

Biochemical Validation of a Fourth Guanidine Riboswitch Class in Bacteria

Hubert Salvail,^{†,‡} Aparajita Balaji,^{†,‡} Diane Yu,[†] Adam Roth,[‡] and Ronald R. Breaker^{†,‡,§,*}

[†]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8103, USA

[‡]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520-8103, USA

[§]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8103, USA

Corresponding Author

*E-mail: ronald.breaker@yale.edu

[‡]These authors contributed equally to this work

ABSTRACT: An intriguing consequence of ongoing riboswitch discovery efforts is the occasional identification of metabolic or toxicity response pathways for unusual ligands. Recently, we reported the experimental validation of three distinct bacterial riboswitch classes that regulate gene expression in response to the selective binding of a guanidinium ion. These riboswitch classes, called guanidine-I, -II and -III, regulate numerous genes whose protein products include previously misannotated guanidine exporters and enzymes that degrade guanidine via an initial carboxylation reaction. Guanidine is now recognized as the primal substrate of many multidrug efflux pumps that are important for bacterial resistance to certain antibiotics. Guanidine carboxylase enzymes had long been annotated as urea carboxylase enzymes but are now understood to participate in guanidine degradation. Herein we report the existence of a fourth riboswitch class for this ligand, called “guanidine-IV”. Members of this class use a novel aptamer to selectively bind guanidine and use an unusual expression platform arrangement that is predicted to activate gene expression when ligand is present. The wide distribution of this abundant riboswitch class, coupled with the striking diversity of other guanidine-sensing RNAs, demonstrates that many bacterial species maintain sophisticated sensory and genetic mechanisms to avoid guanidine toxicity. This finding further highlights the mystery regarding the natural source of this nitrogen-rich chemical moiety.

Numerous riboswitch classes¹⁻³ for RNA-derived compounds, such as enzyme cofactors and ribonucleotide derivatives,^{4,5} have been discovered in recent years. These findings are consistent with the RNA World theory,^{6,7} which encompasses the concept that ancient metabolic processes and their regulation would have required many types of structured noncoding RNAs (ncRNAs) to perform various functions. As efforts to experimentally validate novel riboswitch classes have progressed, some observations revealed that many bacterial species also sense ligands whose biological relevance had previously been obscure. For example, the discovery of fluoride riboswitches^{8,9} has exposed the existence of numerous genes whose expression allows cells to overcome the stresses caused by toxic levels of this elemental anion.⁹⁻¹¹ In other instances, RNAs corresponding to riboswitches for the bacterial second messengers c-di-AMP^{12,13} and c-AMP-GMP^{14,15} had been identified before their signaling molecules were known to exist.

The experimental validation of the function of guanidine riboswitches^{16,17} provides another prominent illustration wherein the investigation of a riboswitch candidate uncovers hidden

biological processes. Three distinct structural classes of riboswitches have been found that respond to guanidine, called guanidine-I,¹⁸ guanidine-II,¹⁹ and guanidine-III.²⁰ These RNAs control a diverse set of genes whose protein products act to protect cells from toxic levels of guanidinium ions. For example, a prominent set of genes associated with guanidine riboswitches codes for proteins implicated in guanidine degradation by first promoting guanidine carboxylation.^{18,21} Another commonly associated set of genes codes for proteins recently demonstrated to expel guanidine from cells.^{18,22} Strangely, it is not yet clear how cells become exposed to toxic levels of guanidine. Presumably they experience a metabolic state wherein guanidine accumulates as a byproduct of other metabolic processes.^{16,18}

In our ongoing effort to find additional riboswitch classes and other structured ncRNAs, we identified nucleotide sequences corresponding to a widespread and well conserved sequence element called the *mepA* motif (**Figure 1A**). The name for this candidate ncRNA was chosen due to its frequent association with *mepA* genes, which commonly reside immediately downstream from the conserved features that define the motif. The *mepA* motif was initially uncovered by newly implementing a bioinformatics approach called ‘GC-IGR analysis’ method.^{23,24} This computational pipeline was used to comprehensively examine the genome of *Clostridium botulinum* Strain A Hall, which produced the initial representatives of the *mepA* motif. Ultimately, we detected 1275 representatives with non-redundant sequences distributed across six phyla: Actinobacteria, Firmicutes, Fusobacteria, Spirochaetes, Synergistetes and Tenericutes.

The diversity of sequences and broad phylogenetic distribution indicates that *mepA* motif RNAs serve an important biochemical role. The conserved features of this class (**Figure 1**) include two short base-paired substructures (P1 and P2) in addition to a typical intrinsic terminator stem.^{25,26} The most highly conserved nucleotide positions cluster both near the loop of P1 and the loop of the terminator stem, which would become proximal on the formation of the predicted P2 stem. A variable-length linker joins the right shoulder of P1 with the left shoulder of the terminator stem, which appears to be too short to permit the loops of these two hairpins as described above on the simultaneous formation of P1 and the full-size terminator stem. This architecture suggests that *mepA* motif RNAs might function as riboswitches wherein ligand binding to an aptamer domain modulates gene expression by controlling transcription via an expression platform employing a terminator stem. Specifically, we speculated that *mepA* motif RNAs are most likely

to function as genetic “ON” switches wherein ligand binding stabilizes the formation of P2 to preclude formation of the full-length terminator stem.

Representatives of the *mepA* motif most commonly associate with genes (**Figure 1B**, **Figure S1**) that are known to be regulated by previously validated classes of guanidine riboswitches.¹⁸⁻²⁰ As noted above, the most common gene association codes for MepA/MATE-like proteins, which are annotated as multidrug efflux proteins. Previous findings^{18,22} demonstrated that the biological function of certain efflux proteins associated with guanidine riboswitches is to eject guanidine from cells, and that variants of these proteins likely emerged later in evolution to function as antibiotics resistance factors.²² Additional genes that frequently occur near the *mepA* motif RNA include EmrE-like and B3/B4 domain proteins that are also commonly regulated by representatives of the known guanidine riboswitch classes.¹⁸⁻²⁰ Therefore, we sought to evaluate the hypothesis that *mepA* motif RNAs function as members of a distinct riboswitch class for guanidine. In addition, we evaluated a series of guanidine analogs for ligand function, including those carrying modest chemical modifications (**Figure 1C**) that reduce affinity with representative aptamers from the known riboswitch classes for guanidine.

Two constructs based on the *mepA* representative from *C. botulinum* B1 strain Okra were prepared for analysis (**Figure 2A**). The 97 *mepA* construct carried 97 nucleotides from the natural sequence plus two additional guanosine nucleotides on the 5' terminus to facilitate production by in vitro transcription. A shorter construct called 81 *mepA* was also prepared to weaken the terminator stem and increase the length of the single-stranded linker. This latter construct takes into consideration the hypothesis that the aptamer domain and the full-length terminator stem might form in a mutually exclusive manner, as would be expected if the RNA functions as a genetic “ON” riboswitch.

Indeed, we failed to observe any evidence that the 97 *mepA* construct binds guanidine by using in-line probing assays^{27,28} (see below). In contrast, the 81 *mepA* construct exhibits robust structural modulation when 1 mM guanidine is added to binding assay reactions (**Figure 2B**). These assays take advantage of the fact that 3',5'-phosphoester bonds in RNA can undergo spontaneous scission by the nucleophilic attack of a ribose 2'-oxygen atom on the adjacent phosphorus center. This reaction is dependent on the local structural context of an RNA molecule, wherein linkages in unstructured regions cleave more rapidly than those in highly structured regions.²⁷ Thus, changes

in the folded structure of an RNA on ligand binding cause changes in the pattern of RNA degradation products as observed by polyacrylamide gel electrophoresis (PAGE).

Two regions of 81 *mepA* observed to undergo suppression of spontaneous RNA strand scission in the presence of 1 mM guanidine (**Figure 2B**) are located in the loop of P1 (site 1) and the loop of the terminator stem (site 3) (**Figure 1A**). Nucleotides in or near sites 1 and 3 comprise the predicted P2 stem, suggesting that this *mepA* motif substructure is formed on ligand binding and therefore might be critical for aptamer function. In contrast, the nucleotides in the linker between the P1 and terminator stems (site 2) appear to be relatively unstructured in the absence of ligand, and these nucleotide linkages experience an increase in spontaneous scission when guanidine is present. This observation is consistent with our hypothesis that the ligand-binding aptamer conformation precludes formation of the full-length terminator stem. Finally, we note that nucleotide linkages preceding P1 and several linkages in the loop of the terminator stem that join poorly conserved nucleotides do not experience a change in spontaneous scission on the addition of guanidine, suggesting these nucleotides are not directly involved in forming the aptamer structure. Similar results were observed with an analogous construct based on the *mepA* motif representative from the bacterium *Lactobacillus brevis* (data not shown).

