HHS Public Access

Author manuscript

Structure. Author manuscript; available in PMC 2018 January 03.

Published in final edited form as:

Structure. 2017 January 03; 25(1): 195–202. doi:10.1016/j.str.2016.11.020.

Structural basis for ligand binding to the guanidine-I riboswitch

Caroline Reiss[1,2], **Yong Xiong**[1], and **Scott Strobel**[1,2,*]

[1]Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, CT 06511, USA

[2]Chemical Biology Institute, Yale University, West Haven, CT 06516, USA

SUMMARY

The guanidine-I riboswitch is a conserved RNA element with approximately 2000 known examples across four phyla of bacteria. It exists upstream of nitrogen metabolism and multidrug resistance transporter genes and alters expression through the specific recognition of free guanidinium cation. Here we report the structure of a guanidine riboswitch aptamer from *Sulfobacillus acidophilus* at 2.7Å resolution. Helices P1, P1a, P1b, and P2 form a coaxial stack that acts as a scaffold for ligand binding. A previously unidentified P3 helix docks into P1a to form the guanidinium binding pocket, which is completely enclosed. Every functional group of the ligand is recognized through hydrogen bonding to guanine bases and phosphate oxygens. Guanidinium binding is further stabilized through cation-π interactions with guanine bases. This allows the riboswitch to recognize guanidinium while excluding other bacterial metabolites with a guanidino group, including the amino acid arginine.

Graphical Abstract

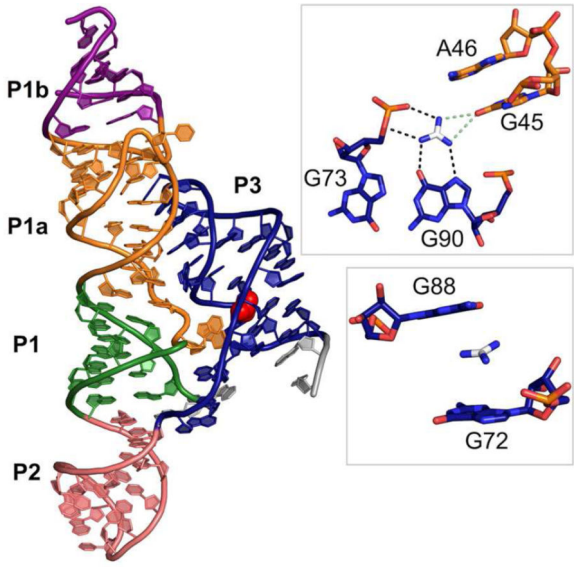
*corresponding author, lead contact, Dr. Scott A. Strobel, scott.strobel@yale.edu.

Accession number

Coordinates for the guanidine-I riboswitch have been deposited in the Protein Data Bank under accession code 5T83.

Author Contributions

C.W.R. determined the crystal structure and performed biochemical experiments. Y.X. helped significantly with data processing and discussion of the project. C.W.R. and S.A.S. prepared the manuscript. S.A.S. advised on the project.



INTRODUCTION

The *ykkC/yxkD* RNA was first identified as a riboswitch by bioinformatics in 2004. It was dubbed an “orphan” riboswitch, because the ligand had not yet been identified (Barrick et al., 2004). Ligand identification has proven elusive due to the wide variety of poorly studied genes under its control (Meyer et al., 2011). Nelson et al. recently demonstrated that guanidine is the *ykkC* ligand (Nelson et al., 2016).

Riboswitches are employed by bacteria to control gene expression for crucial metabolic pathways, including one-carbon and purine metabolism (Kim et al., 2015; Mandal and Breaker, 2004a; Mandal et al., 2003). Some riboswitches, like the fluoride riboswitch, control genes related to detoxification of the ligand (Baker et al., 2012). Although guanidine is a fragment within key metabolites such as guanine and arginine, guanidine had not previously been linked to any metabolic pathway in bacteria, suggesting that there is more to learn about guanidine biochemistry.

Genes associated with this riboswitch include genes annotated as urea carboxylases, urea carboxylase-associated proteins, and a variety of transporters, including multidrug resistance efflux proteins (Nelson et al., 2016). The guanidine riboswitch exists in the 5’ untranslated region of these genes. Nelson et al. examined a gene from *Oleomonas sagaranensis* previously characterized as a urea carboxylase, yet the catalytic efficiency is 40 times better for guanidine than urea, suggesting that it is a guanidine carboxylase that has been mis-annotated as a urea carboxylase. They also investigated a transporter (a SugE protein) in *Clostridiales* controlled by the guanidine-I riboswitch. They found that intrinsic fluorescence of tryptophan residues increases in a guanidine-dependent manner (K_d ~1 mM), providing preliminary evidence that this transporter specifically recognizes guanidine (Nelson et al., 2016).

Nelson et al. also demonstrated that guanidine binding to the aptamer turns on gene expression in *Bacillus subtilis* and *Desulfotomaculum ruminis*. Transcription termination assays in *D. ruminis* show that upregulation of gene expression occurs in a guanidine-dependent manner via destabilization of a terminator stem (Nelson et al., 2016). This is comparable to switching mechanisms used in other riboswitches (For review see Breaker, 2012; Mandal and Breaker, 2004). This, combined with data regarding downstream gene function, suggests that guanidine turns on expression of genes that either enzymatically act on guanidine or transport it out of the cell.

To provide evidence that the riboswitch controlling these genes is specifically recognizing guanidine rather than another guanidino-containing compound, we set out to determine the crystal structure of the guanidine-I riboswitch bound to guanidine via X-ray crystallography. Specific recognition of guanidine by the RNA poses a challenge because the ligand is a common functional group in many other metabolites. The binding pocket must exclude larger guanidino-containing compounds while still maintaining sufficient guanidine affinity. Arginine and agmatine both contain guanidino groups and are present at mid to high micromolar concentrations in bacteria (Bennett et al., 2009; Caldara et al., 2008; Hamana, 1996; Ishii et al., 2007). These compounds do not bind the riboswitch, suggesting they are excluded from the binding site (Nelson et al., 2016). The riboswitch also has no apparent affinity for urea, which is produced by the urea cycle in bacteria and differs from guanidine by only a single functional group, where one amine ($C\text{-}NH_2$) in guanidinium is a carbonyl group (C=O) in urea (Nelson et al., 2016).

