Identification of a novel riboswitch class and its application in *S. aureus*

Dissertation zur Erlangung des

akademischen Grades eines Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Felina Lenkeit

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Chemie

Konstanz, 2021

Tag der mündlichen Prüfung: 03.12.2021

1. Referent/Referentin: Prof. Dr. Jörg Hartig

2. Referent/Referentin: Prof. Dr. David Schleheck

Notes to the reader

The presented thesis is divided into four chapters that can be read independently.

Chapter 1 contains a general introduction providing background knowledge regarding regulatory RNAs. Structures and mechanisms of natural riboswitches are discussed as well as strategies for identification of novel riboswitches. Chapter 2 presents the identification and characterization of a new class of guanidine riboswitches. Chapter 3 contains the optimization of a controllable gene expression system in *S. aureus*. Chapter 4 gives new insights into the characterization of additional orphan riboswitches and their genetic context.

The main part of the experiments was conducted in the laboratory of Prof. Dr. Jörg S. Hartig (Department of Chemistry, University of Konstanz) from December 2017 until July 2021. In-line probing experiments with the *folP* RNA motif (shown in Chapter 4) were conducted in the laboratory of Prof. Dr. Ronald Breaker (Department of Molecular, Cellular and Developmental Biology, Yale University) in August and September 2019. Bioinformatics analysis for identification of novel riboswitch candidates were performed in collaboration by Dr. Zasha Weinberg (Department of Computational Science, University of Leipzig).

Parts of Chapter 2 are published in "Discovery and characterization of a fourth class of guanidine riboswitches" (Lenkeit, F. *et al*. 2020, Nucleic Acids Research).

Table of Contents

1. Abstract

RNA can fold into complex secondary and tertiary structures and thereby use various mechanisms with a wide range of physiological responses to control the expression of certain genes. In bacteria, especially noncoding RNA regions are involved in the regulation of gene expression. Riboswitches are such noncoding regulatory elements in the 5´-UTR of mRNA that control expression of the downstream gene in response to binding of a certain ligand. They are composed of two functional domains: an aptamer and an expression platform. The aptamer domain can form a highly selective binding pocket to sense the cognate ligand and induce structural modulation in the expression platform that leads to a change of gene expression. Their mechanisms to control gene expression include modulation of transcription, translation and mRNA stability. Since riboswitches specifically sense intramolecular metabolites, acting without any protein intermediates, they have various applications. They can be used for controllable expression systems in the laboratory and are promising tools in gene therapy. Riboswitches are also potential targets for antimicrobial agents since they control fundamental pathways in bacteria. Moreover, the discovery of natural riboswitches and characterization of their associated regulatory network gives insight into functions of associated genes and proteins.

This thesis presents in Chapter 2 the identification and complete characterization of a new natural riboswitch class. The *GGAM-1* motif shows several properties that are characteristic for riboswitches and is associated with guanidine related genes. Taken this into account, we hypothesized that the *GGAM-1* motif represents a new class of guanidine sensing riboswitches. We were able to demonstrate *in vivo* and *in vitro* that the *GGAM-1* motif regulates expression of its downstream gene in respond to selective binding of guanidine. Therefore, we call the motif the *guanidine-IV* riboswitch. Moreover, we were able to identify nucleotides involved in selective guanidine sensing and give insight into the mechanism of ligand binding and subsequent structural modulation. All representatives of this riboswitch class were found to be associated with an intrinsic terminator and are demonstrated to function as transcriptional ON-switches in which ligand binding stabilizes a conformation that excludes terminator stem formation.

In Chapter 3 we present the optimization of a plasmid-based system in *S. aureus* that uses the *guanidine-IV* riboswitch to conditionally control reporter expression. *S. aureus* is one of the major pathogens today, causing a wide range of infections and considerable human mortality. With the growing need to understand pathogenic mechanisms, also the need for tools to control and monitor gene expression increases. Using the *guanidine-IV* based system, established and optimized in this work, protein expression can be induced in *S. aureus* at different growth stages from early to late exponential phase with a 120-fold change. The inducible high-fold change makes it not only a

promising system to study unknown protein function, but also for use as biosensor. We show that the system efficiently monitors accumulation and concentration of guanidine and analogues.

In the last chapter of this work, we introduce three riboswitch candidates. These motifs exhibit properties that are characteristic for riboswitches. Based on their genetic context, educated guesses provided metabolites that were tested in *in vitro* as well as *in vivo* assays. Our results give insights into the functionality of the riboswitch candidates and their associated network.

2. Zusammenfassung

RNA kann komplexe Sekundär- und Tertiärstrukturen eingehen und dabei verschiedene Mechanismen nutzen, um die Expression bestimmter Gene mit einer großen Bandbreite an physiologischen Reaktionen zu steuern. In Bakterien sind vor allem nichtcodierende Regionen in der RNA an der Regulation von Genexpression beteiligt. Riboswitche sind solche nichtcodierenden, regulatorischen Elemente in der mRNA, die die Expression eines assoziierten Gens in Abhängigkeit der Bindung des Liganden steuern. Sie bestehen aus zwei funktionellen Domänen: einer Aptamer-Domäne und einer Expressionsplattform. Die Aptamer-Domäne kann eine hochselektive Bindungstasche bilden, um den Liganden spezifisch zu erkennen und eine strukturelle Veränderung in der Expressionsplattform hervorzurufen, die wiederum die Genexpression beeinflusst. Die Mechanismen von Riboswitchen zur Kontrolle der Genexpression umfassen die Modulation der Transkription, Translation und der mRNA-Stabilität. Da Riboswitche spezifisch intramolekulare Metaboliten erkennen und ohne die Beteiligung von Proteinen funktionieren, haben sie viele verschiedene Anwendungsmöglichkeiten. Sie können für kontrollierbare Expressionssysteme im Labor eingesetzt werden und sind vielversprechende Werkzeuge in der Gentherapie. Riboswitche sind auch potenzielle Ziele für antimikrobielle Wirkstoffe, da sie grundlegende Signalwege in Bakterien kontrollieren. Darüber hinaus gibt die Entdeckung von natürlichen Ribowitchen und die Charakterisierung ihres regulatorischen Netzwerkes, Aufschluss über die Funktionen der zugehörigen Gene und Proteine.

