes an extensive set of hydrogen bonding interactions with all four of the exocyclic functional groups on the glucosamine ring. Recognition of the ethanolamine moiety of GlcN6P (the C1-OH and adjacent C2-amine), the minimal substrate for the ribozyme reaction [7], explains the specificity of the GlmS ribozyme. The C1-OH accepts a hydrogen bond from the N1 of G57 and donates a hydrogen bond to the pro-$S_P$ nonbridging oxygen of the scissile phosphate (Figure 4B). The C2-amine donates hydrogen bonds to the G1 5′-O leaving group and to the O4 of U43 (Figure 4B). The contacts to the pro-$S_P$ oxygen and the 5′-O are the only interactions between GlcN6P and atoms involved in bond making and bond breaking. Additional binding specificity for the effector is provided by contacts to the C3 and C4 hydroxyl groups. The C3-OH makes a hydrogen bond with the pro-$R_P$ oxygen of U43, and the C4-OH makes a hydrogen bond with the 2′-OH of A42 (Figure 4B).

Subtle, but potentially important, differences were observed in the hydrogen bonding patterns to the sugars of Glc6P and GlcN6P [8]. In the Glc6P inhibitor structure, the C2-OH also hydrogen bonds to the 5′-O leaving group, but because of the C2 hydroxyl substitution, it can donate

**Figure 4. GlcN6P Binding by the GlmS Ribozyome**

(A) Phosphate coordination by two magnesium ions in the GlmS ribozyme. Nucleotides in P2.1 (blue), A28 (red), and G1 (green) use water-mediated contacts to organize two hydrated magnesium ions (yellow) and orient the GlcN6P (primarily in gray) phosphate oxygens in the active site.

(B) Recognition of the GlcN6P sugar ring by nucleobase functional groups. The sugar contacts nucleotides A42, U43, and G57 (blue), and the sugar and 3′-phosphate of G1 (green). The guanine base at G1, which stacks on top of GlcN6P (see Figure 4A), is not shown to allow all hydrogen bonding contacts to be visualized.

(C) Active site interactions expected to stabilize this conformation are depicted as thin dashed lines. The catalytically critical interactions between the ethanolamine moiety of GlcN6P and the reactive phosphate are shown as thicker dashed lines. The scissile phosphate, 5′-O leaving group (which has been methylated), and 2′-O nucleophile are shown as orange spheres.

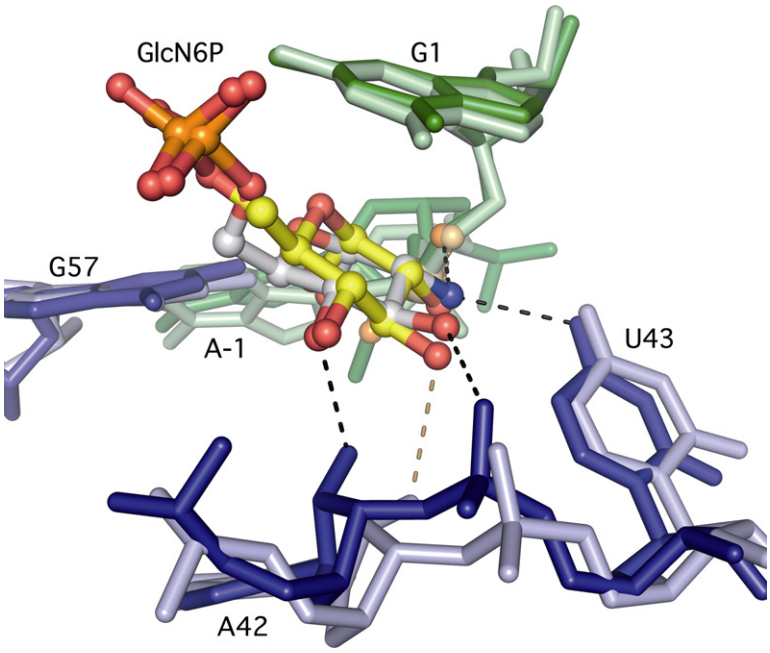



**Figure 5. Comparison of the active sites of GlcN6P bound *B. anthracis* GlmS ribozyme and Glc6P bound *T tengcongenis* GlmS ribozyme**