To establish the basic characteristics of the guanidine aptamer in the 81 *mepA* construct, we conducted in-line probing reactions using various concentrations of guanidine ranging from 214 nM to 10 mM. Quantitation of the intensities of certain bands from the resulting PAGE analysis of products (**Figure 2C**) allowed us to estimate the fraction of RNA bound to guanidine at each concentration tested (**Figure 2D**). The resulting plot indicates that the RNA binds guanidine in a 1-to-1 complex with an apparent dissociation constant (K_D) of approximately 64 μ M.

Similarly, binding assays were conducted with numerous guanidine analogs, which revealed that the 81 *mepA* RNA discriminates strongly against even close chemical derivatives. For example, at 1 and/or 10 mM concentration the methyl, amino, and ethyl derivatives of guanidine (**Figure 1C**) modulate the structure of the 81 *mepA* RNA in a manner essentially identical to that observed for guanidine as determined by in-line probing assays used to screen for ligand function (**Figure 3A**). However, additional assays using a range of ligand concentrations provided estimates of K_D values for these guanidine analogs (**Figure S2**), demonstrating that guanidine is bound by the aptamer 10- to 100-fold more tightly (**Figure 3B**).

Additional analogs were also examined for evidence of binding (see the legend to **Figure 3B**), but all are rejected even when tested at concentrations of 10 mM (data not shown). The strong selectivity and affinity of the 81 *mepA* RNA construct for guanidine is very similar to the binding characteristics observed for members of the three previously studied guanidine riboswitch classes.¹⁸⁻²⁰ The characteristics of these aptamers thus could be harnessed by cells to form precision riboswitches to regulate gene expression in response to elevated guanidine concentrations, despite the presence of numerous other metabolites naturally present in cells that carry a guanidine moiety.

Mutant versions of the 81 *mepA* RNA construct were also prepared to evaluate the importance of the highly conserved nucleotides in the loop of P1 or in the loop of the terminator hairpin. Specifically, mutant constructs M1 (G26U) and M2 (A63U) each carry a single nucleotide change (**Figure 2A**) at a position whose identity remains conserved in more than 97% of the known representatives (**Figure 1A**). As expected, these mutations independently disrupt guanidine binding, even when the ligand is present at 1 mM (**Figure 3C**). These results suggest that the conserved nucleotides observed by comparative sequence analysis are either directly forming the binding pocket for guanidine, or perhaps are required for *mepA* motif RNAs to adopt a structure that permits the pocket to fold properly.

The architecture of *mepA* motif RNAs is unusual compared to most known riboswitch classes. Most commonly, a contiguous aptamer domain is employed by a riboswitch, wherein this structure overlaps some portion of the expression platform to allow ligand binding to control production of either the full-length mRNA⁴ or the protein it encodes.²⁹ However, *mepA* motif RNAs appear to carry a split aptamer, wherein the 3' portion resides in the loop of the terminator stem. This creates a possible problem for aptamer formation because the two portions of the aptamer need to make contact to form the guanidine binding pocket. Again, this structural requirement is consistent with nucleotide covariation supporting the formation of a P2 stem (**Figure 1A**) and the guanidine-induced structural modulation evident from in-line probing assay data (**Figures 2, 3 and 4**).

This unusual architecture is predicted to create distance constraints that are coupled to each other in a manner dictated by the progression of RNA polymerase as it synthesizes the nascent mRNA. Specifically, for the parent *mepA* motif construct used in this study, the linker joining P1 and the left shoulder of the terminator stem could be entirely single-stranded until nucleotides corresponding to the right shoulder of the terminator stem begin to emerge from the exit channel of RNA polymerase. As the terminator stem begins to form base-pairs, the length of the single-

stranded linker between P1 and the terminator becomes progressively shorter. Thus, we speculated that perhaps the length of the nascent terminator stem would dictate the ability of the two portions of the aptamer to make contact and form a guanidine binding site.

To examine this hypothesis, we prepared a series of 3'-end deletion constructs (**Figure 4A**) based on the original 97 *mepA* RNA that encompasses the entire motif, including the run of U nucleotides that are characteristic of an intrinsic terminator device. As predicted, shorter constructs such as 84 *mepA* and 87 *mepA* exhibit guanidine-dependent structural modulation with patterns (**Figure 4B**) that are like those observed for the 81 *mepA* construct (**Figure 2**). However, longer constructs such as 90 *mepA*, 93 *mepA* and 97 *mepA* fail to respond to 1 mM guanidine. These findings indicate that as the length of the base-paired region of the terminator stem increases beyond a certain point, the split aptamer domains cannot merge to form the ligand binding pocket.

This data is consistent with a model for gene control (**Figure 4C**) wherein the ligand must bind to the nascent RNA before the terminator stem becomes too long to permit contact between the two parts of the aptamer. From data gathered using the *mepA* motif representative examined in this study (**Figure 4B**), we conclude that the single-stranded linker is too short (L_1 , **Figure 4C**) when the terminator includes nucleotide position 90. In contrast, the single-stranded linker is sufficiently long (L_2) when the terminator stem included only up to nucleotide position 87. Intriguingly, bacterial RNA polymerases are known to sequester approximately 12 nucleotides of the nascent RNA transcript in the exit channel.³⁰⁻³² Therefore, when RNAP is experiencing a transcriptional pause induced by the run of U nucleotides at the 3'-end of the terminator structure, a portion of the right shoulder of the terminator stem remains sequestered in the exit channel.

Even if RNAP is paused near the end of this U-rich region, for example at nucleotide position 97, the nascent RNA transcript protruding from the protein complex only reaches approximately nucleotide position 86. Therefore, this paused state should still permit guanidine binding. Once guanidine binds, extending the base-pairing of the terminator stem should be disfavored because their complementary nucleotides in the single-stranded linker are structurally unavailable for interaction. This ligand-bound state thereby removes the thermodynamic incentive for base-pairs between the nascent RNA and its DNA template to be displaced in favor of the formation of additional base-pairs in the terminator stem, which is the widely accepted mechanism for this type of transcription termination.^{25,26}

Similarly, we prepared a series of constructs based on the 81 *mepA* RNA with internal deletions within the putative single-stranded linker (**Figure 5A**) to further assess the effects of linker length on guanidine binding. In-line probing analyses reveal that the deletion of 6 nucleotides (construct $\Delta 46-51$) results in a functional aptamer, whereas the deletion of three more nucleotides (construct $\Delta 43-51$) prevents guanidine binding (**Figure 5B**). This additional data also sets constraints on the length of single-stranded nucleotides in the linker needed to permit ligand binding (**Figure 5C**). These findings combined with the length constraint noted earlier (**Figure 4**) indicate that the linker must have no less than 7 or 8 single-stranded nucleotides for the split aptamer portions to make contact and form a ligand-binding pocket.

Thus, given the bioinformatic, biochemical, and mechanistic characteristics of *mepA* motif RNAs, we conclude that they likely comprise as a fourth riboswitch class for guanidine. Thus, we propose naming this riboswitch class ‘guanidine-IV’. Our bioinformatics searches have uncovered 1275 non-redundant representatives of this newly found riboswitch class, which remarkably makes it more abundant in comparison to the numbers¹ published for guanidine-I,¹⁸ guanidine-II,¹⁹ and guanidine-III²⁰ riboswitches. Atomic-resolution structural models have been established for these previously reported riboswitch classes, demonstrating that RNAs can use diverse structures to selectively recognize guanidine.³³⁻³⁷ The biophysical analysis of guanidine-IV riboswitches promises to provide yet another glimpse at how RNAs can employ only four types of nucleotides to form precise binding pockets for this ligand. Taken together, this collection of riboswitch classes also provides further evidence that guanidine toxicity is widely experienced by many bacterial species.^{16,17,38} Interestingly, however, a biological source of the toxic levels of guanidine in so many species remains to be determined. Perhaps with the addition of the genes known to associate with guanidine-IV riboswitches, the expanded ‘super-regulon’ of guanidine-controlled pathways will provide some hints at the natural source(s) of this molecule. Regardless, advances are already being made on biological processes controlled by guanidine riboswitches,^{21,39,40} and guanidine aptamers appear to be promising RNA structures for use in synthetic biology.⁴¹

It is also interesting to consider why the guanidine-IV riboswitch class was not encountered in previous bioinformatics searches for novel ncRNA motifs.^{8,12,42} The consensus model exhibits extensive base-pairing and at least a modest number of highly-conserved nucleotides, which are features that typically designate strong ncRNA candidates during searches exploiting comparative sequence analysis algorithms. However, the consensus model is dominated by the terminator stem,

which is usually ignored when ranking ncRNA candidates because of its prevalence as a known genetic feature. In other words, the presence of a terminator stem immediately upstream of a coding region is very common because this structure is typically the transcription stop site for the preceding transcriptional unit in many bacterial genomes.