Diese Arbeit stellt in Kapitel 2 die Identifizierung und vollständige Charakterisierung eines neuen natürlichen Riboswitches vor. Das *GGAM-1* Motiv zeigt mehrere Eigenschaften die für Riboswitche charakteristisch sind und ist assoziiert mit Genen die im Zusammenhang mit Guanidin stehen. Daher stellten wir die Hypothese auf, dass das *GGAM-1* Motiv eine neue Klasse von Guanidin bindenden Riboswitchen darstellt. Wir konnten sowohl *in vivo*, als auch *in vitro* zeigen, dass das *GGAM-1* Motiv die Expression des nachgelagerten Gens in Reaktion auf die selektive Bindung von Guanidin reguliert. Daher bezeichnen wir dieses Motiv als Guanidin-IV Riboswitch. Er ist immer mit einem intrinsischen Terminator assoziiert und fungiert als transkriptioneller AN-Schalter. Die Bindung des Liganden stabilisiert eine Struktur, in der kein intrinsischer Terminator ausgebildet werden kann.

In Kapitel 3 stellen wir die Optimierung eines Systems in *S. aureus* RN4220 vor, das den Guanidin-IV Riboswitch zur konditionalen Kontrolle der Reporterexpression nutzt. *S. aureus* ist heute einer der wichtigsten Krankheitserreger, der ein breites Spektrum an Infektionen und eine beträchtliche menschliche Sterblichkeit verursacht. Mit der wachsenden Notwendigkeit, die pathogenen Mechanismen zu verstehen, steigt der Bedarf an Werkzeugen zur Kontrolle und Überwachung der Genexpression in *S. aureus*. Mit dem in dieser Arbeit etablierten System kann die Proteinexpression in *S. aureus* in verschiedenen Wachstumsstadien von der frühen bis zur späten

exponentiellen Phase induziert werden. Die induzierbare hohe Expression macht es nicht nur zu einem vielversprechenden System für die Untersuchung unbekannter Proteinfunktionen, sondern auch für den Einsatz als Biosensor. Wir zeigen, dass das System effizient die Akkumulation und Konzentration von Guanidin detektiert.

Im letzten Kapitel dieser Arbeit stellen wir drei Riboswitch-Kandidaten vor. Alle diese Motive weisen Eigenschaften auf, die für Riboswitche charakteristisch sind. Basierend auf ihrem genetischen Kontext lieferten fundierte Vermutungen Metaboliten, die sowohl in *in vitro* als auch in *in vivo* Assays getestet wurden. Unsere Ergebnisse geben Einblicke in die Funktionalität dieser Riboswitch-Kandidaten und ihres zugehörigen Netzwerks.

3. Chapter I: General Introduction

The central dogma of molecular biology, stated by Francis Crick in 1957, describes the cellular flow of genetic information[1]. Sequence information, stored in the deoxyribonucleic acid (DNA), is transferred to ribonucleic acid (RNA) that passes the information into protein[2]. RNA is a polymeric and multifunctional molecule, originated by the process of DNA transcription. Comparing composition and structural features of DNA and RNA there are three main differences. First, RNA contains ribose instead of deoxyribose. Ribose has an additional hydroxy group attached to the 2´ position of the pentose ring. This makes the backbone of RNA chemically less stable than DNA by lowering the activation energy of hydrolysis. Second, like DNA, RNA contains the bases guanine (G), cytosine (C) and adenine (A), but thymine (T) is mostly replaced by its unmethylated version uracil (U). Third, RNA is assembled as chain of nucleotides, but unlike DNA it is mostly found to be single stranded. The lack of a complementary strand gives RNAs the possibility to fold into complex tertiary structures due to its high rotational freedom in the backbone[3]. For instance, RNA can fold back on itself and form base paired segments between complementary sequences, called hairpins. These base paired segments can also exhibit segments of unpaired nucleotides to form loops and bulges.

To synthesize a new strand of RNA, RNA polymerase uses a DNA template and catalyzes the formation of phosphodiester bonds between nucleoside triphosphates (NTPs). In bacteria, RNA Polymerase core enzyme together with sigma factors, form a holoemzyme that can initiate transcription at a specific promotor sequence. Once the polymerase escapes the promotor region to enter the phase of transcription elongation, RNA is synthesized with a rate of ~50 $nt*s^{-1}$ [4]. The transcription process stops at a termination site that can be either intrinsic or factor-dependent. An intrinsic terminator consists of a short palindromic repeat followed by at least six adenosines in the DNA. When this region is transcribed, a terminator stem is formed. This hairpin structure forces the polymerase to pause transcription and the transcription elongation complex (TEC) dissociates from the DNA template. Factor-dependent termination on the other hand is mediated by the protein Rho. It consists of six monomers that form a ring, moving upstream along the RNA until it reaches the RNA Polymerase to trigger dissociation of the TEC. Generally, RNA that is used to convey genetic information stored in the DNA is called messenger RNA (mRNA). The sequence information of the mRNA can be processed by ribosomes to synthesize a certain protein. Besides this messenger function however, RNA performs several tasks within the cell. Today we know that RNA is involved in various cellular processes like coding[5], regulation of gene expression[6], catalysis of chemical reactions and the sensing of cellular signals and its communication[7].

