Structural basis for ligand binding to the guanidine-II riboswitch

CAROLINE W. REISS and SCOTT A. STROBEL

Department of Molecular Biophysics and Biochemistry, Chemical Biology Institute, Yale University, West Haven, Connecticut 06516, USA

ABSTRACT

The guanidine-II riboswitch, also known as *mini-ykkC*, is a conserved mRNA element with more than 800 examples in bacteria. It consists of two stem–loops capped by identical, conserved tetraloops that are separated by a linker region of variable length and sequence. Like the guanidine-I riboswitch, it controls the expression of guanidine carboxylases and SugE-like genes. The guanidine-II riboswitch specifically binds free guanidinium cations and functions as a translationally controlled on-switch. Here we report the structure of a P2 stem–loop from the *Pseudomonas aeruginosa* guanidine-II riboswitch aptamer bound to guanidine at 1.57 Å resolution. The hairpins dimerize via the conserved tetraloop, which also contains the binding pocket. Two guanidinium molecules bind near the dimerization interface, one in each tetraloop. The guanidinium cation is engaged in extensive hydrogen bonding to the RNA. Contacts include the Hoogsteen face of a guanine base and three nonbridging phosphate oxygens. Cation–π interactions and ionic interactions also stabilize ligand binding. The guanidine-II riboswitch utilizes the same recognition strategies as the guanidine-I riboswitch while adopting an entirely different and much smaller RNA fold.

Keywords: guanidine; riboswitch; RNA; kissing loop; tetraloop; hairpin

INTRODUCTION

The guanidine-II riboswitch, originally dubbed the *mini-ykkC* RNA, was first identified through bioinformatics in 2007 (Weinberg et al. 2007). The RNA was named *mini-ykkC* because it controls a similar set of genes as the *ykkC* RNA motif, but is much smaller and simpler. Both riboswitches regulate the expression of genes that encode guanidine carboxylases and SugE-like efflux proteins that transport guanidine out of the cell. The *ykkC* and *mini-ykkC* riboswitches are likely used by bacteria to detoxify the cell in the presence of high guanidine concentrations.

The Breaker laboratory discovered that both riboswitches bind guanidine as their native ligand, and they were renamed guanidine-I and guanidine-II (Nelson et al. 2017; Sherlock et al. 2017). The structure of the class I motif was reported previously (Battaglia et al. 2017; Reiss et al. 2017). The class II motif is made up of two short hairpins connected by a linker that is between 7 and 40 nt in length. Each hairpin is capped by a conserved ACGR tetraloop, and there are two conserved G–C base pairs at the base of each tetraloop, making a total of 16 conserved nucleotides in each riboswitch (eight in each hairpin). In contrast, the guanidine-I aptamer is substantially larger and contains 35 nt that are >97% conserved. Also, unlike the guanidine-I riboswitch, guanidine-II appears to bind two molecules of guanidine in a cooperative manner. Based on the consensus sequence, it was proposed that the two ACGR tetraloops each bind a molecule of guanidine and that the hairpins might interact with each other (Sherlock et al. 2017). Consistent with this, in-line probing data showed that guanidine-dependent modulation occurs in both tetraloops (Sherlock et al. 2017). These alterations to the RNA structure suggest that the guanidine-II riboswitch recognizes its ligand in an entirely different way from the guanidine-I riboswitch.

In addition to adopting distinct RNA folds in the aptamer domain, they also regulate gene expression by two different mechanisms. Based on bioinformatic data, the guanidine-II riboswitch is a translationally controlled on-switch, while guanidine-I is a transcriptionally controlled on-switch (Weinberg et al. 2007; Nelson et al. 2017; Sherlock et al. 2017).