The consensus motif for the guanidine riboswitch generated by bioinformatics provides less structural information than usual due to a particularly high degree of sequence conservation. Three short helices, P1, P1b, and P2, were identified by covariation analysis. Of the remaining 51 nucleotides that do not exhibit covariation or form variable hairpins, 35 are greater than 97% conserved. The predicted secondary structure has a large, highly conserved internal loop between P1 and P1b and a long, highly conserved tail at the 3’ end of the RNA (Nelson et al., 2016). Such a high degree of conservation suggests that the riboswitch contains additional structural elements including many sequence-specific tertiary interactions. Here we present the structure of the aptamer domain of the guanidine-I riboswitch from *Sulfobacillus acidophilus* bound to guanidine at 2.7 Å resolution.

RESULTS

Structure determination of the guanidine-I riboswitch

We targeted the aptamer domain of the guanidine-I riboswitch from *S. acidophilus* for crystallization. This riboswitch controls the expression of a protein annotated as urea carboxylase-associated protein 2, though the gene may actually be a guanidine carboxylase-associated protein. The crystallization construct includes 94 nucleotides from the native sequence, plus an added G on the 5’-end to aid in efficient transcription. The non-conserved P1b helix (nt 23–26 and 30–33) was mutated from the wild type sequence CUAG/CUAG to CGUC/GGUG to introduce a tandem GU wobble motif. This tandem GU motif is an ideal binding site for iridium hexamine, a metal complex with anomalous signal for use in structure determination (Keel et al., 2007). Initial phases were generated using a three-

wavelength MAD dataset on a co-crystal of the riboswitch and iridium hexamine, a SAD dataset collected on a co-crystal of the riboswitch and iridium hexamine, and a native dataset isomorphous to the SAD dataset (Table 1). Three iridium hexamine sites were identified and used for phasing. A model of the *S.* acidophilus guanidine riboswitch was built into the unbiased density and refined using the high remote wavelength dataset from MAD collection.

Overall architecture

The structure of the guanidine-I riboswitch is dominated by a long continuous set of coaxially stacked helices that includes P1, P1a, P1b, and P2. The large internal loop between P1 and P1b forms a helix, P1a, that extends this helical stack. An additional helix, P3, forms between conserved nucleotides in the extended tail at the 3' end of the sequence. Helices P1a and P3 were not previously identified by bioinformatics, because the very high level of sequence conservation and limited canonical pairing prevented conclusions based upon covariation analysis (Meyer et al., 2011; Weinberg et al., 2007, 2010). The P1a and P3 helices form sequence-specific tertiary interactions, consistent with this high degree of conservation. The guanidine-I riboswitch consensus motif has been updated to reflect the secondary structure observed in the crystal structure (Figure 1A). The secondary structure of the crystallized construct and a ribbon diagram of the crystal structure are shown in Figure 1B and 1C, respectively.

Guanidine recognition by the riboswitch

The guanidine binding pocket is formed at an interface between the P3 and P1a helices. Electron density for the guanidinium can be seen in the unbiased electron density map initially generated from phasing (Figure 2A). The nucleotides surrounding the guanidine ligand are highly conserved. Only a small tunnel is left solvent-accessible by the RNA and this is blocked by a metal ion that is visible in the crystal structure (Sr-7) (Figure 2D).

The ligand is exhaustively recognized by the RNA. The pK_a of guanidine is 13.6 (Perrin, 1965) and is present almost exclusively as the guanidinium cation, which has three-fold symmetry and a positive charge stabilized by resonance (Figure 1D). The guanidinium has six hydrogen bond donors, two on each nitrogen. All potential hydrogen bonds are satisfied by the RNA. Guanidinium donates hydrogen bonds to three nucleotides, G45, G90, and G73, one nucleotide for each of its three edges. On one edge, the ligand donates two hydrogen bonds to G90 in the P3 helix in a bifurcated manner. One amine donates to the N7 imino of G90 and another donates to the O6 carbonyl oxygen. On the second edge, two hydrogen bonds are made between the two guanidinium amines and the 5' bridging and pro-R_p non-bridging phosphate oxygens of G73 (Figure 2B). The proximity of the positively charged guanidinium cation and the negatively charged phosphate of G73 is also likely to provide an ionic stabilizing effect on ligand binding. On the third guanidinium edge, two nitrogens are approximately equidistant from the carbonyl oxygen of G45 (3.3 Å and 3.4 Å). They could both donate a hydrogen bond; however, they appear to be mutually exclusive of each other based upon the narrow bond angle that would be formed between them (39.1°). We predict that both are alternatingly sampled upon ligand binding (Figure 2B).

In addition to hydrogen bonding interactions on the guanidinium edges, there are also contacts with the top and bottom faces of the ligand. The guanidinium is positioned to potentially engage in cation-π interactions with G72 and G88 (Figure 2C). In this interaction, positively charged species form noncovalent interactions with the electron-rich π orbitals of an aromatic system, like the purine bases (Dougherty, 2013). The distances between the guanidinium and the 6-membered rings in G72 and G88 are 4.1 Å and 4.3 Å, respectively. The θ angles between guanidinium and the 6-membered rings of each base are 21.0° for G72 and 44.5° for G88. The distances and angles are consistent with cation-π interactions found in proteins and RNA (Gallivan and Dougherty, 1999; Zhang et al., 2014).

Surface models demonstrate that the guanidinium binding pocket is almost completely enclosed (Figure 2D). A calculation of the solvent accessible surface area (SASA) in PyMol shows that the SASA of guanidinium in complex with the riboswitch is 0.30 $Å^2$, compared to 62.68 $Å^2$ of total surface area in solution. Expressed another way, 99.5% of the guanidinium ligand is buried within the riboswitch. The surface model demonstrates that there is little room in the binding pocket for anything larger than guanidinium to bind (Figure 2D). There is a small pocket of extra space next to one guanidinium NH_2 that would potentially be large enough to accommodate one additional atom. Consistent with this, methylguanidine, hydroxyguanidine, and aminoguanidine turn on expression in a *B. subtilis* guanidine riboswitch reporter system, but it is a weak response relative to that of guanidine (Nelson et al., 2016). The small size of the binding pocket supports the finding that other common metabolites containing guanidino groups, including arginine, are too large and are not able to efficiently bind the riboswitch (Nelson et al, 2016).