3.1. Regulatory functions of bacterial RNAs

Due to its property to fold into complex tertiary structures, RNA can use various mechanisms like conformational changes, protein binding and interactions with other nucleic acids, with a wide range of physiological responses. This response can include modulation of transcription, translation and mRNA stability, but also DNA maintenance and silencing to regulate gene expression[8]. For the control of gene expression, especially noncoding RNA segments are important[9]. One example for such regulatory, noncoding RNA in bacteria is small RNA (sRNA). sRNA has the ability to bind the complementary section of a mRNA and controls its expression by transcription termination, translation inhibition or mRNA cleavage[8]. Thereby, sRNA functions as regulatory RNA that acts in *trans*, meaning that it is a lone-standing RNA that acts on another RNA molecule. *Cis* acting regulatory RNAs on the other hand form secondary and tertiary structures to regulate gene expression of a gene encoded on the same molecule. Riboswitches are such *cis* acting regulatory RNAs that are found in noncoding regions and control gene expression in respond to binding of a certain ligand[7].

3.2. Riboswitches

In order to maintain cellular homeostasis, bacteria have to respond to changes in environmental or cellular conditions by adapting gene expression. For a long time, this was thought to be exclusively controlled by metabolite sensing protein factors. However, in 2002, riboswitches were discovered and broadened the understanding of RNA-based metabolic regulation in prokaryotes[10–13]. Riboswitches are complex structured, non-coding regions in the 5´-UTR of mRNA that enable regulation of the downstream gene[7,14–16]. They are able to sense metabolites or ions to control gene expression and thereby maintain cellular homeostasis of the cognate ligand[17]. The first discovered riboswitches were found to bind the coenzymes thiamine triphosphate (TPP) and flavinmononucleotid (FMN). During the next decade, the list of coenzymes bound by riboswitches has been further expanded, for example by S-Adenosyl methionine (SAM)[18,19] and tetrahydrofolate (THF)[20]. However, riboswitches have also been identified to bind nucleotide derivatives like purines[21] and signaling molecules such as ZTP[22], ppGpp[23] and c-di-AMP[24]. Interestingly, riboswitches were also found to bind and regulate gene expression in response to ions, like the fluoride riboswitch[25] or the Ni^{2+}/Co^{2+} riboswitch[26].

The mechanism of action is based on the formation of alternative structures to control transcription elongation and translation initiation depending on the metabolic status of the cell[27,28]. Almost all characterized riboswitches act in *cis* with only one known exception for a S-adenosylmethionine (SAM) sensitive system in *Listeria monocytogenes*[29]. Riboswitches are composed of two functional components: an aptamer domain and an expression platform[30]. The aptamer domain can sense a

certain ligand and upon binding induces structural rearrangement of the expression platform is induced. This conformational change either represses (OFF-switches) or activates (ON-switches) gene expression. Well-known examples of riboswitch-controlled metabolic pathways are the vitamin, amino acid and nucleotide metabolisms[30]. Additionally, riboswitches have also been found to regulate motility[24], biofilm formation[31] and virulence[29].

To date, more than 40 riboswitches have been experimentally validated, distributed among the bacterial divisions. Some representatives like the TPP, AdoCbl and SAM-I riboswitches are present in almost all phyla, others, on the contrary, are only limitedly distributed. Moreover, especially Firmicutes and Proteobacteria exhibit a high density of known riboswitches[32]. Only two riboswitch classes have been shown to be also represented in other domains of life. Fluoride riboswitches[25] and thiamine pyrophosphare (TPP) riboswitches[10] are also found in archaea. The TPP riboswitch is even present in several eukaryotic species of fungi, plants and algae[33–36].

3.2.1. Riboswitch structure formation and ligand binding

The high selectivity of riboswitches for their cognate ligands providing their ability to discriminate against other metabolites derives from their complex structural features. Especially aptamer domains are highly conserved ensuring the exceptional selectivity of these recognition sites[28]. Aptamer domains form specific interactions and shape a binding pocket, perfectly fitting for the cognate ligand[37]. It has been shown for many characterized riboswitches, that a point mutation of these highly conserved nucleotides in the aptamer region, can lead to a diminished or fully eliminated binding affinity. That proves that these nucleotides are essential for the formation of a specific binding pocket and are likely to interact with the cognate ligand. Ligands are recognized by electrostatic, hydrophobic, π-π and π-cation interactions and formation of hydrogen bonds. Although sensing domains of different riboswitch subclasses can vary greatly in size and organization of their secondary and tertiary structure, most of them exhibit multihelical junctions and pseudoknots[38]. The stability of these junctions and helices is directly affected upon ligand binding. Dependent on the abundancy of the ligand in the cell, riboswitches adopt two specified conformations: the ligand-bound fold and ligand-free fold, involving alternative base-pairing. Metabolite binding leads to a stabilization of the ligand-bound fold and thus to a stabilization or disruption of the regulatory hairpin within the expression platform. Moreover, molecular ligand binding processes can be divided into two mechanisms; the induced fit and the conformational selection, also named “binding-first” and “folding-first”, respectively[39,40]. Metal cations like magnesium or potassium support ligand binding by shifting the binding mechanism from induced fit to conformational selection[41,42]. In detail, physiological concentrations of Mg^{2+} govern the aptamer folding towards its ligand-bound state to support folded-like conformations of the RNA, that is then

further stabilized by ligand binding[43]. Moreover, cations can also be important to compensate negatively charged ligands like fluoride[25].