*B. anthracis* nucleotides are in darker shades of blue and green and *T tengcongenis* nucleotides are in lighter shades. GlcN6P is shown primarily in gray and Glc6P is shown primarily in yellow. Hydrogen bonds to GlcN6P are show in black, hydrogen bonds to Glc6P are shown in dark yellow.

only one hydrogen bond and does not interact with the O4 of U43 (U51 in *T. tengcongenis*). There is an approximately 1 Å shift in nucleotides A42 (A50) and U43 (U51) between the GlcN6P bound and Glc6P bound forms of the ribozyme (Figure 5). This shift leads to changes in the hydrogen bonding network around the C3 and C4 hydroxyls of Glc6P [8]. The C3-OH makes a hydrogen bond to the 2′-OH of A42 (A50), while no hydrogen bonding is observed to the C4-OH. Biochemical studies found that a GlcN6P mimic with inverted stereochemistry at C4-OH does not activate the *B. cereus* GlmS ribozyme, which suggests that this functional group contacts the ribozyme[16]. The presence of such a contact to the C4-OH in the *B. anthracis* GlcN6P structure, and its absence in the Glc6P *T. tencongenis* structure, may reflect a subtle but functionally important difference between activator and inhibitor binding.

## Catalysis by the GlmS Ribozyme

The structure suggests that the GlmS ribozyme is not a metalloenzyme. The active site is devoid of metal ions that could directly participate in the cleavage reaction, which is consistent with biochemical observations [9] High concentrations of monovalent metal ions or cobalt hexamine, which mimics a hydrated magnesium, are able to fully support activity [9]. The catalytic potential of the GlmS ribozyme appears to be achieved by nucleobases and GlcN6P acting as a catalytic cofactor.

Based on analogy to other phosphorotransferases such as ribonuclease A, there are at least four strategies the GlmS ribozyme might use to promote its autolytic reaction[6]: (1) alignment of the nucleophile and leaving group, (2) deprotonation of the 2′-OH nucleophile, (3) stabilization of the negative charge on the scissile phosphate, and/or (4) protonation of the 5′-O leaving group. GlcN6P could promote the ribozyme activity of this riboswitch by fulfilling any one or a combination of these roles, through inducing a localized conformational change to create an inline conformation or by directly participating in the chemical reaction.

The structure of the GlmS ribozyme reveals an active site in which the scissile phosphate and flanking nucleotides are oriented for chemistry (Figure 4c). The nucleotides immediately 5′ and 3′ of the reactive center are unstacked and splayed away from each other. The nucleotide 5′ of the cleavage site, A-1, is orientated by an interaction with G57. Consistent with an important role in substrate orientation, mutation of G57 severely impairs ribozyme function, leading to minimal (G57C) or no (G57A) activity [17]. The nucleotide 3′ of the cleavage site, G1, stacks underneath A28 and on top of GlcN6P (Figure 4A). It is positioned through the donation of two hydrogen bonds: one from the N1 of G1 to a phosphate oxygen of GlcN6P, and the second from the G1 2′-OH to the N7 of G30. Substitution of G30 with 7-deazaguanosine is detrimental to GlmS ribozyme activity in interference analysis [15]. These interactions confer a 165° torsion angle between the 2′-OH nucleophile and the 5′-O leaving group in the ground state, very close to the 180° angle predicted in the transition state for inline attack on the scissile phosphate.

The active sites of the uncleaved *T. tencongenis* and the GlcN6P bound *B. anthracis* GlmS ribozyme structures reveal only small changes in geometry upon effector binding [8]. The core region of the GlcN6P bound ribozyme superimposes with only a 1.5Å deviation on the unliganded form of the *T. tengcongenis* GlmS ribozyme. Thus, while the GlmS ribozyme aligns the reactive 2′-OH and 5′-O groups for nucleophilic attack, it does so in a manner largely independent of GlcN6P binding. This is contrary to a model in

which the GlmS riboswitch functions through a global or local conformational change upon GlcN6P binding. Like the in vitro selected Diels-Alderase ribozyme [12], which adopts the same active site conformation independent of substrates or products, the GlmS ribozyme active site appears to be largely preorganized in the absence of the cofactor [5].