Guanidine riboswitches face the challenge of recognizing a small molecule in a specific manner despite the presence of other guanidino-containing metabolites in the cell that are present at micromolar concentrations, such as arginine and

Corresponding author: scott.strobel@yale.edu

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.061804.117.

agmatine (Hamana 1996; Caldara et al. 2008; Bennett et al. 2009). In addition, guanidine riboswitches must select against urea, which occurs at high concentrations and is similar in size and shape, differing only in having a carbonyl group where guanidine has an amino group. The guanidine-I riboswitch accomplishes this by recognizing every possible face of the guanidinium cation (Battaglia et al. 2017; Reiss et al. 2017). In addition to hydrogen bonding, the class I riboswitch utilizes cation–π and ionic interactions.

The guanidine-II riboswitch recognizes two molecules of guanidine in a cooperative manner, suggesting that the tetraloops dimerize. To determine how the guanidine-II riboswitch recognizes guanidine and selects against similar compounds, and to determine the structural basis for tetraloop dimerization, we set out to determine the crystal structure of the riboswitch in complex with its ligand. Here, we present the crystal structure of a dimerized P2 stem–loop of the guanidine-II riboswitch from *Pseudomonas aeruginosa* determined at 1.57 Å resolution.

RESULTS AND DISCUSSION

Structure determination

Biochemical analysis of full-length guanidine-II riboswitch constructs suggested that the RNA forms large, higher order structures, likely due to dimerization of the hairpins in *trans*. Therefore, we used singlet hairpin constructs for crystallography. A previously characterized singlet stem–loop construct shows guanidine-dependent modulation by in-line probing, likely due to the formation of homodimers (Sherlock et al. 2017). Because the conserved tetraloops are identical in the P1 and P2 helix of most guanidine-II riboswitches, we expect that the conserved regions of the homodimer structure will adopt the same structure as a physiologically relevant heterodimer between P1 and P2. We targeted the 16-nt P2 stem of the guanidine-II riboswitch aptamer from *P. aeruginosa* (Fig. 1B), which controls a gene annotated as a SugE efflux protein. A member of the SugE family controlled by the guanidine-I riboswitch was previously shown to associate specifically with guanidine, suggesting that it may act to export guanidine out of the cell to avoid guanidine toxicity (Nelson et al. 2017). A nonconserved U in the sequence was modified to a 5-bromouridine and the structure was determined by SAD phasing. Guanidinium was clearly visible in the unbiased electron density maps. An F_o-F_c map contoured at 5σ was calculated using phases from a model in which guanidinium was omitted, which reveals significant guanidinium-shaped density that supports our assignment of the ligand model (Fig. 2A). X-ray statistics are shown in Table 1.

The crystal structure contains four P2 hairpins in the asymmetric unit (chains A, B, C, and D in the structure). The conserved nucleotides 5–12 are highly similar between the four molecules (average rmsd = 0.39 Å). Chains A + B and chains C + D dimerize through their tetraloops and this appears to be the physiologically relevant interface. The two dimers form crystal-packing interactions with each other.

Guanidinium recognition

Two stem–loops form a dimer via a kissing-loop interaction between the two ACGA tetraloops. The two hairpins are rotated 180° apart from each other and the interaction occurs strictly through the tetraloops and the cytosine in the first G–C base pair below the tetraloops. The tetraloop of each hairpin also creates a binding pocket for the guanidinium cation (Fig. 1C).

Although the guanidine-I and guanidine-II riboswitches adopt completely different RNA folds, they use remarkably similar strategies to recognize the ligand. Guanidinium has a pK_a of 13.6 (Perrin 1965), which means that under