When we first encountered representatives of the *mepA* motif, they appeared to be unusual to us because the loop sequences of the terminator stem were remarkably well-conserved, which is unnecessary for the structure to serve as a simple intrinsic transcription terminator.^{25,26} Indeed, the architecture of the *mepA* motif was reminiscent of a recently reported riboswitch class for the thiamin pyrophosphate biosynthetic precursor HMP-PP, which appears to exploit highly conserved nucleotides in a terminator stem to selectively bind its ligand.^{24,43} Likewise, we predict that nucleotides in the terminator loop are involved in forming the aptamer structure required to bind guanidine.

Given that riboswitches such as the HMP-PP class and the guanidine-IV class are more difficult to identify by previous comparative sequence analysis algorithms, there might be many other classes hidden in existing bacterial genomic databases that exploit the loop of terminator stems to form an aptamer structure. The unique structural constraints placed on riboswitches that embed all or a portion of their aptamer in the loop of an intrinsic transcriptional terminator hairpin are also likely to place functional constraints on the riboswitch. Thus, we predict that most RNAs with this architecture will function as genetic “ON” switches wherein the ligand binding structure and the terminator stem will be mutually exclusive structures.

MATERIALS AND METHODS

Chemicals. All chemicals were purchased from Sigma-Aldrich. Enzymes and oligonucleotide sources are described elsewhere.

Bioinformatic analysis of *mepA* motif RNAs. The *mepA* motif was initially identified through the application of the GC-IGR genome analysis process^{23,24} to the genome *Clostridium botulinum* Strain A Hall as previously described.²³ Briefly, the bioinformatic search was conducted with the Infernal 1.1 software,⁴⁴ which was utilized to search genomic DNA sequences (RefSeq version 80) and microbial environmental sequence collections as described previously.⁴¹ Hypothetical

gene annotations were validated manually by conducting protein sequence homology searches with the NCBI Basic Local Alignment Search Tool (BLAST).⁴⁵ Covariation features were predicted using the comparative sequence analysis algorithm CMfinder.⁴⁶ The RNA consensus model was generated with R2R software⁴⁷ and manually adjusted to generate the final figure.

RNA oligonucleotide preparation. RNA oligonucleotides were prepared as previously described⁴⁸ using the appropriate synthetic DNA (Keck Oligo Synthesis Resource) templates for *in vitro* transcription (**Table S1**). Template DNAs were hybridized in equimolar amounts with their complementary strand to form a double-stranded DNA template containing a T7 RNA polymerase promoter sequence. The subsequent transcription reactions were incubated at 37°C for 4 hours. The resulting RNA transcripts were separated by using denaturing (8 M urea) 10% PAGE. RNAs were recovered from the gel, dephosphorylated using calf intestinal alkaline phosphatase, (New England Biolabs), and subsequently 5'-³²P-labeled using [γ -³²P]-ATP and T4 polynucleotide kinase (New England Biolabs) according to the manufacturer's protocols.

In-line probing assays. In-line probing reactions were conducted generally as described previously.^{27,28} Briefly, ³²P-labeled RNAs (trace amounts) were incubated in the absence or presence of ligand candidates as indicated at room temperature in the presence of 20 mM MgCl₂, 100 mM KCl, and 50 mM Tris-HCl (pH 8.3 at ~23°C). The reaction products were separated by denaturing (8 M urea) 10% PAGE and were visualized by using a Typhoon phosphorimager (GE Healthcare). As described previously,⁴⁸ band intensities were determined and used to estimate the fraction of RNAs bound to ligand. Values were plotted relative to the logarithm of the molar concentration of ligand, wherein half-maximal binding represents the K_D .

ACKNOWLEDGMENTS

We thank members of the Breaker laboratory for helpful discussions. Aparajita Balaji was supported in part by the National Institutes of Health Chemical Biology Training Grant (T32 GM067543). This work was also supported by NIH grants (GM022778 and AI136794) to R.R.B. Research in the Breaker laboratory is also supported by the Howard Hughes Medical Institute.

REFERENCES

- (1) McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E., Breaker, R. R. (2017) Riboswitch diversity and distribution. *RNA* 23 (7), 995-1011.
- (2) Lotz, T. S., and Suess, B. (2018) Small-molecule-binding riboswitches. *Microbiol. Spectr.* 6 (4), RWR-0025-2018.
- (3) Vézina Bédard, A.-S., Hien, E. D. M., and Lafontaine, D. A. (2020) Riboswitch regulation mechanisms: RNA, metabolites and regulatory proteins. *Biochim Biophys Acta Gene Regul. Mech.* 1863 (3), 194501.
- (4) Breaker, R. R. (2012) Riboswitches and the RNA World. *Cold Spring Harb. Perspect. Biol.* 4 (2), a003566.
- (5) Breaker, R. R. (2009) Riboswitches: from ancient gene control systems to modern drug targets. *Future Microbiol.* 4 (7), 771-773.
- (6) Benner, S. A., Ellington, A. D., and Tauer, A. (1989) Modern metabolism as a palimpsest of the RNA world. *Proc. Natl. Acad. Sci. USA* 86 (18), 7054-7058.
- (7) Chen, X., Li, N., and Ellington, A. D. (2007) Ribozyme catalysis of metabolism in the RNA World. *Chem. Biodivers.* 4 (4), 633-655.
- (8) Weinberg, Z., Wang, J.X., Bogue, J., Yang, J., Corbino, K., Moy, R.H., and Breaker, R.R. (2010) Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* 11, R31.
- (9) Baker, J.L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R.B., and Breaker, R.R. (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* 335 (6065), 233-235.
- (10) Stockbridge, R. B., Robertson, J. L., Kolmakova-Partensky, L., and Miller, C. (2013) A family of fluoride-specific ion channels with dual-topology architecture. *eLife* 2, e01084.
- (11) Li, S., Smith, K. D., Davis, J. H., Gordon, P. B., Breaker, R. R., and Strobel, S. A. (2013) Eukaryotic resistance to fluoride toxicity mediated by a widespread family of fluoride export proteins. *Proc. Natl. Acad. Sci. USA* 110 (47), 19018-19023.
- (12) Barrick, J. E., Corbino, K. A., Winkler, W. C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., Wickiser, J. K., and Breaker, R. R. (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. USA* 101 (17), 6421-6426.
- (13) Nelson, J. W., Sudarsan, N., Furukawa, K., Weinberg, Z., Wang, J. X., and Breaker, R. R. (2013) Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nat. Chem. Biol.* 9, 834-839.
- (14) Kellenberger, C. A., Wilson, S. C., Hickey, S. F., Gonzalez, T. L., Su, Y., Hallberg, Z. F., Brewer, T. F., Iavarone, A. T., Carlson, H. K., Hsieh, Y.-F., and Hammond, M. C. (2015) GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP. *Proc. Natl. Acad. Sci. USA* 112 (17), 5383-5388.
- (15) Nelson, J. W., Sudarsan, N., Phillips, G. E., Stav, S., Lünse, C. E., McCown, P. J., and Breaker, R. R. (2015) Control of bacterial exoelectrogenesis by c-AMP-GMP. *Proc. Natl. Acad. Sci. USA* 112 (17), 5389-5394.