The ligand sensing aptamer domains show binding affinities in a range from nM to low µM. These affinities can decrease from 10- to 100-fold on proceeding transcription from the aptamer domain alone to the whole riboswitch. When testing "full-length" RNA constructs, carrying aptamer and expression platform of a riboswitch, often *in vitro* binding to the ligand cannot be observed. However, when only the aptamer domain is used, high binding affinities can be determined[44–48]. That is proof that ligand binding to a riboswitch motif is a cotranscriptional process with sequential folding [49–52]. It has also been shown that the polymerase slows down at regulatory pause sites to allow the formation of aptamer structures and ligand binding[53]. This demonstrates that riboswitch folding is regulated dynamically with the influence of transcription speed, transcriptional pausing, kinetics and thermodynamics of RNA structure formation and ligand binding[49,54]. Riboswitches therewith regulate gene expression in both a temporal and a spatial manner.

3.2.2. Gene control mechanisms

Ligand-dependent aptamer stabilization during co-transcriptional RNA folding is crucial for the function of riboswitches since ligand binding occurs within a limited time window before transcription of the expression platform[55,44]. Gene expression control can take place on the level of transcription[13,18] or on the level of translation[12,56] of metabolite transport and biosynthetic genes. In controlling translation, riboswitch structures either mask and prevent translation or release the ribosomal binding site (RBS) by formation of an antisequestor, allowing ribosomal attachment and thus protein biosynthesis[57,58].

For riboswitches controlling transcription there are two pathways how transcription termination can be induced: the intrinsic Rho-independent termination and the factor-depending termination, for example depending on helicase Rho[59,60] (Figure 1). Most reported riboswitches controlling on the level of transcription, fold into structures that serve as Rho-independent terminator and promote intrinsic transcription termination. But also a mechanism of riboswitch action has been reported to control gene expression based on regulation of the RNA-dependent helicase Rho to terminate transcription[61]. The hexamer Rho binds the primary binding site, the C-rich, G-poor *rut*-sequence (*rho utilization site*). Conformational change of Rho from an open-ring conformation to a locked-ring conformation triggers Rho's ATP-depending translocase activity. Rho moves along the transcript in 5' to 3' direction thus translocating the nascent RNA. When Rho encounters the RNA polymerase, the transcription elongation complex (TEC) is destabilized leading to its dissociation and transcription termination. In

case of transcription regulation based on a Rho independent terminator, the formation of a G-rich terminator-stem-loop followed by a short poly-U sequence is responsible for transcription termination (Figure 1). The hairpin causes irreversibly inactivation of TEC at the termination point (usually at U7 and U8 positions) and destabilizes the TEC[62–64]. Destabilization of the TEC and its dissociation induces the transcription termination.

The glmS riboswitch is the only known riboswitch to control gene expression using a different mechanism. The ligand Glucosamine-6-phosphate (GlcN6P) binds to the aptamer and is directly involved in a self-cleavage reaction (lower right) that leads to degradation of the mRNA (Figure 1).

Coenzymes	Ions	Nucleotide derivatives	Signaling molecules	Amino Acids	Other Metabolites
AdoCbl, AqCbl, FMN, Moco/Tuco, SAM, SAH, THF, TPP, NAD+	F^-, Co^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+}	Adenine, ADP. dADP, CDP, dCDP, 2´-dG, Guanine, preQ1, Xanthine	c-di-AMP, c-AMP-GMP, c-di-GMP, ppGpp, ZTP	Glutamine, Glycine, Lysine	Azaaromatics, GlcN6P, Guanidine, PRPP