If effector binding does not induce a conformational change, how does the GlcN6P trigger the cleavage activity of this riboswitch? The structure suggests that GlcN6P directly participates in the reaction as a catalytic cofactor using the ethanolamine moiety on the glucosamine sugar. The C1-OH and C2-$NH_2$ groups are in direct hydrogen bonding contact with the scissile phosphate and the 5′-O leaving group, respectively (Figure 4C). Based upon the following structural and biochemical data, we propose that the C2-amine serves as a general acid to activate the 5′-O leaving group, and the C1-OH stabilizes developing charge on the scissile phosphate.

## General Acid Catalysis by the GlcN6P Amine

All biochemical data collected to date have shown the amine of GlcN6P to be required for catalysis [3, 7, 16]. The GlmS ribozyme positions the C2-amine of GlcN6P 2.7Å from the 5′-O leaving group of G1 (Figure 4C). An equivalent interaction with the C2-OH is made in the structure of the Glc6P bound ribozyme[8], which indicates that the chemical properties, not just the physical position, of the amine are critical to its catalytic contribution. The $pK_a$ of the amine is 8.2 [7], closer to neutrality than any RNA functional group and substantially lower than the $pK_a$ of a hydroxyl group at the same position ($pK_a > 12$). The structure suggests a mechanism involving general acid catalysis with the ammonium form of the GlcN6P donating a proton to the 5′-O leaving group (Figure 4C).

Previous experiments suggested that the amino rather than the ammonium form of GlcN6P activates the ribozyme [7]. This would be most consistent with GlcN6P acting as a general base to activate the 2′-OH nucleophile. In the *T. tengcongenis* GlmS structure, waters observed in the active site were proposed to mediate C2-amine deprotonation of the A-1 2′-OH nucleophile [8]. We do not observe equivalent water molecules in the active site of the *B. anthracis* GlcN6P bound ribozyme, which argues against this mechanism. How can the observed pH dependence be reconciled with the structural proximity of GlcN6P to the 5′-O leaving group?

The reported pH dependence was determined at subsaturating concentrations of effector (10 μM GlcN6P, 10 mM alternative ligands) [7]. As a result, the measured $k_{obs}$ values do not distinguish between a pH dependence on the $K_m$ or the $k_{cat}$ of the reaction. The pH dependence of the effector $K_m$ was tested and found to be minimal, but the experiments were performed across only a narrow pH range (7.5–8.8) near the $pK_a$ (8.2) where the predicted effect on $K_m$ is expected to be small [7].

Because these biochemical data significantly affect the interpretation of the GlmS structures, we determined the pH dependence of both $K_m$ and $k_{cat}$ for the GlcN6P promoted reaction. Consistent with the previous report[7], we found that the reaction rate ($k_{obs}$) at subsaturating concentrations of effector (10 μM GlcN6P) was highly pH dependent (~50 fold change in rate between pH 6.5 and 9.0). However, at saturating concentrations of GlcN6P (10 mM), $k_{cat}$ showed less than a 2-fold change from pH 6.5–9.0 (Figure 6A). By contrast, there was a dramatic GlcN6P $K_m$ dependence on pH (Figure 6B). The $K_m$ was about 1000-fold higher at low pH (~65000 μM at pH 5.5) than at high pH (~50 μM at pH 9.0). In fact the $K_m$ was so large at low pH (5.5–6.5) that it became impossible to reach saturating concentrations of the ligand (Figure 6A). This may explain why Klein and Ferré-D'Amaré were unable to observe a GlcN6P bound complex; the pH of their crystals was between 5.2 and 5.8 [8]. Although the observed rate of 6 $min^{-1}$ is faster than the rate reported in the previous kinetic analysis [7], it is likely that it still does not reflect the rate of chemistry ($k_{cat} \neq k_{chem}$). For example, a rate of 16 $min^{-1}$ has been reported for another GlmS ribozyme construct [13]. Consequently, the pH dependence of $k_{chem}$ remains to be determined, but the data demonstrate a clear effect of pH on $K_m$. These data obviate a model that invokes GlcN6P as a general base in catalysis and do not preclude the structurally predicted mechanism of GlcN6P as a general acid. Physical organic chemistry implicates stabilization of the 5′-O leaving group as the critical factor in the mechanism of phosphodiester cleavage [18]. In the GlmS ribozyme, leaving group stabilization appears to be achieved by the cofactor.