The architecture of the binding site is formed using an S-turn motif involving G90 and G88 (Figure 2E). This motif is characterized by two successive bends in the RNA backbone, resulting in the eponymous “S” shape. A single base, in this case G88, is oriented in the reverse direction with the sugar adopting a C2’-endo conformation (Hendrix et al., 2005). This allows G88 to be positioned above the ligand to form a putative cation-π interaction with guanidinium. Nucleotide G90 is positioned in the normal orientation to hydrogen bond with the guanidinium. G89 is bulged out and forms a Watson-Crick (WC) base pair with highly conserved C69, cinching the two strands of the P3 helix together near the binding pocket (Figure S1A).

The nucleotides interacting with guanidinium form base-pairing networks with each other and with other nucleotides in the P3 helix. G90 and G73 form a *trans* Watson-Crick/Hoogsteen base pair, which orients G90 and G73 to form their respective hydrogen bonds with guanidinium (Figure S1B). G45 and A46 flip out from helix P1a and insert into P3 (Figure S1C). The 2’-hydroxyl of A46 forms a single hydrogen bond with a non-bridging phosphate oxygen on C71, which likely plays a role in stabilizing its position in the P3 helix (Figure S1C). G45 forms a sugar-edge/sugar-edge base pair in the *trans* orientation with A70 as well as a hydrogen bond between its N1 imino and the carbonyl oxygen of G72, possibly to stabilize G45 in the flipped-out position (Figure S1D). A sheared base pair is formed between A91, which stacks on G90, and G72 (Figure S1D). This positions G72 to form the predicted cation-π interactions with guanidinium (Figure 2C). The N2 of G72 donates a hydrogen bond to the pro-R_p non-bridging phosphate oxygen of G90.

To further probe the importance of bases near the guanidinium, we used equilibrium dialysis to measure the K_d for individual mutants at four conserved residues within the binding pocket (Figure 3). Under these conditions, the guanidinium (a chaotropic agent) is present in trace amounts, eliminating the possibility of denaturing effects on the RNA. The wild type *S. acidophilus* guanidine riboswitch and the crystal construct with an engineered iridium hexamine binding site have an equilibrium binding affinity of 20 μM, which is comparable to the affinity measured by in-line probing for guanidine-I riboswitches from other bacteria (Nelson et al, 2016). In contrast, G45A and G90A both exhibited less than 10% binding at the highest RNA concentration possible in the assay. This corresponds to binding defects of more than 100-fold. Mutation of either G45 or G90 is predicted to disrupt a hydrogen bond between a guanidinium NH_2 and a carbonyl oxygen on the RNA base. We also measured binding to the G72C mutation to explore the effects of possible cation-π interactions. G72C showed less than 10% binding at the highest concentration range tested. The A46C mutant showed a 40-fold binding defect, consistent with the structural observation that guanidinium H-bond donors are too distant to form strong hydrogen bonds with the N1 of A46 (3.5 Å). This nucleotide may sterically constrain the size of the binding pocket.

Critical tertiary contacts between helix P1a and helix P3

The main tertiary contacts within the aptamer are formed by docking the P3 helix into P1a, an interaction that is critical for creating the binding pocket. P3 specifically recognizes the P1a minor groove through A-minor interactions. A highly conserved series of two base triples and a base quadruple in P1a open the minor groove, which allows contact with a highly conserved run of A's in the hairpin loop of P3. A81, A82, and A83 form A-minor interactions with the P1a helix (Figures 4A and 4B). A81 and A82 form type II A-minor interactions with G17 and G16, respectively, while A83 forms a type I A-minor interaction with the G15-C42 base pair (Figure 4B).

Two highly conserved base triples and a highly conserved base quadruple distort the P3 minor groove. To form internal base triples from only two RNA strands, a number of nucleotides bulge out and reinsert into either the major or minor groove of neighboring base pairs. The G22-C36 base pair in P1a (>90% conserved) stacks directly below the C23-G33 base pair at the bottom of helix P1b (not pictured). U35 is flipped outward. G34 (>97% conserved) bulges and re-inserts into the major groove of P1a. G34 forms a base triple with the C21-G38 WC base pair (both >97% conserved) (Figure 4C). Non-conserved nucleotide U37 bulges out and forms a weak dinucleotide platform with C36 (not pictured). The G39-C20 WC base pair (>97% conserved) forms a base triple with U18 (>97% conserved), in which the primary amine of C20 donates a hydrogen bond to the O1 of U18 (Figure 4C). U19 bulges into the minor groove and engages in a base quadruple.

A base quadruple is formed by U40, G17, U19, and A80 (Figure 4C). U40 and G17 engage in a pseudo-GU wobble with only one hydrogen bond between the N1 imino of G17 and the O4 of U40. The O4 of U19 accepts two hydrogen bonds, one from the N3 imino of U40 and one from the primary amine of G17. The primary amine of G17 also donates a hydrogen bond to the N1 imino of A80. This base quadruple is further stabilized by stacking of U19 and A80. The set of base triples distorts the P1a helix, allowing a base quadruple to form

that specifically interacts with A80 to help anchor the A-rich hairpin loop into the minor groove of P1a.

Two metal ions also appear to stabilize the tertiary docking of P3 into P1a. The two metals (Sr-3 and Sr-7) facilitate the close approach of the phosphate backbones of the two helices by coordinating between the non-bridging phosphate oxygens of G45, C86, and C85 (Fig. 4D). Both of these metal ions are Sr^{2+} in the crystal structure, since the crystallization condition contains $SrCl_2$. Under cellular conditions, we expect that these are magnesium ions that make similar inner-sphere coordinations. One metal ion (Sr-7 in the crystal structure) coordinates to the pro-S_P oxygen of G45 (2.3 Å), the pro-S_P oxygen of C85 (2.3 Å), and the pro-R_P oxygen of C86 (2.3 Å). The bond distances are consistent with inner sphere coordination. The other metal ion (Sr-3) coordinates to the pro-R_P oxygen of G45 (2.3 Å) and the pro-S_P oxygen of C86 (2.2 Å), again with bond distances consistent with inner sphere coordination. The distance of Sr-3 to the pro-S_P oxygen of A44 is 3.3Å suggestive of water-mediated coordination to this phosphate. As previously noted, Sr-7 appears to close the guanidinium binding pocket, thus limiting the size of the ligand (Figure 4D).

DISCUSSION

Compounds containing guanidino groups, including arginine and agmatine, are prevalent in bacterial cells. In this environment, the guanidine riboswitch faces the challenge of specifically recognizing guanidinium while excluding larger compounds with a guanidino functional group. The riboswitch accomplishes this by recognizing every possible surface of guanidinium, including all three edges and both faces, leaving almost no solvent-exposed surface area around the ligand.