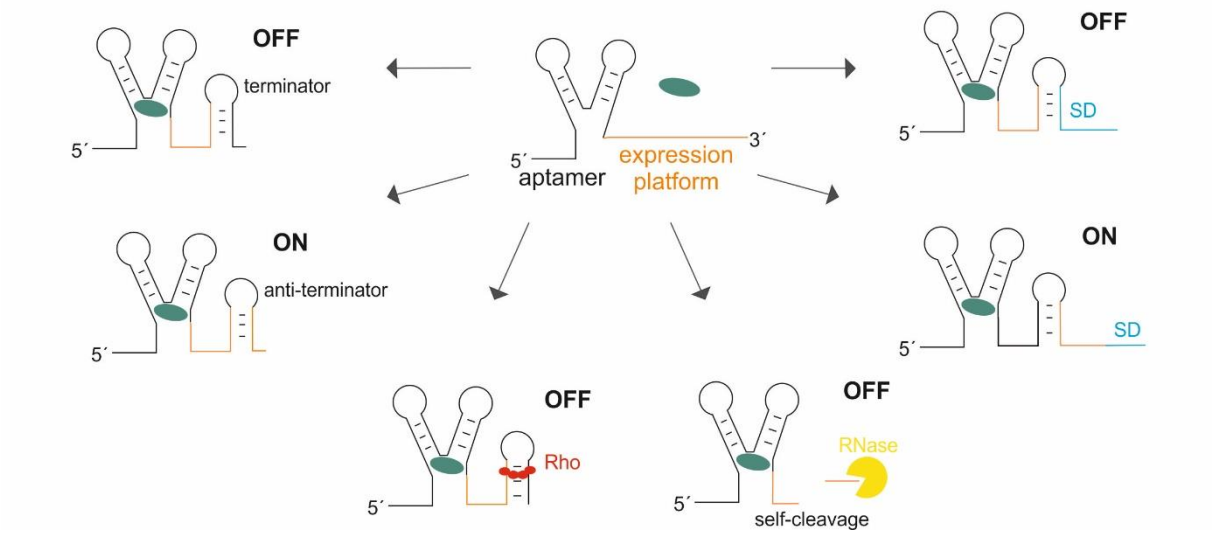


Figure 1: Known riboswitch ligands and gene control mechanisms. In the table on top, all riboswitch ligands identified until to date, are listed. Schematic shows the diversity of riboswitch-dependent gene control mechanisms. In the center the general structure of a riboswitch in an unfound state is represented, including the ligand binding aptamer and the expression platform. Depending on the riboswitch class, binding of the ligand (green) leads to different structural rearrangements that enable gene expression control on the transcriptional or translational level. On the translational level, upon ligand binding the Shine-Dalgarno sequence (SD, blue) can be sequestered (upper right) or masked (middle right). On the level of transcription, ligand binding can result in the formation of an intrinsic (upper left), or Rho-dependent terminator stem (middle left), preventing further transcription, or in the formation of an anti-terminator stem and destabilizing of the intrinsic terminator (lower left), allowing the polymerase to continue transcription. The glmS riboswitch uses a mechanism where the ligand Glucoseamine-6-phosphate (GlcN6P) is directly involved in a self-cleavage reaction (lower right) that leads to mRNA degradation. The figure was adapted and updated from Serganov and Nudler[7] and from Malte Sinn[65].

3.3.Riboswitch identification and characterization

Since riboswitches are unique in their ability to directly function as a sensitive and selective sensors for intracellular metabolites in a protein-independent manner, they can be used in various applications [66]. For example, they are a promising tool in gene therapy[67]. Indeed, aptamers have been already used as a drug in treatment of age-related macular-degradation[68,69]. Additionally, riboswitches control fundamental genes that for example play roles in metabolic pathways and transport in bacteria, which makes them potential targets for antimicrobial agents[70,71]. They can also be used for the development of compounds that artificially regulate gene expression and thus effect the bacterial growth[66]. On the other hand, riboswitches provide the opportunity for new laboratory tools for controllable expression of gene constructs. The collection of gene control system could be even more expanded by engineering of those natural riboswitches. Riboswitch-based expression systems are at an advantage over protein-based systems, since those are limited to the permeability of the regulatory protein to the cell and therewith dependent on the expression of the protein on a stable level. These systems can also be used as detecting tools, called biosensors, where binding of a specific ligand leads to a detectable signal production like fluorescence[66,72,73]. The discovery of additional riboswitch classes and their associated regulatory networks will also help to understand functions of associated genes and their encoded proteins[74]. Thus, it is important to expand the list of characterized riboswitches[18].

In order to identify novel riboswitches, bioinformatics tools are used. Using comparative sequence analysis algorithms, bacterial genomes are analyzed for elements that are characteristic for riboswitches. The common element in such approaches is that they analyze homologous intergenic regions for “covariation”, which are mutations that change the primary sequence but conserve an RNA secondary structure. Good riboswitch candidates show, among other characteristics, significant covariation and are consistently located upstream of protein-coding genes, which they are expected to regulate[75–79]. Besides, these riboswitch candidates whose ligands remain to be identified are called orphan riboswitches. In the past, most of the ligand identification for these orphan riboswitches were constructed from the genetic context. Educated guesses provided metabolites that were tested and experimentally verified to be sensed and bound by the particular riboswitch[80]. Nearly all characterized riboswitches have been verified and characterized using in-line probing.

3.3.1. in-line probing

The in-line probing method is based on the inherent chemical instability of RNA and its tendency to undergo a spontaneous cleavage of phosphodiester linkages[81,82]. The nucleophilic oxygen of the 2´-hydroxy group is in proximity to the phosphodiester linkage enabling intramolecular