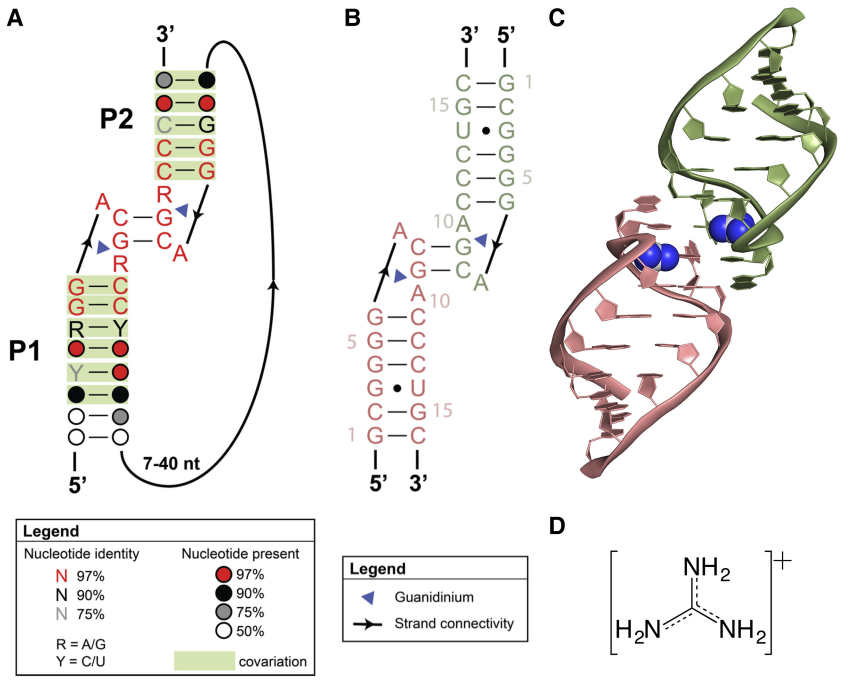

FIGURE 1. Overall structure of the guanidine-II riboswitch P2 hairpin from *P. aeruginosa*. (*A*) Updated consensus sequence to reflect secondary structure observed in the crystal structure. (*B*) Secondary structure of the crystal construct used. The pink and green hairpins represent chain A and chain B of the asymmetric unit and are equivalent in sequence. Blue triangles represent guanidinium. (*C*) Overall crystal structure of the guanidine-II riboswitch P2 hairpin dimer from *P. aeruginosa*. All figures are generated using chains A and B and all distances are calculated using chain A. (*D*) Chemical structure of the guanidinium cation.

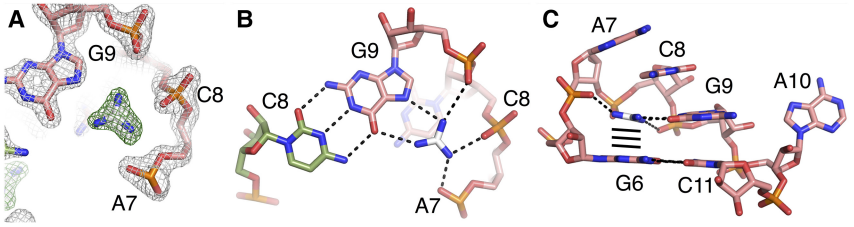

FIGURE 2. Guanidinium binds within the conserved ACGR tetraloop. (*A*) $2F_o-F_c$ map of binding pocket contoured at 2σ (gray) and F_o-F_c map of binding pocket contoured at 5σ (green), where the input model lacked guanidinium. Pink sticks represent chain A and white sticks represent guanidinium. (*B*) Hydrogen bonds donated from guanidinium to the riboswitch. Pink sticks represent chain A, green sticks represent chain B, and white sticks represent guanidinium. (*C*) The ACGA tetraloop and G–C base pair below. Hydrogen bonds from guanidinium to RNA are represented by gray dashes. Possible cation–π interactions are represented by black dashes. Color scheme is the same as in panel *A*.

physiological conditions it is present almost exclusively as a cation and has six potential hydrogen bond donors (Fig. 1D). In the guanidine-II riboswitch, five of the six positions make direct contact to the RNA.