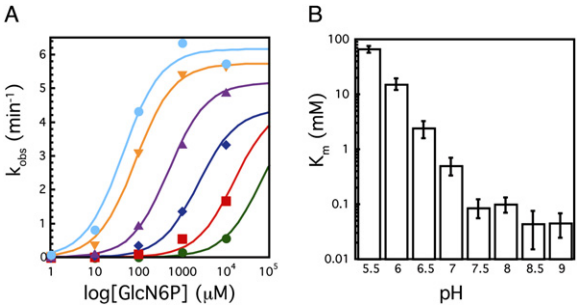



**Figure 6. pH Dependence of the Glms Ribozyme Reaction**
(A) Plot of $k_{obs}(min^{-1})$ versus log concentration of GlcN6P (μM) at various pHs, each curve represents the average of two experiments. Green circles, pH 5.5; red squares, pH 6.0; blue diamonds, pH 6.5; purple triangles, pH 7.0; orange triangles, pH 7.5; blue circles, pH 8.5. Experiments at pH 8 and pH 9 were omitted to simplify the graph. Lines represent best fit to the data. At pH 5.5, 6.0, and, to a less degree, 6.5, saturating GlcN6P concentrations could not be achieved. These data were fit assuming a $k_{cat}$ of 4.5 $min^{-1}$, and the resulting fits are likely to provide a lower estimate for the $K_m$ at the lowest pHs. (B) $K_m$ (mM) versus pH, each bar represents the average of two experiments.

## Transition State Charge Stabilization

GlcN6P could also provide transition state stabilization through interaction between the C1-OH and the scissile phosphate. In this ground state structure, the pro-$S_P$ nonbridging oxygen of the scissile phosphate accepts

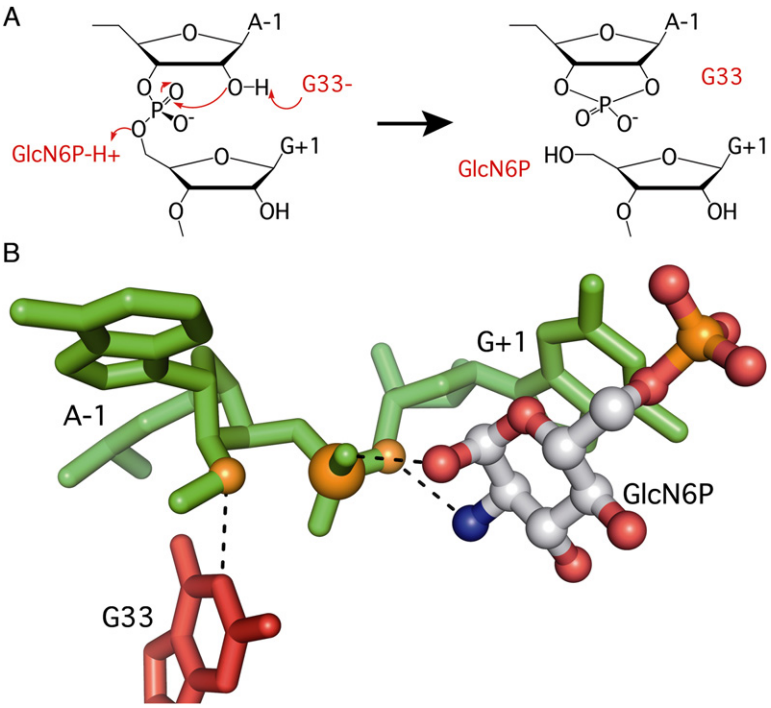



**Figure 7. Proposed Catalytic Mechanism of the GlmS ribozyme at Optimal pH**
(A) In this mechanism, G33 functions as a general base to deprotonate the 2′-OH nucleophile and GlcN6P as a general acid to protonate the 5′-O leaving group. Other functional groups, including the C1-OH of GlcN6P, stabilize charge developing on the scissile phosphate. See text for complete discussion and mechanistic alternatives at neutral pH.
(B) Position of the key functional groups as observed in the ground state structure of the GlcN6P bound GlmS ribozyme.