- (16) Breaker, R. R., Atilho, R. M., Malkowski, S. N., Nelson, J. W., and Sherlock, M. E. (2017) The biology of free guanidine as revealed by riboswitches. *Biochemistry* 56 (2), 345-347.
- (17) Battaglia, R. A., and Ke, A. (2018) Guanidine-sensing riboswitches: How do they work and what do they regulate? *Wiley Interdiscip. Rev. RNA* 26, e1482.
- (18) Nelson, J. W., Atilho, R. M., Sherlock, M. E., Stockbridge, R. B., and Breaker, R. R. (2017) Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. *Mol. Cell* 65 (2), 220-230.
- (19) Sherlock, M. E., Malkowski, S. N., and Breaker, R. R. (2017) Biochemical validation of a second guanidine riboswitch class in bacteria. *Biochemistry* 56 (2), 352-358.
- (20) Sherlock, M. E. and Breaker, R. R. (2017) Biochemical validation of a third guanidine riboswitch class in bacteria. *Biochemistry* 56 (2), 359-363.
- (21) Schneider, N. O., Tassoulas, L. J., Zeng, D., Laseke, A. J., Reiter, N. J., Wackett, L. P., and St Maurice, M. (2020) Solving the conundrum: widespread proteins annotated for urea metabolism in bacterial are carboxyguanidine deiminases mediating nitrogen assimilation from guanidine. *Biochemistry* 59 (35), 3258-3270.
- (22) Kermani, A. A., Macdonald, C. B., Gundepudi, R., and Stockbridge, R. B. (2018) Guanidinium export is the primal function of SMR family transporters. *Proc. Natl. Acad. Sci. USA* 115 (12), 3060-3065.
- (23) Meyer, M. M., Ames, T. D., Smith, D. P., Weinberg, Z., Schwalbach, M. S., Giovannoni, S. J., and Breaker, R. R. (2009) Identification of candidate structured RNAs in the marine organism ‘*Candidatus Pelagibacter ubique*’. *BMC Genom.* 10, 268.
- (24) Stav, S. S., Atilho, R. M., Mirihana Arachchilage, G., Nguyen, G., Higgs, G., and Breaker, R. R. (2019) Genome-wide discovery of structured noncoding RNAs in bacteria. *BMC Microbiol.* 19, 66.
- (25) Wilson, K. S., and von Hippel, P. H. (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc. Natl. Acad. Sci. USA* 92 (19), 8793-8797.
- (26) Yarnell, W. S., and Roberts, J. W. (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* 284 (5415), 611-615.
- (27) Soukup, G.A., and Breaker, R.R. (1999) Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* 5 (10), 1308-1325.
- (28) Regulski, E.E., and Breaker, R.R. (2008) In-line probing analysis of riboswitches. In: Wilusz J. (eds) Post-transcriptional Gene Regulation. *Methods Mol. Biol.* 419, 53-67.
- (29) Breaker, R. R. (2018) Riboswitches and translation control. *Cold Spring Harb. Perspect. Biol.* 10 (11), a032797.
- (30) Monforte, J. A., Kahn, J. D., and Hearst, J. E. (1990) RNA folding during transcription by *Escherichia coli* RNA polymerase analyzed by RNA self-cleavage. *Biochemistry* 29 (34), 7882-2890.

- (31) Komissarova, N., and Kashlev, M. (1998) Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proc. Natl. Acad. Sci. USA* 95 (25), 14699-14704.
- (32) Vassylyev, D. G., Senkine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417 (6890), 712-719.
- (33) Battaglia, R. A., Price, I. R., and Ke, A. (2017) Structural basis for guanidine sensing by the *ykkC* family of riboswitches. *RNA* 23 (4), 578-585.
- (34) Reiss, C. A., Xiong, Y., and Strobel, S. A. (2017) Structural basis for ligand binding to the guanidine-I riboswitch. *Structure* 25 (1), 195-202.
- (35) Huang, L., Wang, J., and Lilley, D. M. J. (2017) The structure of the guanidine-II riboswitch. *Cell Chem. Biol.* 24 (6), 695-702.
- (36) Reiss, C. A., and Strobel, S. A. (2017) Structural basis for ligand binding to the guanidine-II riboswitch. *RNA* 23 (9), 1338-1343.
- (37) Huang, L., Wang, J., Wilson, T. J., and Lilley, D. M. J. (2017) Structure of the guanidine-III riboswitch. *Cell Chem. Biol.* 24 (11), 1407-1415.
- (38) Lilley, D. M. J. (2017) The guanidine riboswitch – a poor orphan no longer. *Cell Chem. Biol.* 24 (2), 130-131.
- (39) Higgins, D. A., Gladden, J. M., Kimbrel, J. A., Simmons, B. A., Singer, S. W., and Thelen, P. (2019) Guanidine riboswitch-regulated efflux transporters protect bacteria against ionic liquid toxicity. *J. Bacteriol.* 201 (13), e00069-19.
- (40) Slipski, C., Jamieson, T. R., Zhanel, G. G., and Bay, D. C. (2020) Riboswitch associated guanidinium selective efflux pumps frequently transmitted on proteobacterial plasmids increase *Escherichia coli* biofilm tolerance to disinfectants. *J. Bacteriol.* 14, JB.00104-20.
- (41) Huang, L., Wang, J., Wilson, T. J., and Lilley, D. M. J. (2019) Structure-guided design of a high-affinity ligand for a riboswitch. *RNA* 25 (4), 423-430.
- (42) Weinberg, Z., Lünse, C. E., Corbino, K. A., Ames, T. D., Nelson, J. W., Roth, A., Perkins, K. R., Sherlock, M. E., and Breaker, R. R. (2017) Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.* 45 (18), 10811-10823.
- (43) Atilho, R. M., Mirihana-Arachchilage, G., Greenlee, E. B., Knecht, K. M., and Breaker, R. R. (2019) A bacterial riboswitch class for the thiamin precursor HMP-PP employs a terminator-embedded aptamer. *eLife* 8, e45210.
- (44) Nawrocki, E. P., and Eddy, S. R. (2013) Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29 (22), 2933-2935.
- (45) Altschul, S., Gish, W., Miller, W., Myers, E. W., and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215 (3), 403-410.
- (46) Yao, Z., Weinberg, Z., and Ruzzo, W. L. (2006) CMfinder - a covariance model based RNA motif finding algorithm. *Bioinformatics* 22 (4), 445-452.

- (47) Weinberg, Z., and Breaker, R. R. (2011) R2R - software to speed the depiction of aesthetic consensus RNA secondary structures. *BMC Bioinformatics* 12, 3
- (48) Malkowski, S. N., Spencer, T. C. J., and Breaker, R. R. (2019) Evidence that the *nadA* motif is a bacterial riboswitch for the ubiquitous enzyme cofactor NAD⁺. *RNA* 25 (12), 1616-1627.

Figure 1. The *mepA* motif RNA associates with genes regulated by experimentally validated guanidine riboswitches. (A) Consensus sequence and secondary structure model of the *mepA* motif RNA based on 1275 homologous but non-identical sequences. P1 and P2 denote predicted base-paired stems, wherein P2 is incompatible with the formation of the predicted intrinsic terminator stem. (B) Genetic contexts of the *mepA* motif. Genes located immediately downstream of the motif are identified and reported (total of 1250). The difference between the number of adjacent genes recorded and the number of non-identical representatives is due to the fact that some representatives present in metagenomic sequence data lack sufficient read length to identify the neighboring gene. Additional details regarding the associated genes are presented in **Figure S1**. (C) Chemical structures of guanidinium and its analogs that exhibit the ability to be bound by a representative *mepA* RNA aptamer.

Figure 2. A *mepA* motif RNA construct undergoes a structural change when binding guanidine. (A) Sequence and secondary structure model of a representative *mepA* motif RNA from *C. botulinum* B1 strain Okra. Two constructs examined in this study, 81 *mepA* and 97 *mepA*, have identical nucleotides on the 5' side, but terminate at nucleotide positions 81 and 97, respectively. The two non-native guanosine nucleotides on the 5' terminus (lowercase g letters) are not included in the numbering system used. Mutations present in constructs M1 (G26U) and M2 (A63U) are boxed. Results from in-line probing reactions using the 81 *mepA* RNA construct are derived from the data depicted in B. Colored circles indicate the sites undergoing substantial spontaneous strand scission in the absence and/or presence of guanidine. (B) Autoradiogram of the polyacrylamide gel electrophoresis (PAGE) separation of the products of in-line probing reactions. ³²P-labeled 81 *mepA* RNA was subjected to in-line probing reactions in the absence (–) or presence (+) of 1 mM guanidine. Full-length 81 *mepA* precursor (Pre) RNAs present in the first three lanes were loaded without incubation (NR, no reaction), or were subjected to partial digestion with RNase T1 (T1, cleaves after G nucleotides) or alkali (OH⁻, cleaves after each nucleotide). Bands corresponding to strand scission after certain G nucleotides are labeled. In addition, regions corresponding to predicted P1, P2 and partial terminator (T*) base-paired structures are annotated. Regions exhibiting substantial modulation on the addition of guanidine are denoted sites 1, 2 and 3. (C) Portion of an autoradiogram of the PAGE separation of in-line probing reaction products generated near sites 1, 2 and 3 in the presence of various concentrations of guanidine ranging from

214 nm to 10 mM at every third log unit. (D) Plot of the fraction of 81 *mepA* RNA bound to ligand at various concentrations (c) of guanidine. Fraction bound values were derived from the data depicted in C, wherein the bands measured were G26 (site 1), U41 (site 2) and G65 (site 3). The line represents a theoretical curve for a 1:1 receptor-ligand interaction with a K_D of 64 μ M.