One edge of guanidinium donates two hydrogen bonds to the Hoogsteen face of G9, a highly conserved nucleotide located in the tetraloop. One hydrogen bond is donated to the O6 carbonyl oxygen (3.1 Å) and one is donated to the N7 imino (3.1 Å) (Fig. 2B). On the other two edges of guanidinium, three hydrogen bonds are donated to nonbridging phosphate oxygens, specifically the pro-S_p of A7, pro-R_p of C8, and pro-R_p of G9 (Fig. 2B). It is likely that the negatively charged phosphate groups in the binding pocket also stabilize guanidinium cation binding through ionic effects. The ionic interactions likely contribute to selection of guanidinium over urea, which is uncharged. The last hydrogen in guanidinium is not recognized by the riboswitch and appears to be solvent exposed. In one set of the two dimers in the asymmetric unit (chains C and D), there is a solvent bridge comprised of two water molecules that hydrogen bond to the two guanidiniums and to each other. Ligand-dependent RNA structure modulation is observed with methylguanidine and aminoguanidine (Sherlock et al. 2017). However, the modulation is weaker than with guanidine. The solvent exposed region on each of the two binding pockets are oriented toward each other, suggesting that the riboswitch may be selecting against larger guanidine-containing molecules, like arginine and agmatine, by sterically preventing the two from binding simultaneously.

The guanidinium is also sandwiched above and below by conserved nucleotides G6 and C8 (Fig. 2C) and is potentially forming cation–π interactions, in which the positively charged ligand is interacting with the electron-rich π orbitals of RNA bases (Blanco et al. 2013; Dougherty 2013). Based on a published analysis of cation–π interactions found in 282 crystal structures containing nonredundant protein–RNA interfaces, interactions between the guanidine moiety of arginine and guanine bases are very common (Zhang et al. 2014). Cytosine is sometimes observed forming cation–π interactions with guanidino moieties in arginine, but much less commonly compared to guanine. Empirical energy calculations between cytosine and the guanidino moiety of arginine suggest that this interaction is unfavorable on average (Zhang et al. 2014). Nucleotide C8 may make little or no direct contribution to ligand-binding energetics, but is important for its role in inter-hairpin base-pairing at the dimerization interface. These cation–π interactions are another way that the riboswitch can select for the guanidinium cation over urea.

Structural aspects of the dimerization interface

The main structural motif that forms the dimerization interface is a kissing-loop interaction between the two tetraloops. C8 and G9 form Watson–Crick base-pairing with the equivalent C8 and G9 of the dimerization partner (Fig. 3A). The G9 nucleotide also forms bifurcated hydrogen bonds to guanidinium via the Hoogsteen face (Fig. 2B). This kissing-loop interaction positions the two guanidiniums within 10 Å of each other. In the tetraloop, there are two adenines, one on each side of the kissing-loop nucleotides C8 and G9 (Fig. 2C). Highly conserved nucleotide A7 stacks on top of C8 and forms two hydrogen bonds to the sugar edge of C11 of the dimerization partner (Fig. 3B). Nucleotide A10, which is highly conserved as a purine (A or G), flips out from the tetraloop and forms stacking interactions with A10 of the other hairpin (Fig. 3C). Guanidinium binding may result in a rearrangement of the tetraloop such that A7 and A10 are more accessible to form dimerization interactions. In this case, guanidinium binding would shift the equilibrium toward the dimeric state. Consistent with this model, in-line probing of a P1 hairpin from the *Gloeobacter violaceus* guanidine-II riboswitch shows that the equivalents of nucleotides A7 and A10 in the crystal structure modulate in a guanidine-dependent manner (Sherlock et al. 2017).

Guanidinium recognition in class-I and class-II guanidine riboswitches

The guanidine-I and guanidine-II riboswitches adopt completely different RNA folds, but use similar strategies to recognize guanidinium (Supplemental Fig. S1A). In the guanidine-I riboswitch, one edge of guanidinium forms two hydrogen bonds with the Hoogsteen face of G90 in a bifurcated manner. The same hydrogen bonds are formed in the guanidine-II riboswitch, involving G9 (Supplemental Fig. S1B). Both riboswitches also use phosphate oxygens in the RNA backbone to recognize guanidinium. In the guanidine-I riboswitch, the guanidinium donates two hydrogen bonds