hydrogen bonds from the exocyclic amine of G57 and the C1-OH of GlcN6P (Figure 4C). Both groups have been shown to contribute significantly to the reaction [7, 15, 16]. The importance of the G57 exocyclic amine is evidenced by interference resulting from inosine or N-methylguanosine substitutions [15]. The importance of the vicinal hydroxyl is further demonstrated in experiments with the two enantiomers of serine [7]. L-serine, which presents the C1-OH and C2-amine functional groups in the same stereochemical orientation as GlcN6P, supports ribozyme activity, but D-serine does not. Introduction of any additional functionalities at the C1 position fully inhibits the GlmS reaction, while removal of the oxygen results in a 70-fold rate reduction [16]. As the reaction proceeds toward the trigonal bipyrimidal transition state, the pro-$S_P$ phosphate oxygen is expected to shift still closer to the GlcN6P C1-OH and the G57 amine. Likewise, the pro-$R_P$ oxygen could form stronger interactions with the G32 N1, the importance of which is still untested. These interactions would provide selective stabilization of the growing negative charge at the transition state.

## Nucleobase Contributions to Catalysis

GlcN6P does not directly interact with the A-1 2′-O nucleophile, so if nucleophilic activation occurs at all, it is likely to come from another source. The closest functional group to the A-1 2′-O, which is methylated in this structure, is the N1 of G33. G33 is orientated through a hydrogen bond network between its 2′-OH and N3 and the 2′-OH of U59. While mutation of U59 does not have a significant effect on catalysis [17], interference analysis found that the U59 2′-OH is absolutely required for activity [15]. Given the proximity of G33 to the nucleophile, we tested its importance by mutating it to A, C, or U. At 10 mM GlcN6P (pH 7.5), the G33A mutation resulted in complete inhibition of the reaction (at least a 10,000-fold effect, $k_{obs} = 0.0005\ min^{-1}$), while the G33U and G33C ribozymes were over 1000-fold slower ($k_{obs} = 0.002\ min^{-1}$ and $0.004\ min^{-1}$, respectively). G33 is completely conserved in GlmS ribozymes [9], while the nucleotide immediately 3′ is less conserved and can be mutated with only a modest (~2-fold) effect on the rate [3].

Catalytically important G residues have been observed in the active sites of other autolytic ribozymes [19, 20]. Structures of the hairpin [20] and hammerhead [19] ribozymes also revealed that the N1 of a G donates a hydrogen bond to the 2′-O. Mutagenesis of these Gs also led to large rate reductions [21, 22]. There are two main hypotheses to explain the mechanistic role of the guanosine in these catalytic RNAs. Nucleobase rescue of an abasic site in the hairpin ribozyme suggested that the G stabilizes the negative charge in the transition state [23]. Alternatively, despite its high $pK_a$, the N1 of G has been proposed to act as the general base to the 2′-OH nucleophile, possibly by adopting an alternative tautomeric form [24, 25]. It is unclear what contribution G33 makes to the GlmS mechanism, though geometric constraints favor the possibility of G33 acting as a general base (Figure 7). Formation of the transition state would appear to move the nonbridging oxygens further from, rather than closer to, the G33 N1, which argues against charge stabilization. A similar orientation of G33 (G40) was observed in the structures from *T. tengcongenis* [8]. In this 2′-O-methyl inhibited structure, there are no functional groups other than G33 sufficiently close to the nucleophile to activate the 2′-OH.

Key determinants in efficient general acid-base catalysis are the $pK_a$ values of the general acid and base and the difference between them [26, 27]. Ideally the $pK_a$s are close to neutral pH and separated by less than 2 pH units to allow both protonation states of the acid and the base to exist in reasonable proportions. The $pK_a$ of the GlcN6P amine is 8.2, while that of the G33 N1 is between 9.3 and 10.4 [28]. At the elevated pH optimum (pH 8–9) of GlmS catalyzed cleavage, the GlcN6P and the G33 N1 are predicted to be in the correct protonation state a sufficient fraction of the time for both to contribute significantly to catalysis. At neutral pH, the predominant species of GlcN6P is the ammonium form, capable of protonating the 5′-O leaving group. The G33 N1 would also be protonated and would function as a poor general base. This suggests that at physiological pH, GlcN6P would contribute substantially more catalytic power than G33.