Figure 3. Requirements for ligand binding by the 81 *mepA* guanidine aptamer. (A) In-line probing analysis of 32 P-labeled 81 *mepA* RNA in the presence of 1 or 10 mM of the ligands indicated. Other annotations are as described in the legend to Figure 2B. (B) Plot of the K_D values determined (filled circles) or estimated (open circle) for the ligands examined in A. See **Figure S2** for additional K_D analyses. Other guanidine analogs tested (1,1-dimethylguanidine; 1,3-diaminoguanidine; 1,1-dimethylbiguanidine; urea) exhibited no evidence of binding at 10 mM. (C) In-line probing analysis of radiolabeled WT (wild-type), M1 and M2 versions of the 81 *mepA* RNA construct. Guanidine concentrations tested were 0.1 and 1 mM. Other annotations, including the three sites of structural modulation exhibited by the WT RNA are as described in the legend to **Figure 2B**.

Figure 4. The mechanism of gene regulation employed by guanidine-IV riboswitches. (A) Truncated versions of the *mepA* RNA sequence from *C. botulinum* B1 strain Okra used to evaluate the importance of the terminator stem length on guanidine binding. (B) In-line probing analysis of various 32 P-labeled *mepA* RNA constructs in the absence (–) or presence (+) of 1 mM guanidine. The bands corresponding to the terminator stem (T) are indicated. Other annotations are as described in the legend to **Figure 2B**. (C). Model depicting architectural features of *mepA* motif RNAs expected to be important for gene regulation by guanidine-IV riboswitches. The aptamer domain is likely formed by two regions of the RNA that carry highly conserved nucleotides (red letters). The length (L) of the single-stranded linker joining P1 and the terminator stem is proposed to be critical for the hairpin loops to make contact and form the guanidine aptamer.

Figure 5. Internal deletions to the 81 *mepA* RNA construct support the proposed regulatory mechanism. (A) Internal deletions of the 81 *mepA* RNA were conducted in group of three nucleotides to create five additional constructs truncated within the proposed linker region (gray shading). (B) In-line probing analysis of 32 P-labeled internal deletion constructs based on the 81

mepA RNA in the absence (–) or presence (+) of 1 mM guanidine. Annotations in gray apply only to the $\Delta 49-51$ construct. Other annotations are as described in the legend to **Figure 2B**. (C) Model depicting the key features of a guanidine-IV riboswitch, indicating the requirements for the length of the single-stranded linker. Other annotations are as described in the legend to **Figure 4C**.