TABLE 1. Crystal statistics

	PaeP2 hairpin with 5-bromoU SAD	PaeP2 hairpin native
PDB	5VJB	5VJ9
Data collection		
Beamline	24-ID-C at APS	24-ID-C at APS
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell		
a, *b*, *c* (Å)	50.2, 60.5, 71.9	50.3, 60.7, 72.3
α, β, γ (°)	90, 90, 90	90, 90, 90
Wavelength (Å)	0.91930	0.91930
Resolution (Å)	40.00–2.10 (2.14–2.10)	40.00-1.57 (1.60–1.57)
R_{merge}	0.068 (1.058)	0.052 (1.052)
$I/\sigma I$	31.3 (1.6)	46.3 (1.7)
$CC_{1/2}$ in highest resolution shell	0.552	0.565
CC* in highest resolution shell	0.843	0.850
Completeness (%)	99.9% (100.0%)	99.8% (98.8%)
Redundancy	6.8 (7.0)	11.6 (6.9)
Total reflections	172716	361260
Unique reflections	25445	31438
Refinement		
Resolution (Å)	40.00–2.10	40.00–1.57
No. of reflections	13040	29787
R_{work}/R_{free}	0.22/0.24 (0.32/0.39)	0.19/0.23 (0.40/0.42)
No. of atoms		
Total	1456	1727
RNA	1372	1368
Ligand	16	16
Cations	57	62
Water	11	281
B-factors		
Overall (Å^2)	42.1	32.4
RNA (Å^2)	41.2	29.4
Ligand (Å^2)	37.3	26.1
Cations (Å^2)	66.4	49.8
Water (Å^2)	36.1	43.1
Root-mean-square deviations		
Bond lengths (Å)	0.012	0.013
Bond angles (°)	1.995	1.841

to phosphate oxygens of G73. In the guanidine-II riboswitch, guanidinium forms hydrogen bonds to three nonbridging phosphate oxygens. In both classes of riboswitch, the proximity of the negatively charged phosphate groups stabilizes the binding of the guanidinium cation through ionic interactions (Supplemental Fig. 1B). Both riboswitches use guanine bases to form cation–π interactions with guanidinium.

The strategies used to recognize guanidinium by the two classes of riboswitch are the same, but the RNA context differs. In the class I guanidine riboswitch, the single binding site is created at the interface between two helical elements, P1a and P3. The class II guanidine riboswitch has two binding sites, one in each of the hairpins that dimerize upon guanidine binding. Each guanidinium ligand makes contacts to only a single hairpin, but the binding sites are located at the dimerization interface.

Proposed switching model in the full-length aptamer

In a wild-type guanidine-II aptamer, the P1 and P2 hairpins are separated by a linker that is 7 to 40 nt in length (Fig. 1A). In the crystal structure, the 3′-nucleotide of one hairpin and the 5′-nucleotide of the second hairpin are ∼50 Å apart. This is the distance that must be spanned by the random linker between the two hairpins. Nucleotides in a splayed conformation extend a distance of ∼7.2 Å per nucleotide, measuring from phosphate to phosphate. This provides enough length to cover the 50 Å distance, even for the shortest of linker regions.

The guanidine-II riboswitch is a translational on-switch. The structure of the two hairpins provides a model for how guanidine could regulate expression of the downstream message. In the case of this *P. aeruginosa* riboswitch, the Shine–Dalgarno sequence is sequestered by base-pairing with part of the linker region and nucleotides at the 5′ end of the P1 helix (Fig. 4). In the guanidine-bound structure, head-to-head dimerization of P1 and P2 enforces the spatial separation of the Shine–Dalgarno sequence from the anti-Shine–Dalgarno sequence. This would expose the Shine–Dalgarno sequence and promote translation in the presence of guanidine (Fig. 4). The crystal structure does not allow us to distinguish between a conformational selection mechanism or an induced fit mechanism for switching. Exposure of the Shine–