## SIGNIFICANCE

**Unlike other examples of riboswitches, the GlmS riboswitch is not triggered by a conformational change. Comparison of apo, inhibitor, and activator bound GlmS ribozyme structures reveal that the active site is preorganized in the absence of effector [8]. The nucleophile and leaving group are aligned prior to ligand binding, G33 is properly positioned, but cleavage is slow in the absence of GlcN6P. The GlmS ribozyme is activated by GlcN6P binding, which positions a primary amine next to the leaving group. This is an RNA enzyme that uses a cofactor to achieve efficient catalysis. Like many protein enzymes, this ribozyme catalyzes a reaction by relying upon the unique chemical properties of a small molecule cofactor. Such a catalytic strategy may have broadened the chemical repertoire accessible to an evolving pre-biotic RNA world.**

## EXPERIMENTAL PROCEDURES

### Complex Formation and Crystallization

RNA genes were constructed by fusing the Ban-11U GlmS gene, which has an 11-nucleotide deletion at the 5′ end and the recognition sequence for U1A at the end of P1, and the gene for the antigenomic HDV ribozyme. RNA was prepared by in vitro transcription with T7 polymerase. Posttranscriptional processing of the ag-HDV ribozyme generated homogenous 3′-termini containing a 2′–3′ cyclic phosphate. Oligonucleotides containing active site substitutions were purchased from Dharmacon Research (Lafayette, CO), deprotected by manufacturer-recommended procedures, and used in crystallization trials. The U1A double mutant (Y31H and Q36R) was expressed and purified as described (D. Schechner, personal communication) [10, 29]. RNA (60 μM), the 5′-oligonucleotide (oMeS, 60 μM), and GlcN6P (2 mM) were heated to 70°C for 3 min in folding buffer (10 mM sodium cacodylate [pH 6.8], 10 mM $MgCl_2$, 10 mM KCl) and allowed to cool before addition of the U1A protein (65 μM). Two volumes of the complex were combined with one volume of 11% polyethylene glycol-8000 (PEG-8000), 9% dimethyl sulfoxide (DMSO), 20 mM sodium cacodylate [pH 6.8], 20 mM $MgCl_2$, and 200 mM KCl at 25°C. Crystals were grown by vapor diffusion to a maximum size of 500 × 100 × 100 mm in about 3 days. The crystals were stabilized in a solution containing 35% PEG-8000, 1.5 M 1,6-hexanediol, 20 mM sodium cacodylate (pH 6.8), 20 mM $MgCl_2$, and 200 mM KCl before flash freezing.

### Structure Determination

The iridium hexamine derivative was obtained by soaking native crystals in stabilization solution supplemented with 5 mM iridium hexamine. A pseudonative crystal that was isomorphous to the iridium hexamine derivative was obtained by soaking a native crystal in 2 mM cobalt hexamine. Additional phasing information was obtained from crystals containing an oligonucleotide iodinated at position +11 and soaked in 1 mM cobalt hexamine. The 2.8 Å iridium derivative, 2.75 Å iodine derivative, and 2.75 Å pseudonative data sets were collected at beamline X26C at the National Synchrotron Light Source (NSLS) at 1.154 Å, 1.6 Å, and 1.1 Å wavelengths, respectively. A 2.5 Å native data set was collected at beamline X25 at the NSLS at 0.985 Å wavelength. Data were processed using HKL2000 [30].

To determine the structure of the GlmS ribozyme, 20 iridium hexamine sites were located using single isomorphous replacement and anomalous scattering techniques (SIRAS). The initial sites were found with SHELXD software [31]. SIRAS phasing was performed with SHARP and additional iridium sites (36 total); cobalt hexamine (12) and iodine (4) sites were determined using difference Fourier synthesis. Multiple isomorphous replacement and anomalous scattering (MIRAS) was then used for the final phasing of the experimental data and density modified using SOLOMON and an optimized solvent content of 50% [32]. The initial model was built manually using the MIRAS electron density map and the program O [33]. Strong density was observed for the phosphate backbone. The model was then used in PHASER as a molecular replacement search model for the higher resolution (2.5 Å) native data set [34]. The model was refined against the 2.5 Å native data set using the program REFMAC[35]. Figures were made with PyMOL [36].