Figure 1

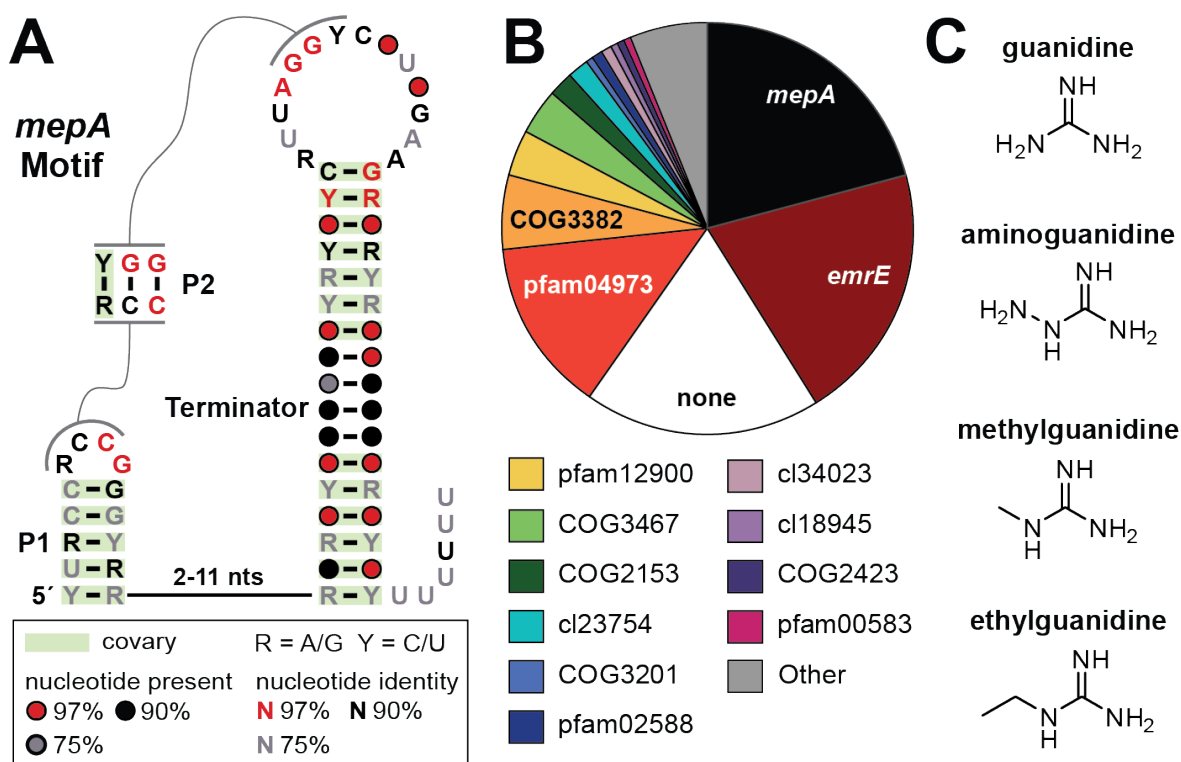


Figure 2

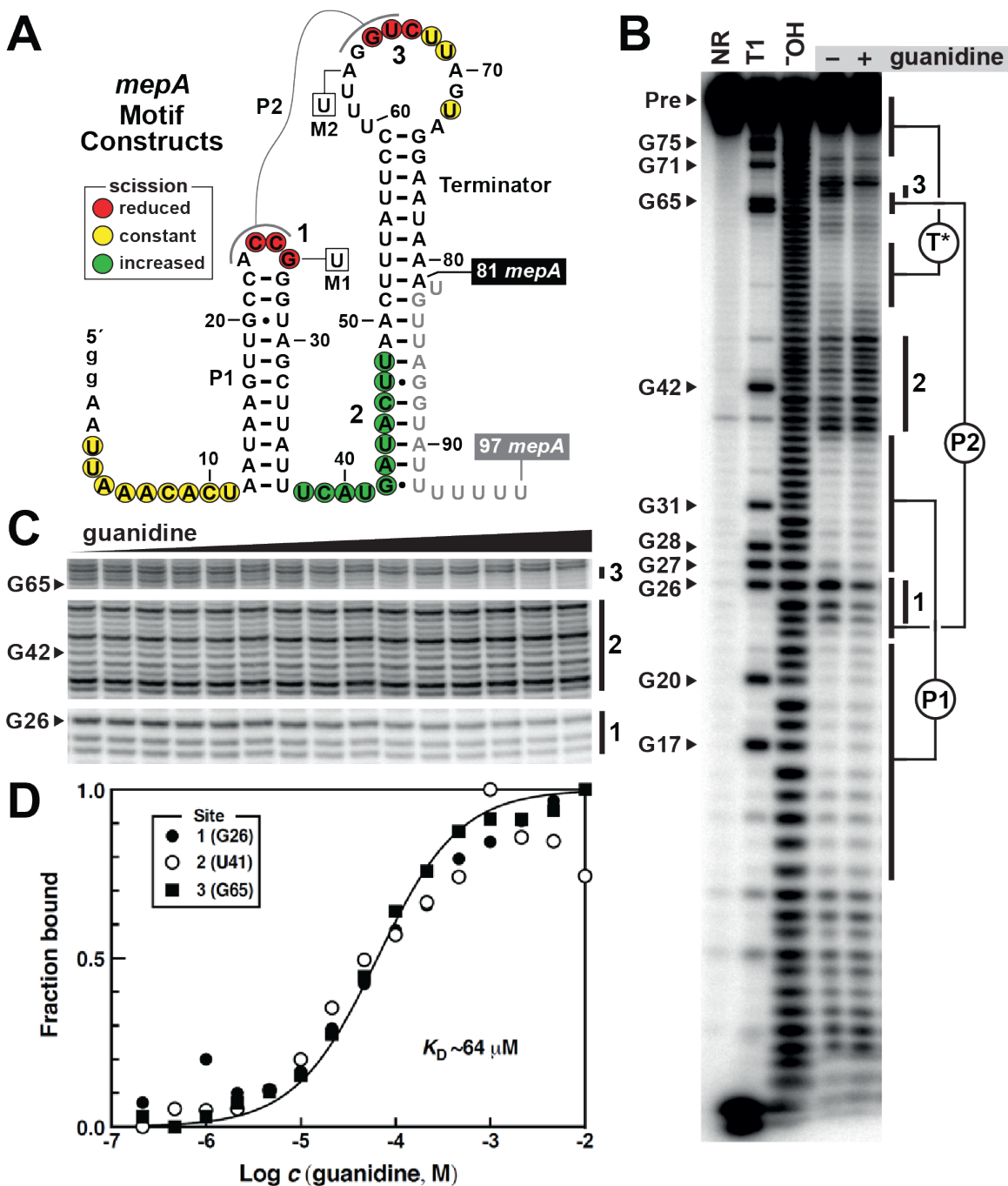


Figure 3

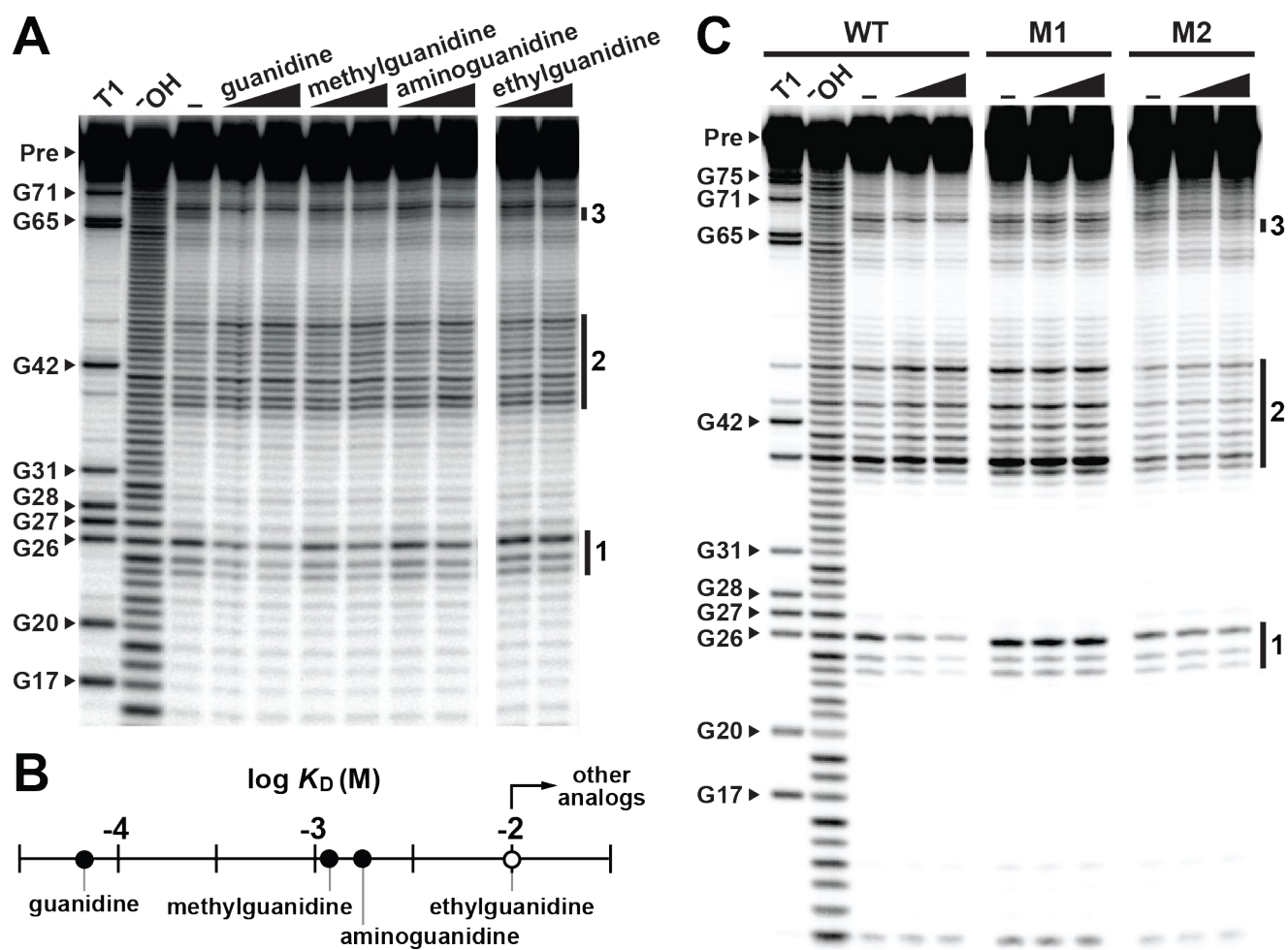


Figure 4

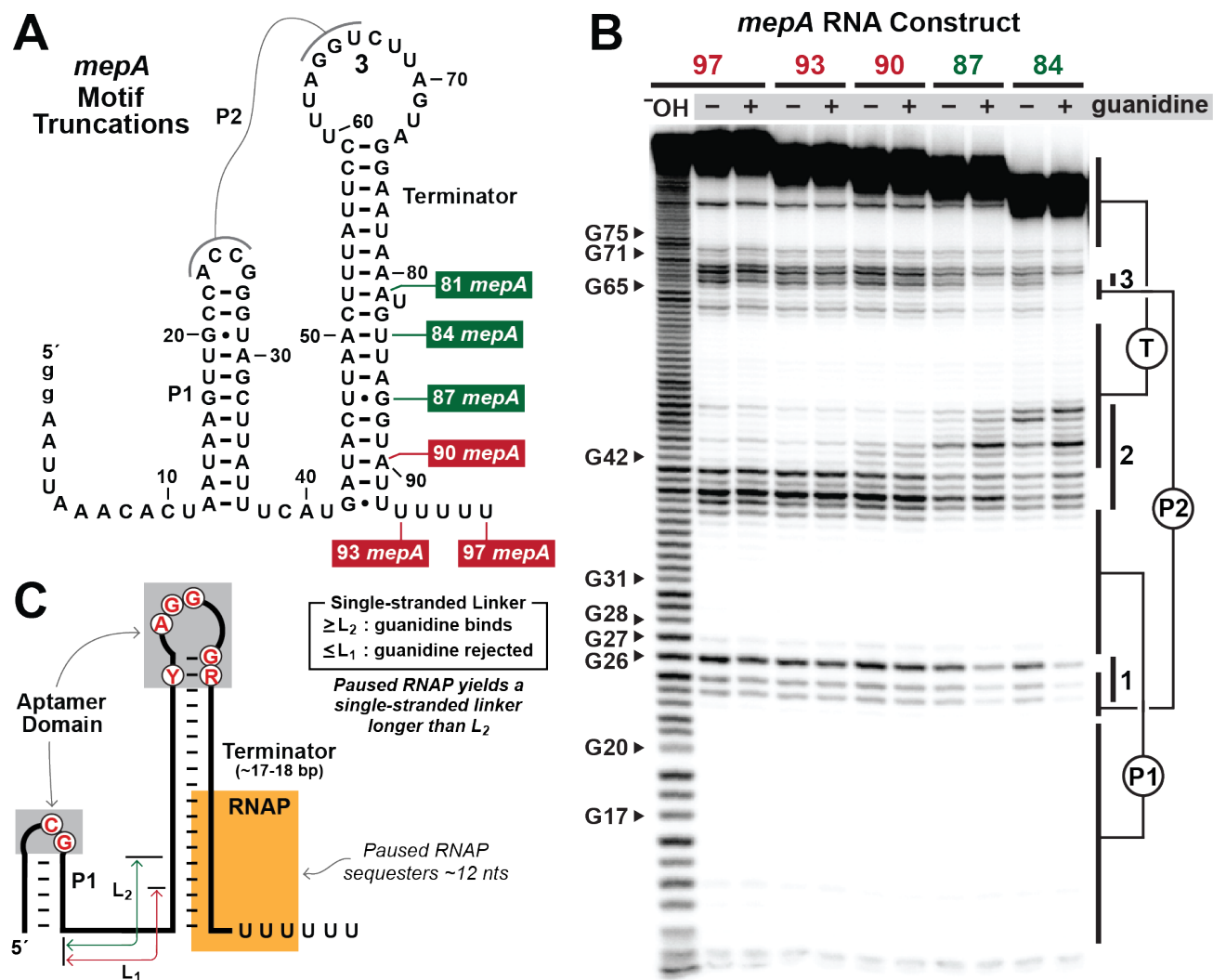
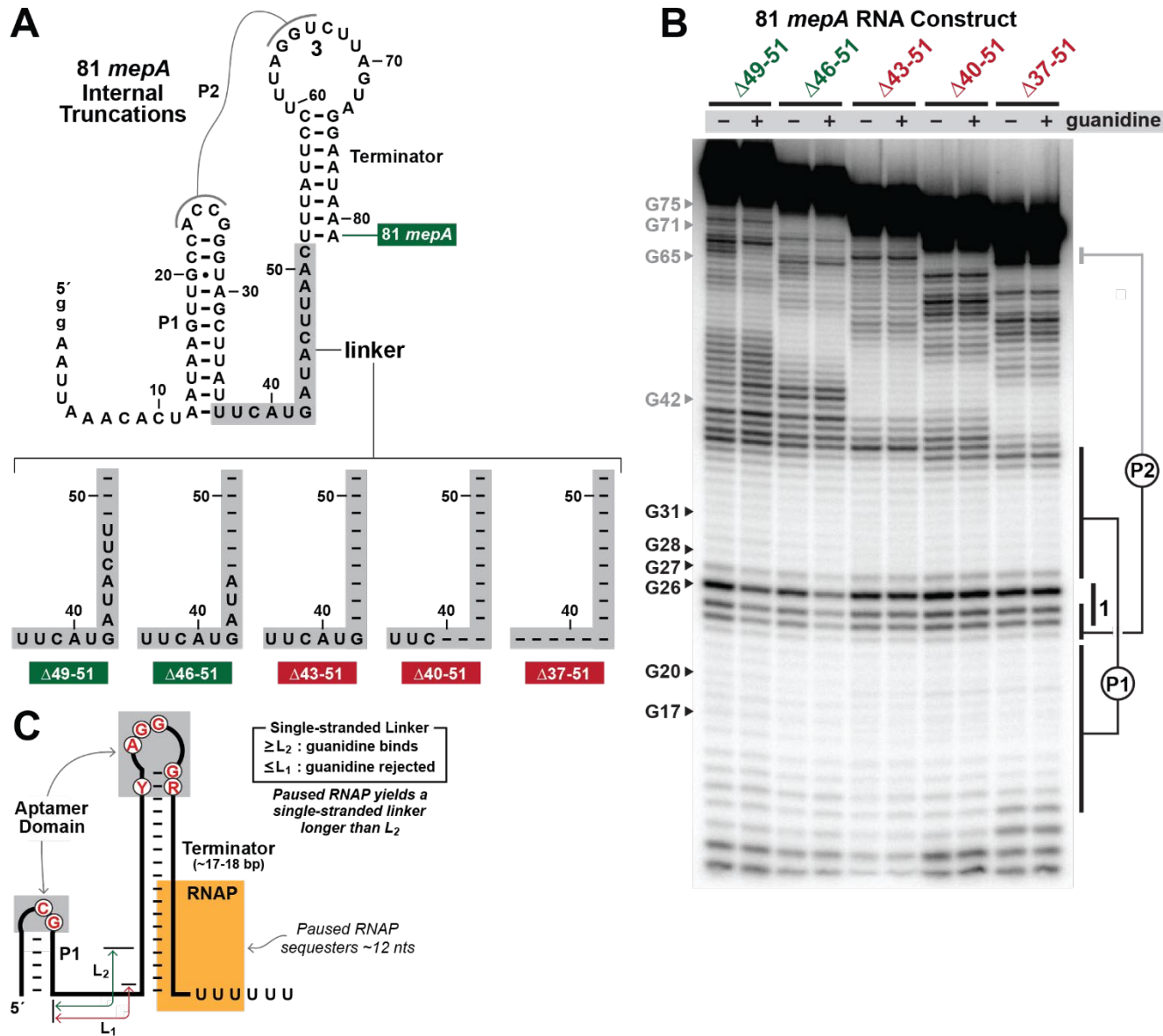


Figure 5



Supplementary Information

Biochemical Validation of a Fourth Guanidine Riboswitch Class in Bacteria

Hubert Salvail,^{†,‡} Aparaaajita Balaji,^{†,‡} Diane Yu,[†] Adam Roth,[‡] and Ronald R. Breaker^{†,‡,§,*}

[†]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8103, USA

[‡]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520-8103, USA

[§]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8103, USA

Corresponding Author

*E-mail: ronald.breaker@yale.edu

[‡]These authors contributed equally to this work

Contents:

Fig. S1: Gene Associations

Fig. S2: Analog K_D Values

Table S1: Oligonucleotides (pp. S2-S4).

Gene Name	Count	Predicted Function
<i>mepA</i>	261	Multidrug and toxic compound extrusion (MATE)-like transporter (CD13143)