One clear pattern in the recognition of guanidinium is the presence of electronegative oxygen atoms in the binding pocket, which act as hydrogen bond acceptors for the NH_2 groups in guanidinium. The only non-oxygen hydrogen bond acceptor interacting with guanidinium is the N7-imino on G90. Despite possessing a three-fold axis of symmetry, the three identical NH_2 groups in guanidinium are recognized in an asymmetric fashion. Above and below the two-fold plane of symmetry, the recognition is pseudo-symmetric. Guanine bases above and below guanidinium (G72 and G88) are predicted to form cation-π interactions. In protein-RNA interactions, the guanidino group of arginine commonly interacts with RNA bases via cation-π interactions, suggesting that the guanidinium ion alone could use the same mechanism of binding to RNA (Gallivan and Dougherty, 1999; Lustig et al., 1997; Zhang et al., 2014).

The residues involved in guanidine binding within the crystal structure are the same ones that undergo a conformational change upon guanidine addition. In-line probing data shows six nucleotides have reduced strand cleavage upon guanidine binding (Nelson et al., 2016). Four of these nucleotides (G45, A46, G72, and G73 in *S. acidophilus*) are in the immediate vicinity of guanidinium. The other two nucleotides are A70 and C71. As previously mentioned, A70 contacts G45, which directly contacts guanidinium. C71 forms a Watson-

Crick base pair with G92 at the base of the P3 helix, suggesting that guanidine binding stabilizes the P3 helix.

Two distinct subtypes of the *ykkC* RNA have been identified (Nelson et al., 2016). The guanidine-I riboswitch, subtype 1, is the most common. Subtype 2 does not bind guanidine and the natural ligand has not yet been determined. The two subtypes differ at nucleotides 46, 72, and 88–91, all of which cluster in the binding site, further suggesting that these nucleotides are crucial for specific recognition of guanidine in subtype 1. This also implies that subtype 2 achieves specificity for a different ligand by altering only these six binding site nucleotides.

The P3 helix, which was not predicted by bioinformatics, is formed at the 3' end of the riboswitch. This helix appears to be stabilized by guanidine binding and is most likely where switching of gene expression occurs. A potential terminator stem in the *S. acidophilus* riboswitch (Figure S2) includes nucleotides 84 through 94 of the aptamer. Formation of this helix would sequester six nucleotides that are more than 97% conserved, including G90 in the binding pocket and G89 in the S-turn. In the guanidinium-bound state, nucleotides 84–94 form eight base pairs with other nucleotides in the aptamer, four of which are canonical GC pairs. In the unbound state, assuming that the P3 helix does not form, residues 84–94 could form eleven base pairs within the terminator stem, six of which are canonical GC pairs. Guanidine-dependent switching is likely to occur by stabilization of the P3 helix, allowing it to outcompete terminator stem formation in the presence of the ligand.

The structure of this riboswitch reveals the basis for specific binding of guanidinium to RNA. It also reveals two previously unpredicted helices, P1a and P3. Both of the unpredicted helices form complex tertiary interactions crucial for riboswitch folding and guanidinium recognition. P3 contains the binding pocket and a portion of the predicted terminator stem, providing a possible mechanism for switching through stabilization of the P3 helix upon guanidinium binding. This allows guanidinium binding to modulate the expression of downstream genes through the riboswitch. The extensive and selective recognition of guanidinium by the riboswitch provides further evidence that free guanidine is a biologically relevant compound encountered by bacteria. Other biochemical data shows that genes controlled by this riboswitch act on guanidine as a substrate or remove it from the cell. Collectively, these data suggest that guanidine is biologically relevant compound in bacteria whose concentration must be managed to avoid toxicity.

EXPERIMENTAL PROCEDURES

RNA preparation and purification

Plasmids containing the *ykkC* RNA from *Sulfobacillus acidophilus* were obtained from GeneArt at ThermoFisher Scientific. RNA was transcribed using a T7 RNA polymerase and purified natively by gel filtration in the presence of 1 mM guanidine (Sigma Aldrich).

Crystallization and structure determination

RNA (130 μM) in a folding buffer (10 mM $MgCl_2$, 10 mM KCl, 10 mM HEPES-KOH, pH 7.5, and 10 mM guanidine) was mixed 1:1 with a solution of 32% MPD, 40 mM $SrCl_2$, 12

mM spermine, 40 mM Na-acetate, pH 4.6, and 800 μM iridium (III) hexamine. Crystals were grown at 23°C using the microbatch-under-oil method using 2:1 paraffin:silicon oil. Crystals appeared within 24 hours. The crystals were flash frozen in liquid nitrogen without any further preparation.

Unbiased phases were generated using a combination of SIRAS and MAD phasing. SAD data collected at the iridium hexamine peak wavelength and a native dataset were collected at beamline 19BM at the Advanced Photon Source (APS). Three-wavelength MAD data were collected at beamline 8.2.2 at the Advanced Light Source (Berkeley, CA). Data were processed using HKL2000. Phenix AutoSol was used to locate heavy atom sites and to generate initial phases with both the SIRAS data and the MAD data. Three iridium sites were located with AutoSol and verified using anomalous difference Fourier methods. Model building was done in Coot. Refmac and Phenix Refine were used to for refinement. Each strontium site was identified by first modeling a magnesium ion and then using difference Fourier methods to determine if density was still unaccounted for. Figures of the crystal structure were made in PyMol.

K_d determination of RNAs by equilibrium dialysis

Natively purified RNAs were buffer exchanged into an RNA-folding solution containing 50 mM HEPES-KOH, pH 7.5, 200 mM KCl, and 20 mM $MgCl_2$. The RNA and [^{14}C]-guanidine-HCl (Moravek Biochemicals) were added to separate sides of a 10,000 MWCO Dispo Equilibrium Dialyzer from Harvard Apparatus and were equilibrated overnight at room temperature. The amount of [^{14}C]-guanidine on each side of the dialysis membrane was determined by scintillation counting in Ultima Gold on a PerkinElmer Tri-Carb 4910 TR scintillation counter. The fraction bound was determined by the following equation:

$$\text{Fraction bound} = \frac{(\text{cpm on RNA side} - \text{cpm on ligand side})}{(\text{cpm on RNA side})}$$

The data were fit in GraphPad Prism to the following equation:

$$F = \frac{B_{max} * X}{(K_d + X)}$$

where F= fraction bound, B_{max}= fraction bound at saturation, and X= riboswitch concentration. For mutants that did not reach saturation over the riboswitch concentrations tested, B_{max} was fixed at 1.0, in agreement with the B_{max} of the wild type and crystallization constructs.