### Kinetic Assays

The ribozyme construct (1 μM) was preincubated in a solution containing 25 mM acetate, 25 mM cacodylate, 25 mM HEPES, 25 mM TAPSO, 10 mM $MgCl_2$,and various concentrations of GlcN6P. This four buffer system was used so a single buffering scheme could be implemented over the entire pH range. Buffers were avoided that either competed with GlcN6P binding (such as MES and MOPS) or that had ethanolamine moieties that could promote chemistry in the absence of GlcN6P (such as Tris and glycine). The pH of the solution was adjusted after the addition of GlcN6P because high sugar concentrations were not adequately buffered. The reaction was started by addition of trace 5′ radiolabeled rS oligonucleotide (Dharmacon). Aliquots were taken at appropriate time points, quenched in formamide loading buffer (95% formamide, 2.5 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol) and analyzed by denaturing polyacrylamide gel electrophoresis. Gels were visualized using a STORM PhosphoImager (GE Healthcare), and the data quantitated using ImageQuant (GE Healthcare). Rates ($k_{obs}$) were determined using the equation

$$f = (1 - f_\infty) * \exp(-k_{obs} * t) + f_\infty;$$

$$f = \text{fraction unreacted}, f_\infty = \text{fraction unreacted at } t = \infty, t = \text{time}$$

and fit using the least squares method implemented in KaleidaGraph (Synergy Software). $K_m$ and $k_{cat}$ at each pH were determined using the equation

$$k_{obs} = (k_{cat} * C)/(C + K_m); C = \text{concentration GlcN6P}$$

### Supplemental Data

Supplemental data include one table and are available at http://www.chembiol.com/cgi/content/full/14/1/97/DC1/.

## ACKNOWLEDGMENTS

We would like to thank M. Becker, H. Robinson, A. Héroux, S. Myers, and the beamline staff at X25, X26C, and X29 at the NSLS at Brookhaven National Laboratory; M. Strickler and the CSB core staff; R. Breaker, Y. Modis, J. Wang, A. Roth, D. Shechner, I. Sudyam, D. Hiller, M. Stahley, R. Voorhees and other members of the Strobel Lab for discussions and comments on the manuscript; Y. Xiong for all of his help;

R. Batey for the gift of iridium hexamine; and M. Calabrese for mass spectrometry. This work was supported by grants from the National Science Foundation (MCB0544255) and the National Institutes of Health (GM022778) to S.A.S.

Received: November 17, 2006
Accepted: December 8, 2006
Published Online: December 28, 2006