<i>emrE</i>	255	Multidrug and toxic compound extrusion (MATE)-like transporter (CD13143)
No first gene	231	N/A
pfam04973	173	Nicotinamide mononucleotide (NMN) transporter
COG3382	70	B3/B4 domain (DNA/RNA-binding domain of Phe-tRNA-synthetase)
pfam12900	47	Pyridoxamine 5'-phosphate (PMP) oxidase
COG3467	45	NimA - Nitroimidazol reductase
COG2153	23	ElaA - Predicted N-acyltransferase (GNAT family)
c123754	20	EamA super family - Hypothetical membrane proteins of unknown function (DUF6)
COG3201	10	PnuC - Nicotinamide riboside transporter
pfam02588	10	YitT - Uncharacterized bacterial ABC transporter permease
c134023	9	Tyw1 super family - Tyrosine [tRNA(Phe)-imidazo G37] synthetase, radical SAM superfamily
c118945	8	Aspartate aminotransferase (AAT) superfamily (fold type I)
COG2423	7	OCDMu - Ornithine cyclodeaminase/archaeal alanine dehydrogenase, mu-crystallin family
pfam00583	6	Acetyltransferase (GNAT family)
Other (Count <5)	75	
Total	1250	

Figure S1. Existing annotations for genes associated with *mepA* motif representatives. N/A indicated not applicable.