Acknowledgments

We thank the synchrotron beamline staff at the Structural Biology Center (SBC) and Northeastern Collaborative Access Team (NE-CAT) at the Advanced Photon Source as well as the Berkeley Center for Structural Biology (BCSB) at the Advanced Light Source for their assistance; J. Wang, who provided important help with data processing, and M. Strickler from the Yale Center for Structural Biology; D. Hiller and other members of the Strobel Lab for valuable advice and discussion; and R. Breaker, M. Sherlock, R. Atilho, and J. Nelson for providing data on the riboswitch in advance of publication and for helpful discussion. C.W.R. was supported by the NIH

Cellular and Molecular Biology Training Grant (T32GM007223). This work was also supported by an NIH grant to S.A.S. (GM022778).

References

Baker JL, Sudarsan N, Weinberg Z, Roth A, Stockbridge RB, Breaker RR. Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. Science. 2012; 335:233–235. [PubMed: 22194412]

Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, Lee M, Roth A, Sudarsan N, Jona I, et al. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. Proc. Natl. Acad. Sci U.S.A. 2004; 101:6421–6426. [PubMed: 15096624]

Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nat. Chem. Biol. 2009; 5:593–599. [PubMed: 19561621]

Breaker RR. Riboswitches and the RNA World. Cold Spring Harb. Perspect. Biol. 2012; 4

Caldara M, Dupont G, Leroy F, Goldbeter A, Vuyst LD, Cunin R. Arginine biosynthesis in Escherichia coli: experimental perturbation and mathematical modeling. J. Biol. Chem. 2008; 283:6347–6358. [PubMed: 18165237]

Dougherty DA. The Cation-π Interaction. Acc. Chem. Res. 2013; 46:885–893. [PubMed: 23214924]

Gallivan JP, Dougherty DA. Cation-π interactions in structural biology. Proc. Natl. Acad. Sci. 1999; 96:9459–9464. [PubMed: 10449714]

Hamana K. Distribution of diaminopropane and acetylspermidine in Enterobacteriaceae. Can. J. Microbiol. 1996; 42:107–114. [PubMed: 8742354]

Hendrix DK, Brenner SE, Holbrook SR. RNA structural motifs: building blocks of a modular biomolecule. Q. Rev. Biophys. 2005; 38:221–243. [PubMed: 16817983]

Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, et al. Multiple High-Throughput Analyses Monitor the Response of Ecoli to Perturbations. Science. 2007; 316:593–597. [PubMed: 17379776]

Keel AY, Rambo RP, Batey RT, Kieft JS. A General Strategy to Solve the Phase Problem in RNA Crystallography. Structure. 2007; 15:761–772. [PubMed: 17637337]

Kim PB, Nelson JW, Breaker RR. An Ancient Riboswitch Class in Bacteria Regulates Purine Biosynthesis and One-Carbon Metabolism. Mol. Cell. 2015; 57:317–328. [PubMed: 25616067]

Lustig B, Arora S, Jernigan RL. RNA base-amino acid interaction strengths derived from structures and sequences. Nucleic Acids Res. 1997; 25:2562–2565. [PubMed: 9185564]

Mandal M, Breaker RR. Adenine riboswitches and gene activation by disruption of a transcription terminator. Nat. Struct. Mol. Biol. 2004a; 11:29–35. [PubMed: 14718920]

Mandal M, Breaker RR. Gene regulation by riboswitches. Nat. Rev. Mol. Cell Biol. 2004b; 5:451–463. [PubMed: 15173824]

Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. Riboswitches Control Fundamental Biochemical Pathways in Bacillus subtilis and Other Bacteria. Cell. 2003; 113:577–586. [PubMed: 12787499]

Meyer MM, Hammond MC, Salinas Y, Roth A, Sudarsan N, Breaker RR. Challenges of ligand identification for riboswitch candidates. RNA Biol. 2011; 8:5–10. [PubMed: 21317561]

Nelson JW, Atilho RM, Sherlock ME, Reiss CW, Strobel SA, Stockbridge RB, Breaker RR. Metabolism of Free Guanidine in Bacteria is Regulated by Widespread Riboswitch Classes. Mol. Cell TBD. 2016

Perrin, DD. Dissociation constants of organic bases in aqueous solution. London: Butterworths; 1965.

Weinberg Z, Barrick JE, Yao Z, Roth A, Kim JN, Gore J, Wang JX, Lee ER, Block KF, Sudarsan N, et al. Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. Nucleic Acids Res. 2007; 35:4809–4819. [PubMed: 17621584]

Weinberg Z, Wang JX, Bogue J, Yang J, Corbino K, Moy RH, Breaker RR. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. Genome Biol. 2010; 11:R31. [PubMed: 20230605]

Zhang H, Li C, Yang F, Su J, Tan J, Zhang X, Wang C. Cation-pi interactions at non-redundant protein--RNA interfaces. Biochem. (Moscow). 2014; 79:643–652.

Highlights

- Crystal structure of the ligand-bound guanidine-I riboswitch aptamer
- Guanidinium recognized via H-bonding, ionic interactions, and cation-π interactions
- Unpredicted P3 helix contains binding site and reveals possible switching mechanism