transesterification reactions. The rate of transesterification can be increased by deprotonation of the 2´-hydroxy group. The resulting nucleophilic 2´-oxyanion group attacks the adjacent phosphorus center to form a pentacoordinated intermediate. Product of this S_N2-like reaction is a 2´3´-cyclic phosphate and a hydroxy group at the 5´ end (Figure 2A). Not only deprotonation of the hydroxy group favors transesterification, but also its position is crucial for the transesterification rate[83]. For a nucleophilic attack the 2´-hydroxy group position is optimal when it is found in an in-line conformation with the 5´ leaving group. Due to the conformation in an RNA A-form helix, the 2´-hydroxy group and the 5´ leaving group cannot adopt this in-line conformation. Thus, in a helical conformation 3´5´-phosphodiester linkages are stabilized in a helix. This explains increased instability of unstructured RNA regions compared to well folded regions[84]. The conformation-dependent rate of RNA transesterification provides the basis for the principle of in-line probing. Incubation of 5´-^{32}P-labeled RNA gives a unique cleavage pattern, showing more RNA cleavage within unstructured regions than in more structured regions. To analyze and confirm the ability of an aptamer to bind a certain metabolite, the ligand is included in the incubation. Aptamer and ligand form a complex stabilizing the ligand-bound structure changing conformation-dependent rate of RNA transesterification (Figure 2B). For example, using in-line probing to characterize the FMN riboswitch, the aptamer-ligand complex maintains RNA linkages due to a non-in-line conformation. FMN binding changes the conformation of the 3´5´ phosphodiester linkages from heterogeneous to homogenous conformation disfavoring RNA cleavage[11]. The shift in transesterification rates can be observed in the cleavage pattern[81].

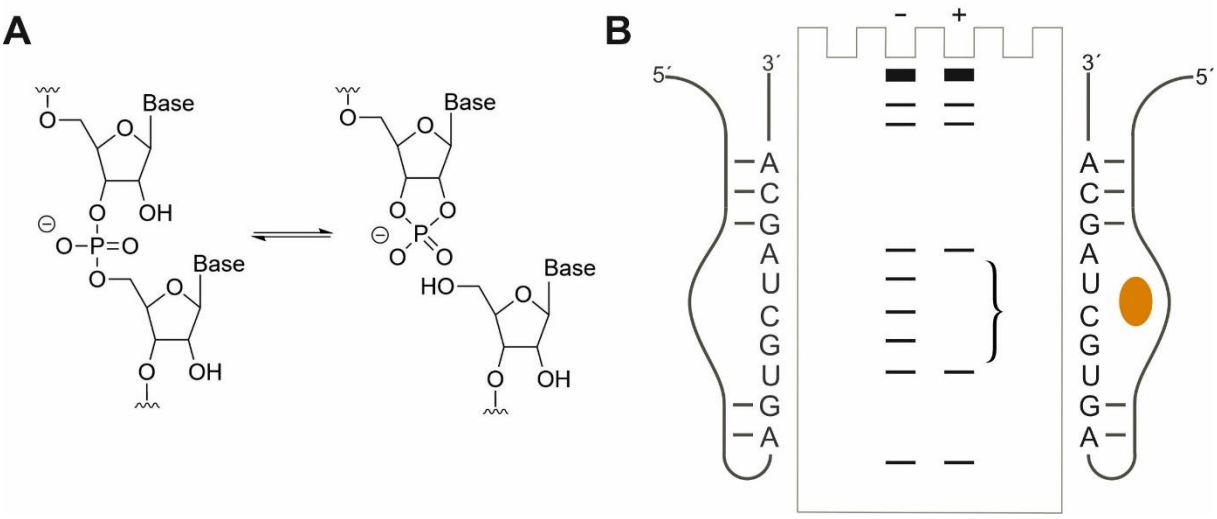


Figure 2: Basic mechanism of in-line probing reactions. (**A**) Spontaneous self-cleavage reaction of an RNA phosphodiester linkage. The nucleophilic oxygen of the 2´-hydroxy group attacks the phosphorus, cleaving the phosphodiester linkage and resulting in a 2´3´ cyclic phosphate and a hydroxy group at the 5´ end. In in-line probing reactions the rate of transesterification is increased by raising the pH value and thereby enabling deprotonation of the 2´-hydroxy group. (**B**) Scheme of an in-line probing gel. Binding of the ligand (+) (shown in green) to the RNA, changes the cleavage pattern in comparison to unbound RNA (-).

4. Chapter II: Identification of a new class of guanidine riboswitches

4.1.Introduction

4.1.1. Guanidine riboswitches

In 2004, the *ykkC-yxkD* RNA motif[75] was identified, but its ligand was unknown for well over a decade, due to challenges posed by the wide variety of associated genes and their unknown function. Being a common motif in various bacterial phyla, it is associated with genes encoding for multidrug efflux pumps and other transporters, urea carboxylases, purine and amino acid metabolism enzymes, among other gene products[75,80]. Many of these genes are found downstream of the two other riboswitch candidates mini-*ykkC*[76] and *ykkC*-III[78]. Their consensus sequences did not show structural similarity between each other, but the similar genetic contexts of all three motifs suggested the hypothesis that they sense the same ligand[76,78]. After many years of efforts, guanidine was eventually revealed as the cognate ligand of the three motifs, now renamed *guanidine-I*, *-II* and *-III* riboswitches (Figure 3)[85–87]. However, they differ in their sequence, structural features and mechanisms of action. The *guanidine-I* riboswitch class activates the expression of its downstream gene in the presence of guanidine by preventing the formation of an intrinsic terminator stem[85]. Whereas *guanidine-II* and *–III* riboswitches act as translational ON-switches[86,87]. In their ligand-unbound state, the Shine-Dalgarno (SD) sequence is masked. Upon guanidine binding, structural rearrangements lead to the liberation of the SD sequence, allowing the ribosome to bind and initiate protein translation.