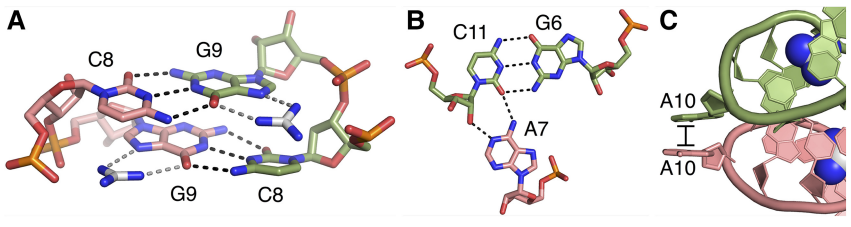

FIGURE 3. The two hairpins dimerize head-to-head through a kissing-loop interaction. (*A*) The kissing-loop interaction at the dimerization interface. Color scheme is the same as in Figure 2B. (*B*) Hydrogen bonds formed at the dimer interface between the Watson–Crick face of A7 in chain A and the sugar edge of C11 in chain B. (*C*) Stacking interactions formed by A10 in chain A and A10 in chain B.

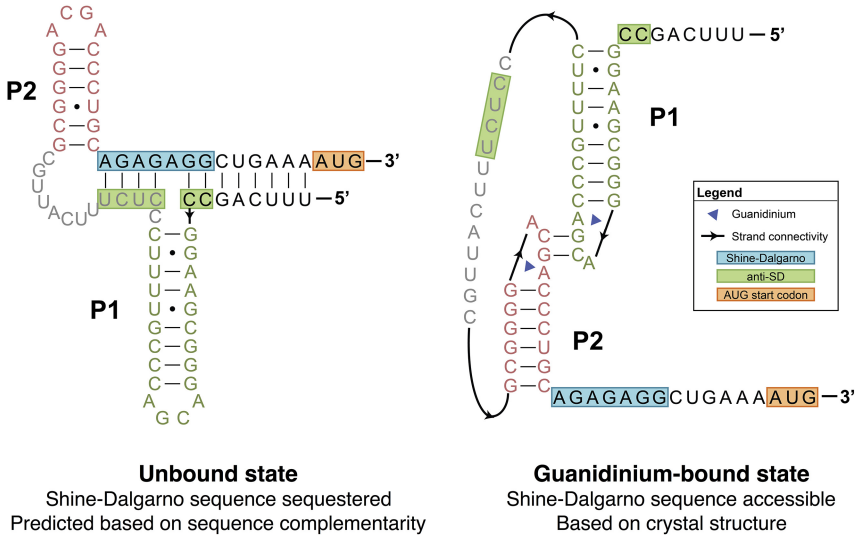

FIGURE 4. Predicted switching model for the *P. aeruginosa* guanidine-II riboswitch. Upon guanidinium binding, the Shine–Dalgarno sequence is made accessible by spatial separation from the anti-Shine–Dalgarno (anti-SD) sequence. Green nucleotides represent hairpin P1 and pink represents hairpin P2.

Dalgarno sequence of a bacterial mRNA allows the ribosome to be recruited for translation of the downstream genes that control guanidine toxicity in the presence of ligand.

Bioinformatic analysis has revealed that operons controlled transcriptionally by class I guanidine riboswitches often contain a class II riboswitch that controls translation of the 125 kDa carboxylase gene (Sherlock et al. 2017). The combination of these two riboswitches allows the cell to exercise guanidine-dependent gene control at both the transcriptional and translational level. The large difference in both RNA fold and the mechanism of gene control enables tight regulation of genes meant to relieve guanidinium toxicity.

MATERIALS AND METHODS

RNA preparation

Unmodified RNA was purchased from Sigma-Aldrich. RNA containing a single 5-bromouridine at position 14 was purchased from Dharmacon. All synthesized RNAs were purified by denaturing polyacrylamide gel electrophoresis.