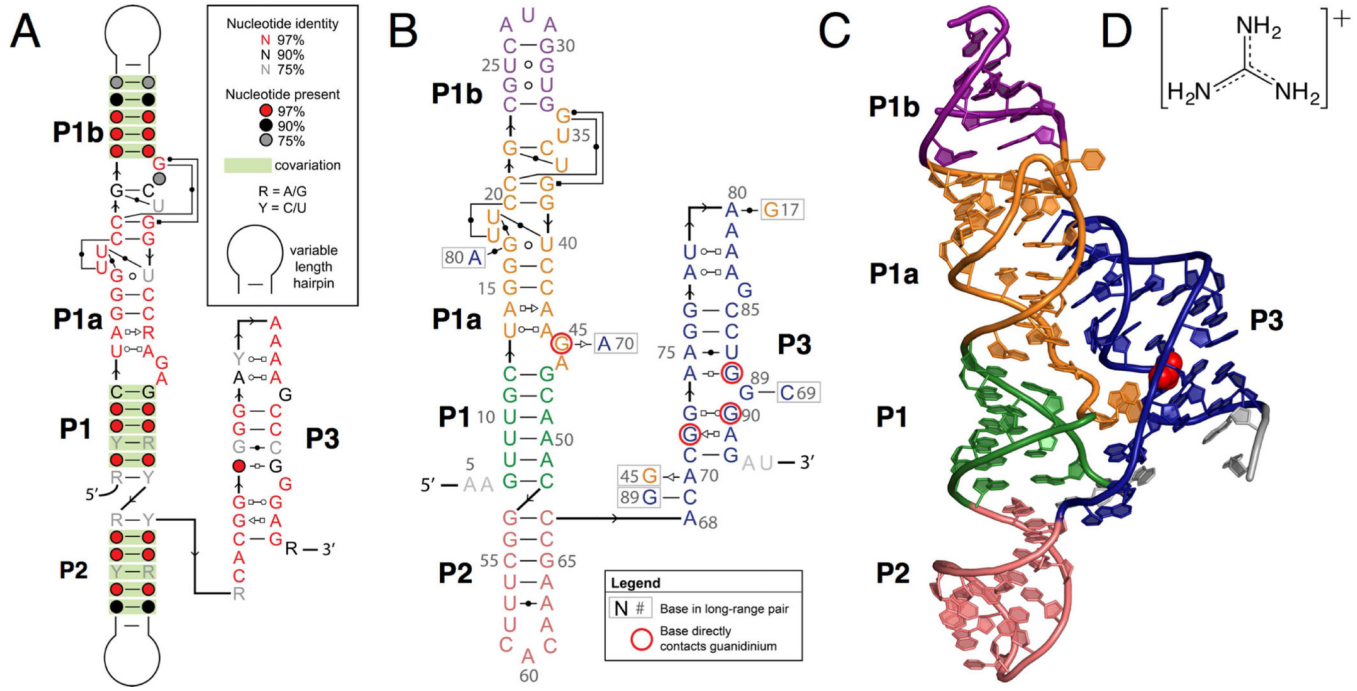


Figure 1. Consensus secondary structure and overall tertiary structure of the guanidine riboswitch bound to its ligand

(A) The consensus motif for the *ykkC* RNA, updated to reflect the secondary structure observed in the crystal structure. Base pairing notation is the same as previously published (Leontis and Westhof, 2001).

(B) The secondary structure of the crystal construct from *S. acidophilus*, including an engineered tandem GU wobble motif in the P1b helix for the purpose of iridium hexamine phasing. Helices P1, P1a, P1b, P2, and P3 are shown in green, orange, purple, salmon, and navy blue, respectively. Bases coming into direct contact with the guanidinium ligand are circled in red. Base pairing notation is the same as previously published (Leontis and Westhof, 2001).

(C) Ribbon diagram depicting the tertiary structural fold of the *S. acidophilus* guanidine riboswitch based on the crystal structure. The colors in the crystal structure diagram correspond to the colors in the secondary structure diagram in B. The guanidinium ligand is shown in red in the sphere representation.

(D) The chemical structure of the guanidinium ion ligand.

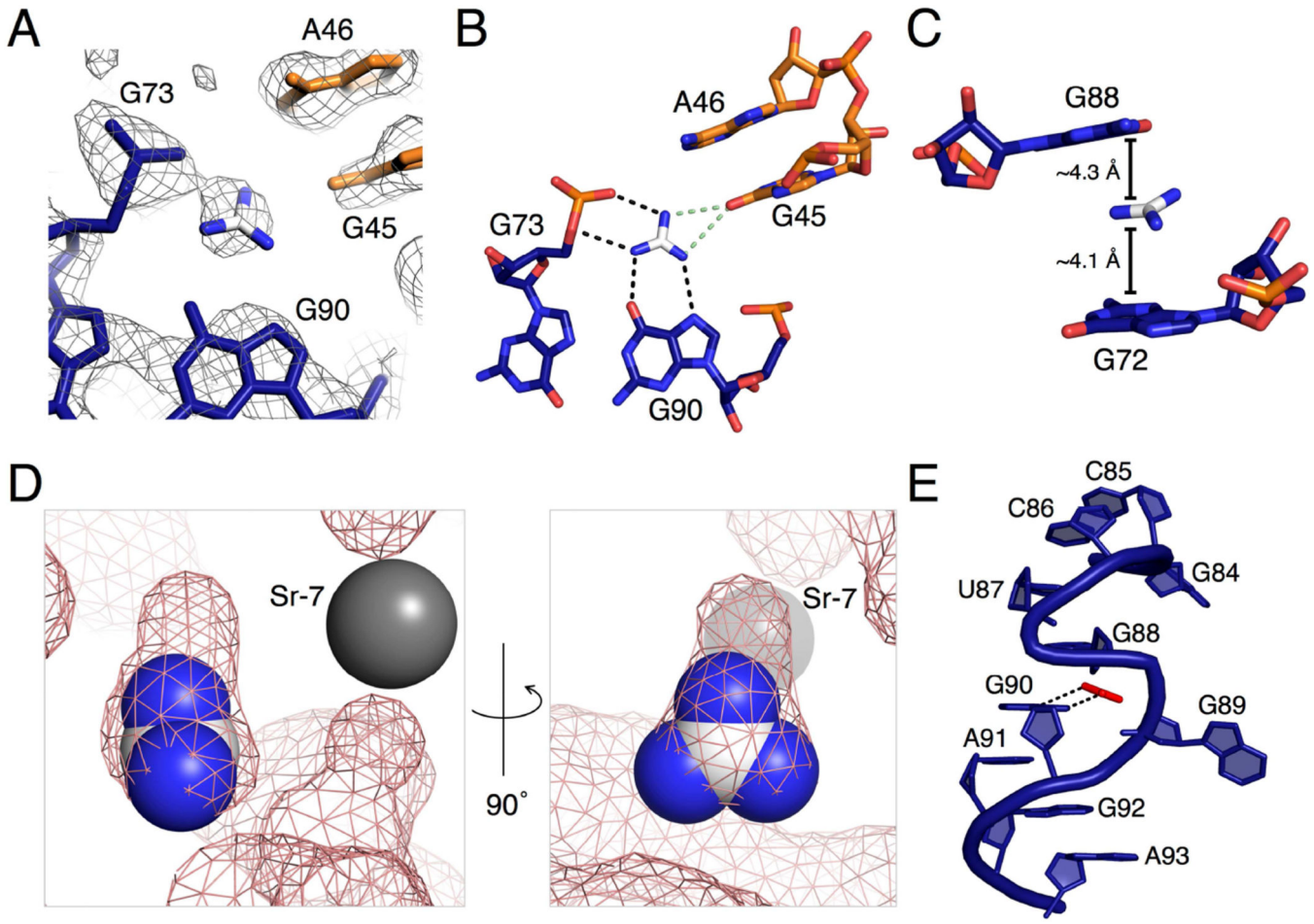


Figure 2. Binding pocket electron density and proposed contacts made by the guanidinium ligand to the riboswitch

(A) The unbiased electron density map contoured at 1σ, generated through SIRAS/MAD phasing and density modification. Nucleotides are colored using the same color scheme in 1B and C. For the guanidinium, carbon is in white and nitrogen is in blue.