At the time of discovery of guanidine-binding riboswitches and their widespread occurrence, guanidine was not known to play a role in biology. Rather it has been used as a propellant, an additive in plastics, as well as a chaotropic substance in protein biochemistry[88]. The existence of a riboswitch sensing guanidine was remarkable and suggested that it occurs naturally. The most widely associated genes *sugE* or *emrE,* were predicted to encode small multidrug resistance (SMR) efflux pumps[89,90]. For example, EmrE multidrug transporters in *Escherichia coli* (*E. coli*) have a wide spectrum of substrates and contribute to overcome toxicity of a variety of compounds[91]. The function, encoded by genes *sugE* or *emrE* associated with guanidine riboswitches, was subsequently demonstrated to export guanidine, termed Gdx[92]. Accordingly, guanidine riboswitches control genes whose protein products are crucial for overcoming its toxicity. Besides, earlier work established the existence of urea carboxylase enzymes, but riboswitch-associated genes that had been predicted to encode these enzymes have since been demonstrated to favor guanidine over urea[85]. Furthermore, it has been shown that this carboxylase pathway enables the utilization of guanidine as sole nitrogen source. Taken together, guanidine riboswitches predominantly induce Gdx transporters in order to export this compound from bacterial cells before it reaches toxic concentrations. Other gene functions controlled by guanidine

riboswitches enable the carboxylation and subsequent degradation. Since the three known guanidine riboswitches are found widespread in various bacterial phyla riboswitch classes, it has been speculated that further, structurally unrelated guanidine riboswitch classes could exist[93]. Apart from guanidine, four unrelated riboswitch classes are currently known that bind SAM[18,19], three for $preQ_1$[94] and two for cyclic di-GMP[32,95]. It seems reasonable to speculate that further structural classes will be found for guanidine, SAM, preQ1 and cyclic-di-GMP.

In order to identify additional classes of guanidine riboswitches, Zasha Weinberg exploited the gene contexts of known guanidine riboswitches in combination with a discovery strategy based on comparative genomics. The earliest application of a similar strategy was in the discovery of the SAM-III riboswitch (also called the S_{MK}-box riboswitch), which binds *S*-adenosylmethionine (SAM)[96]. This riboswitch was found because certain species conspicuously lacked examples of the then-known SAM riboswitch classes. Since the known SAM riboswitch classes often occur upstream of *metK* genes, encoding SAM synthetase, a manual analysis was conducted screening for conserved patterns in the 5´-UTR of these genes. Using a similar strategy, followed by detailed analysis of the computationally predicted alignments, Zasha Weinberg established six candidates of potential guanidine riboswitches. We call these Guanidine-Gene-Associated Motifs (*GGAM*)[97].

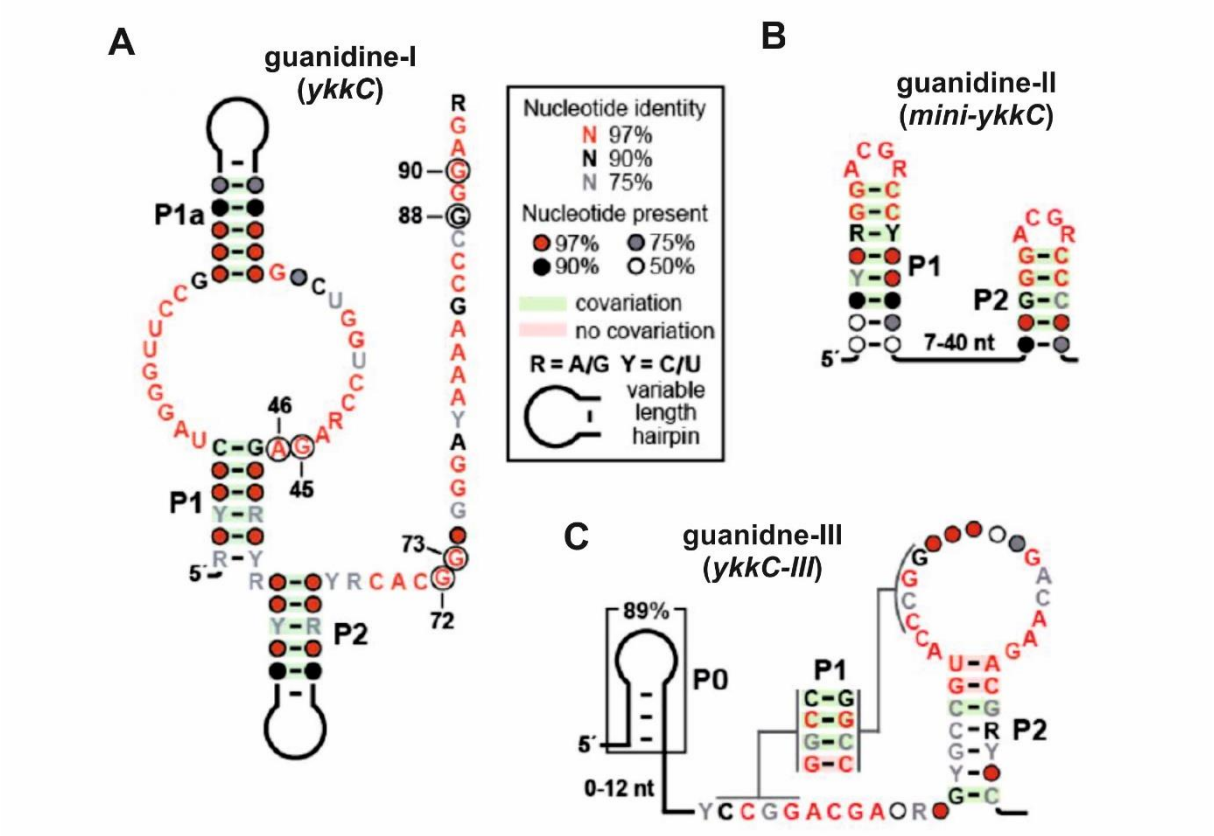


Figure 3: The family of guanidine riboswitches. Diagram showing conserved features of RNA sequences of (**A**) *guanidine-I*, (**B**) *guanidine-II*, and (**C**) *guanidine-III* riboswitches, determined by comparative sequence analysis. A legend (box) explains other symbols. Adopted from[98].