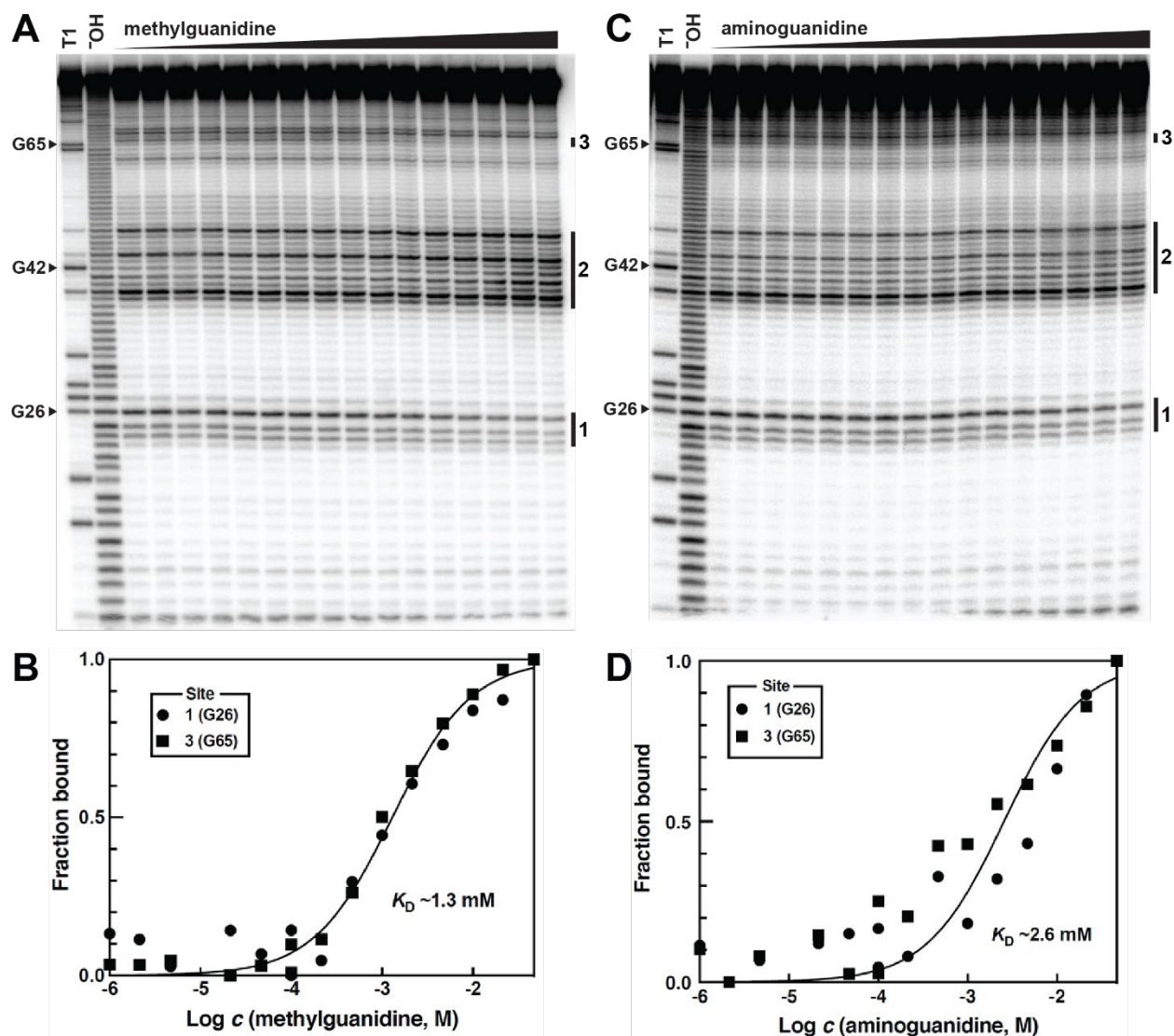


Figure S2. In line probing analyses to establish the apparent K_D values for guanidine analogs. N/A indicated not applicable. (A) The ^{32}P -labeled 81 *mepA* RNA construct was subjected to in-line probing analysis with concentrations of methylguanidine ranging from 1 μM to 46 mM at each $1/3^{\text{rd}}$ log unit. (B) Plot of the fraction of 81 *mepA* RNA bound to ligand at various concentrations (c) of methylguanidine. (C) The analysis described in A using aminoguanidine. (D) KD analysis plot as described in B. Additional annotations are as described in the legend to **Figure 2**.

Table S1. Sequences of synthetic DNA oligonucleotides used in this study.

Name	Sequence (5' to 3')	Annotation
HS003	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTAGGTCTTAGTAGGAATAA ATGTTAGGTATTTTTT	Forward template for transcription of the WT <i>C. botulinum</i> B1 Okra 97 <i>mepA</i> RNA
HS004	AAAAAATACCTAACATTTATTCCTACTAAGAC CTAAAGGAATAAAGTTAAGTATCATGAAATAA GCTACCCGGTGGCAACTTATTAGTGTTTAATTC CTATAGTGAGTCGTATTA	Reverse template for transcription of the WT <i>C. botulinum</i> B1 Okra 97 <i>mepA</i> RNA
HS005	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTAGGTCTTAGTAGGAATAA A	Forward template for transcription of the WT <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS006	TTTATTCCTACTAAGACCTAAAGGAATAAAGTT AAGTATCATGAAATAAGCTACCCGGTGGCAACT TATTAGTGTTTAATTccTATAGTGAGTCGTATTA	Reverse template for transcription of the WT <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS023	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTAGGTCTTAGTAGGAATAA ATGT	Forward template for transcription of the WT <i>C. botulinum</i> B1 Okra 84 <i>mepA</i> RNA
HS024	ACATTTATTCCTACTAAGACCTAAAGGAATAAA GTTAAGTATCATGAAATAAGCTACCCGGTGGCA ACTTATTAGTGTTTAATTCCTATAGTGAGTCGTA TTA	Reverse template for transcription of the WT <i>C. botulinum</i> B1 Okra 84 <i>mepA</i> RNA

HS025	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTAGGTCTTAGTAGGAATAA ATGTTAG	Forward template for transcription of the WT <i>C. botulinum</i> B1 Okra 87 <i>mepA</i> RNA
HS026	CTAACATTTATTCCTACTAAGACCTAAAGGAAT AAAGTTAAGTATCATGAAATAAGCTACCCGGTG GCAACTTATTAGTGTTAATTCCTATAGTGAGTC GTATTA	Reverse template for transcription of the WT <i>C. botulinum</i> B1 Okra 87 <i>mepA</i> RNA
HS027	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTAGGTCTTAGTAGGAATAA ATGTTAGGTA	Forward template for transcription of the WT <i>C. botulinum</i> B1 Okra 90 <i>mepA</i> RNA
HS028	TACCTAACATTTATTCCTACTAAGACCTAAAGG AATAAAGTTAAGTATCATGAAATAAGCTACCCG GTGGCAACTTATTAGTGTTAATTCCTATAGTG AGTCGTATTA	Reverse template for transcription of the WT <i>C. botulinum</i> B1 Okra 90 <i>mepA</i> RNA
HS029	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTAGGTCTTAGTAGGAATAA ATGTTAGGTATTT	Forward template for transcription of the WT <i>C. botulinum</i> B1 Okra 93 <i>mepA</i> RNA
HS030	AAATACCTAACATTTATTCCTACTAAGACCTAA AGGAATAAAGTTAAGTATCATGAAATAAGCTA CCCGGTGGCAACTTATTAGTGTTAATTCCTATA GTGAGTCGTATTA	Reverse template for transcription of the WT <i>C. botulinum</i> B1 Okra 93 <i>mepA</i> RNA

HS033	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCTGGTAGCTTATTTTCATGATACTT AACTTTATTCCTTTAGGTCTTAGTAGGAATAAA	Forward template for transcription of the G26U mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS034	TTTATTCCTACTAAGACCTAAAGGAATAAAGTT AAGTATCATGAAATAAGCTACCAGGTGGCAACT TATTAGTGTTAATTCCTATAGTGAGTCGTATTA	Reverse template for transcription of the G26U mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS043	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TAACTTTATTCCTTTTGGTCTTAGTAGGAATAAA	Forward template for transcription of the A63U mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS044	TTTATTCCTACTAAGACCAAAGGAATAAAGTT AAGTATCATGAAATAAGCTACCCGGTGGCAACT TATTAGTGTTAATTCCTATAGTGAGTCGTATTA	Reverse template for transcription of the A63U mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS051	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTATTCCTTTA GGTCTTAGTAGGAATAAA	Forward template for transcription of the Δ 37-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS052	TTTATTCCTACTAAGACCTAAAGGAATAAAATA AGCTACCCGGTGGCAACTTATTAGTGTTAATT CCTATAGTGAGTCGTATTA	Reverse template for transcription of the Δ 37-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA

HS053	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTCTTATTCCT TTAGGTCTTAGTAGGAATAAA	Forward template for transcription of the Δ 40-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS054	TTTATTCCTACTAAGACCTAAAGGAATAAAGAA ATAAGCTACCCGGTGGCAACTTATTAGTGTTA ATTCCTATAGTGAGTCGTATTA	Reverse template for transcription of the Δ 40-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS055	TAATACGACTCACTATAggAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTCATGTTTATT CCTTTAGGTCTTAGTAGGAATAAA	Forward template for transcription of the Δ 43-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS056	TTTATTCCTACTAAGACCTAAAGGAATAAACAT GAAATAAGCTACCCGGTGGCAACTTATTAGTGT TTAATTCCTATAGTGAGTCGTATTA	Reverse template for transcription of the Δ 43-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS057	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTCATGATATT TATTCCTTTAGGTCTTAGTAGGAATAAA	Forward template for transcription of the Δ 46-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS058	TTTATTCCTACTAAGACCTAAAGGAATAAATAT CATGAAATAAGCTACCCGGTGGCAACTTATTAG TGTTAATTCCTATAGTGAGTCGTATTA	Reverse template for transcription of the Δ 46-51 deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA

HS059	TAATACGACTCACTATAGGAATTAACACTAAT AAGTTGCCACCGGGTAGCTTATTTTCATGATACT TTTTATTCCTTTAGGTCTTAGTAGGAATAAA	Forward template for transcription of the $\Delta 49-51$ deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA
HS060	TTTATTCCTACTAAGACCTAAAGGAATAAAAAG TATCATGAAATAAGCTACCCGGTGGCAACTTAT TAGTGTTTAATTCCTATAGTGAGTCGTATTA	Reverse template for transcription of the $\Delta 49-51$ deletion mutant <i>C. botulinum</i> B1 Okra 81 <i>mepA</i> RNA