(B) A top-down view of the binding pocket with black dashed lines representing predicted hydrogen bonds to the guanidine. The green dashed lines represent two mutually exclusive hydrogen bonds discussed in the text. The ligand and carbon atoms in the nucleotides are colored corresponding to the scheme in A. Oxygen atoms are represented in red and nitrogen in blue.

(C) A side view of the binding pocket showing two guanines sandwiching the guanidinium ligand. Distances from the guanidinium to the six-membered rings of G72 and G88 are shown. The color scheme used is the same as in B.

(D) Images of the binding pocket, the second rotated 90° from the first. Guanidinium is shown in the sphere representation using the same color scheme as in A. The grey sphere is a strontium ion in the crystal structure. The pink mesh is a surface view of the riboswitch, including metal ions and excluding the ligand.

(E) An S-turn involving two bases located in the binding pocket. Guanidinium is shown in red for clarity. The two black, dashed lines are the hydrogen bonds from guanidinium to G90, also represented in B.

See Figure S1 for additional hydrogen bonding contacts surrounding the binding pocket.

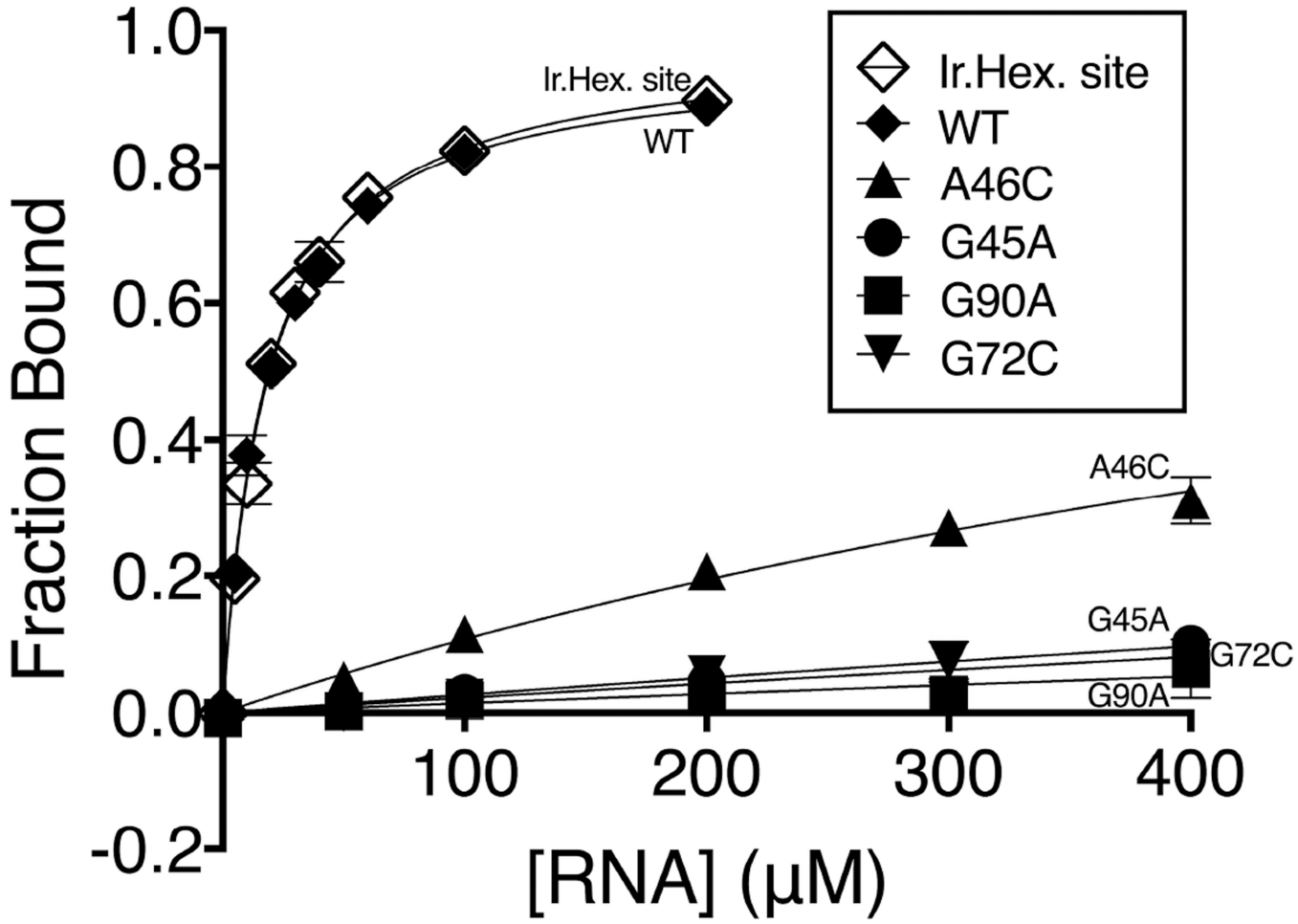


Figure 3. Binding affinity of [^{14}C]-guanidinium to the wild type riboswitch and point mutants by equilibrium dialysis

All point mutants are done in the background of the wild type RNA. WT, G45A, G90A, A46C, and G72C are represented by closed diamonds, circles, squares, triangles, and inverted triangles, respectively. The crystal construct, which contains an iridium hexamine binding site, is labeled “Ir.Hex. site” in the figure and represented by open diamonds. The data are fit to a hyperbolic curve. For the point mutants, Bmax is constrained to 1.0 (see experimental procedures section for more information). Fraction bound is shown with standard error bars with two replicates for each data point.