## REFERENCES

1. Winkler, W.C., and Breaker, R.R. (2005). Regulation of bacterial gene expression by riboswitches. Annu. Rev. Microbiol. *59*, 487–517.
2. Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., and Breaker, R.R. (2003). Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. Cell *113*, 577–586.
3. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. Nature *428*, 281–286.
4. Milewski, S. (2002). Glucosamine-6-phosphate synthase–the multi-facets enzyme. Biochim. Biophys. Acta *1597*, 173–192.
5. Hampel, K.J., and Tinsley, M.M. (2006). Evidence for preorganization of the glmS ribozyme ligand binding pocket. Biochemistry *45*, 7861–7871.
6. Fedor, M.J., and Williamson, J.R. (2005). The catalytic diversity of RNAs. Nat. Rev. Mol. Cell Biol. *6*, 399–412.
7. McCarthy, T.J., Plog, M.A., Floy, S.A., Jansen, J.A., Soukup, J.K., and Soukup, G.A. (2005). Ligand requirements for glmS ribozyme self-cleavage. Chem. Biol. *12*, 1221–1226.
8. Klein, D.J., and Ferre-D'Amare, A.R. (2006). Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. Science *313*, 1752–1756.
9. Roth, A., Nahvi, A., Lee, M., Jona, I., and Breaker, R.R. (2006). Characteristics of the glmS ribozyme suggest only structural roles for divalent metal ions. RNA *12*, 607–619.
10. Ferre-D'Amare, A.R., and Doudna, J.A. (2000). Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module. J. Mol. Biol. *295*, 541–556.
11. Ferre-D'Amare, A.R., Zhou, K., and Doudna, J.A. (1998). Crystal structure of a hepatitis delta virus ribozyme. Nature *395*, 567–574.
12. Serganov, A., Keiper, S., Malinina, L., Tereshko, V., Skripkin, E., Hobartner, C., Polonskaia, A., Phan, A.T., Wombacher, R., Micura, R., et al. (2005). Structural basis for Diels-Alder ribozyme-catalyzed carbon-carbon bond formation. Nat. Struct. Mol. Biol. *12*, 218–224.
13. Wilkinson, S.R., and Been, M.D. (2005). A pseudoknot in the 3′ non-core region of the glmS ribozyme enhances self-cleavage activity. RNA *11*, 1788–1794.
14. Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R., and Patel, D.J. (2006). Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. Nature *441*, 1167–1171.
15. Jansen, J.A., McCarthy, T.J., Soukup, G.A., and Soukup, J.K. (2006). Backbone and nucleobase contacts to glucosamine-6-phosphate in the glmS ribozyme. Nat. Struct. Mol. Biol. *13*, 517–523.
16. Lim, J., Grove, B.C., Roth, A., and Breaker, R.R. (2006). Characteristics of ligand recognition by a glmS self-cleaving ribozyme. Angew. Chem. Int. Ed. Engl. *45*, 6689–6693.
17. Soukup, G.A. (2006). Core requirements for glmS ribozyme self-cleavage reveal a putative pseudoknot structure. Nucleic Acids Res. *34*, 968–975.
18. Oivanen, M., Kuusela, S., and Lonnberg, H. (1998). Kinetics and mechanisms for the cleavage and isomerization of the phosphodiester bonds of RNA by bronsted acids and bases. Chem. Rev. *98*, 961–990.
19. Martick, M., and Scott, W.G. (2006). Tertiary contacts distant from the active site prime a ribozyme for catalysis. Cell *126*, 309–320.
20. Rupert, P.B., and Ferre-D'Amare, A.R. (2001). Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. Nature *410*, 780–786.
21. Pinard, R., Hampel, K.J., Heckman, J.E., Lambert, D., Chan, P.A., Major, F., and Burke, J.M. (2001). Functional involvement of G8 in the hairpin ribozyme cleavage mechanism. EMBO J. *20*, 6434–6442.
22. Ruffner, D.E., Stormo, G.D., and Uhlenbeck, O.C. (1990). Sequence requirements of the hammerhead RNA self-cleavage reaction. Biochemistry *29*, 10695–10702.
23. Kuzmin, Y.I., Da Costa, C.P., and Fedor, M.J. (2004). Role of an active site guanine in hairpin ribozyme catalysis probed by exogenous nucleobase rescue. J. Mol. Biol. *340*, 233–251.
24. Han, J., and Burke, J.M. (2005). Model for general acid-base catalysis by the hammerhead ribozyme: pH-activity relationships of G8 and G12 variants at the putative active site. Biochemistry *44*, 7864–7870.
25. Wilson, T.J., Ouellet, J., Zhao, Z.Y., Harusawa, S., Araki, L., Kurihara, T., and Lilley, D.M. (2006). Nucleobase catalysis in the hairpin ribozyme. RNA *12*, 980–987.
26. Bevilacqua, P.C. (2003). Mechanistic considerations for general acid-base catalysis by RNA: revisiting the mechanism of the hairpin ribozyme. Biochemistry *42*, 2259–2265.
27. Cleland, W.W. (1977). Determining the chemical mechanisms of enzyme-catalyzed reactions by kinetic studies. Adv. Enzymol. Relat. Areas Mol. Biol. *45*, 273–387.
28. Da Costa, C.P., and Sigel, H. (2003). Acid-base and metal ion binding properties of guanylyl(3′->5′)guanosine (GpG(-)) and 2′-deoxyguanylyl(3′-5′)-2′-deoxyguanosine [d(GpG)(-)] in aqueous solution. Inorg. Chem. *42*, 3475–3482.
29. Oubridge, C., Ito, N., Evans, P.R., Teo, C.H., and Nagai, K. (1994). Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature *372*, 432–438.
30. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.
31. Schneider, T.R., and Sheldrick, G.M. (2002). Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. *58*, 1772–1779.
32. Abrahams, J.P., and Leslie, A.G. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D Biol. Crystallogr. *52*, 30–42.
33. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.
34. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. *61*, 458–464.
35. Winn, M.D., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. D Biol. Crystallogr. *57*, 122–133.
36. Delano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).

### Accession Numbers

Coordinates for the 2.5 Å complex of GlcN6P bound to the GlmS ribozyme have been deposited in the Protein Data Bank under accession code 2NZ4 and in the Nucleic Acid Database under accession code PR0249.