Crystallization

Crystallization was performed using the microbatch-under-oil method with a 2:1 paraffin:silicon oil overlay. 200 µM *P. aeruginosa* Pae P2 hairpin RNA in crystallization buffer (10 mM $MgCl_2$, 10 mM KCl, 10 mM HEPES-KOH, pH 7.5, and 40 mM guanidine) was mixed in a 1:1 ratio with crystallization reagent 1 for the 5-bromoU-modified RNA (45% MPD, 50 mM MES, pH 5.6, 4 mM NaCl, 40 mM KCl, and 12 mM spermine) or crystallization reagent 2 for the unmodified RNA (40% MPD, 50 mM sodium acetate, pH 5.0, 12 mM NaCl, 80 mM KCl, and 12 mM spermine). Crystals appeared overnight at 23°C. Crystals were flash frozen in liquid nitrogen without further preparation, since 40%–45% MPD is a cryoprotectant.

Structure determination

Unbiased phases were obtained by single wavelength anomalous diffraction (SAD) phasing. SAD data were collected at the peak wavelength for the bromine K edge at beamline 24-ID-C at the Advanced Photon Source (APS). Data were indexed, integrated, and scaled using HKL2000. The SHELXC/D/E suite was used to evaluate anomalous signals, locate heavy atoms, and obtain the unbiased phases. Four bromine sites were identified, indicating four molecules in the asymmetric unit and allowing us to orient molecules within the density. Model building was performed in Coot. Refmac5 and phenix.refine were used for refinement. R_{free} test sets were matched between the derivative and native data sets to prevent bias. Figures of the crystal structure were made using open source PyMol.

DATA DEPOSITION

Coordinates for the guanidine-II riboswitch have been deposited in the protein data bank (PDB) under accession codes 5VJ9 (native) and 5VJB (5-BrU). There are four molecules (two dimers) in the asymmetric unit.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank the synchrotron beamline staff at the Northeastern Collaborative Access Team (NE-CAT) at the Advanced Photon Source for their assistance; J. Wang, who provided important help with data processing; M. Strickler from the Yale Center for Structural Biology; D. Hiller and other members of the Strobel laboratory for valuable advice and discussion; and R. Breaker and M. Sherlock for helpful discussion. C.W.R. was supported by the National Institutes of Health Cellular and Molecular Biology Training Grant (T32GM007223). This work was also supported by a National Institutes of Health grant to S.A.S. (GM022778).

Received April 19, 2017; accepted May 22, 2017.

REFERENCES

Battaglia RA, Price IR, Ke A. 2017. Structural basis for guanidine sensing by the *ykkC* family of riboswitches. *RNA* **23:** 578–585.

Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. 2009. Absolute metabolite concentrations and

implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* **5:** 593–599.

Blanco F, Kelly B, Sánchez-Sanz G, Trujillo C, Alkorta I, Elguero J, Rozas I. 2013. Non-covalent interactions: complexes of guanidinium with DNA and RNA nucleobases. *J Phys Chem B* **117:** 11608–11616.

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

Dougherty DA. 2013. The cation-π interaction. *Acc Chem Res* **46:** 885–893.

Hamana K. 1996. Distribution of diaminopropane and acetylspermidine in *Enterobacteriaceae*. *Can J Microbiol* **42:** 107–114.

Nelson JW, Atilho RM, Sherlock ME, Stockbridge RB, Breaker RR. 2017. Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. *Mol Cell* **65:** 220–230.

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

Reiss CW, Xiong Y, Strobel SA. 2017. Structural basis for ligand binding to the guanidine-I riboswitch. *Structure* **25:** 195–202.

Sherlock ME, Malkowski SN, Breaker RR. 2017. Biochemical validation of a second guanidine riboswitch class in bacteria. *Biochemistry* **56:** 352–358.

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

Zhang H, Li C, Yang F, Su J, Tan J, Zhang X, Wang C. 2014. Cation-π interactions at non-redundant protein-RNA interfaces. *Biochemistry (Mosc)* **79:** 643–652.

Structural basis for ligand binding to the guanidine-II riboswitch

Caroline W. Reiss and Scott A. Strobel

RNA 2017 23: 1338-1343 originally published online June 9, 2017
Access the most recent version at doi:10.1261/rna.061804.117