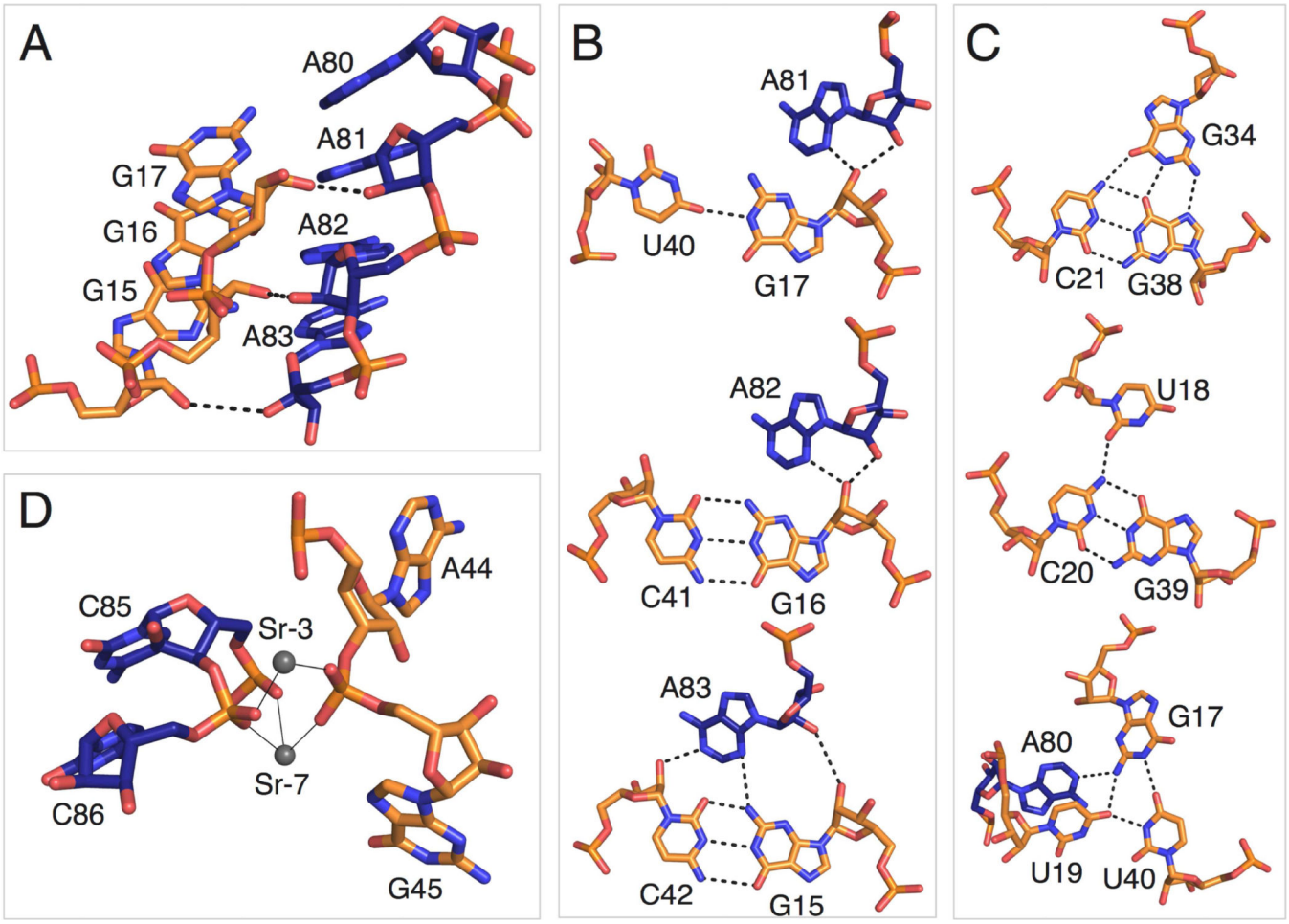


Figure 4. Long-range tertiary interactions involving highly conserved bases

(A) Side view of the highly conserved A patch (A80-A83) and adjacent bases (G15-G17) involved in A-minor interactions. The color scheme is the same as in 1B and C.

(B) Detailed view of A-minor interactions between helix P1a and the A patch in P3.

(C) Highly conserved base triples and a base quadruple in helix P1a.

(D) Strontium ions Sr-3 and Sr-7 in the crystal structure coordinating P1a and P3 phosphate oxygens. Thin, black lines indicate inner sphere coordination.

Table 1

Crystal statistics

				+ guanidine and $[Ir(NH_3)_6]$	Native + guanidine
PDB		**5T83**			
Data collection					
Beamline		8.2.2 at ALS		19-BM at APS	19-BM at APS
Space group		$P4_32_12$		$P4_32_12$	$P4_32_12$
Unit cell					
a, b, c (Å)		49.091, 49.091, 246.341		48.273, 48.273, 246.673	48.284, 48.284, 246.624
α, β, γ (°)		90, 90, 90		90, 90, 90	90, 90, 90
	High remote	*Inflection*	*Peak*		
Wavelength (Å)	0.999946	1.10532	1.10482	1.10062	1.10062
Resolution (Å)	40.00-2.70 (2.80-2.70)	40.00-2.80 (2.85-2.80)	40.00-2.80 (2.85-2.80)	40.00-2.92 (2.97-2.92)	40.00-2.93 (2.98-2.93)
R_{merge}	0.074 (0.712)	0.093 (0.729)	0.107 (0.784)	0.104 (>1.000)	0.165 (>1.000)
I/σI	35.2 (2.5)	29.1 (1.3)	30.1 (1.3)	23.0 (0.9)	18.3 (0.94)
$CC_{1/2}$ in highest resolution shell	0.928	0.831	0.832	0.447	0.448
CC* in highest resolution shell	0.981	0.953	0.953	0.786	0.787
Completeness (%)	99.8% (100.0%)	99.8% (100.0%)	99.9% (100.0%)	99.8% (99.8%)	99.6% (100.0%)
Redundancy	6.9 (6.7)	6.6 (4.2)	6.9 (4.3)	8.0 (6.3)	13.0 (11.9)
Total reflections	107588	92522	97493	96712	91329
Unique reflections	15638	14071	14090	12135	7022
Refinement					
Resolution (Å)	40.00-2.71				
No. reflections	8451				
R_{work}/R_{free}	0.20/0.23 (0.39/0.44)				
No. of atoms					
Total	1985				
RNA	1917				

	+ guanidine and $[Ir(NH_3)_6]$	Native + guanidine
PDB	**5T83**	
Ligand	4	
Cations	46	
Water	18	
B-factors		
Overall	88.9	
RNA	87.7	
Ligand	81.8	
Cations	129.0	
Water	116.1	
Root-mean-square deviations		
Bond lengths (Å)	0.007	
Bond angles (°)	1.815	