4.1.2. The *GGAM-1 motif*

After the identification of six guanidine riboswitch candidates (Figure 4A), the *GGAM-1* motif was of particular interest, because it exhibits properties that are expected of riboswitches. First, it has several highly conserved nucleotides. Moreover, the motif includes sequences present in multiple phyla. Most often it occurs in the phylum Firmicutes, and is also present in species from six other phyla. Since nucleotides are highly conserved, despite the RNAs being highly diverged across phyla, the RNA appears to be subject to strong biochemical constraints, which is expected of an RNA that specifically binds a small molecule. Second, the *GGAM-1* includes a potential pseudoknot which are often associated with riboswitches. Third, *GGAM-1* RNAs consistently occur upstream of protein-coding genes, and they encode multiple non-homologous proteins. This observation is strongly consistent with a *cis*-regulatory function. Finally, the *GGAM-1* motif's structure contains a predicted Rho-independent transcription terminator. Such terminators consist of a hairpin followed by a poly-U signal, and can cause inhibition of the transcription process. They are a common expression platform in riboswitches and also found to be associated with the *guanidine-I* riboswitch[85].

Besides. *GGAM-1* motif RNAs are most commonly found upstream of *sugE* genes, as expected for a guanidine riboswitch. However, the motif is also found to be associated with several other gene classes, some of which are rarely or never associated with previously established guanidine riboswitches (Figure 4B). Like the *sugE* gene, these new gene associations could suggest additional genes with a guanidine-related function. Here, we demonstrate that the riboswitch candidate *GGAM-1* represents a fourth class of guanidine riboswitches acting via transcription termination control[97,48].

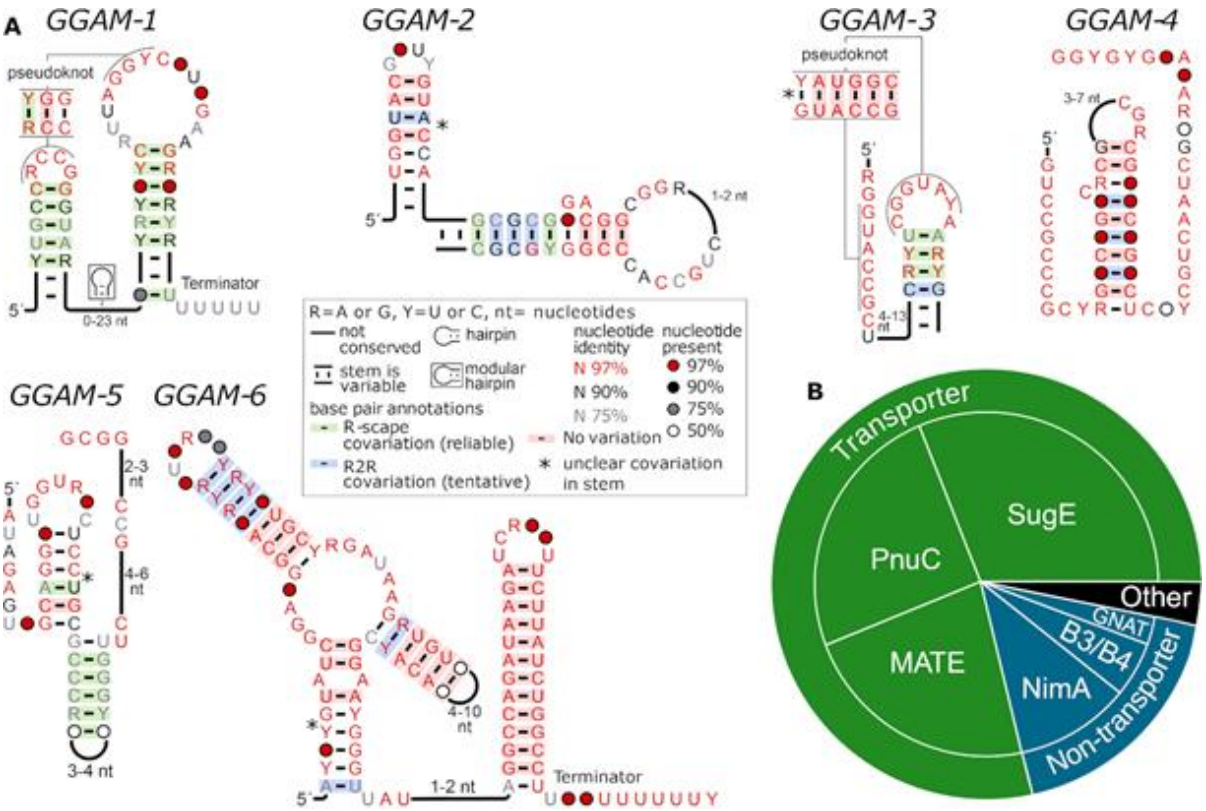


Figure 4: Candidate guanidine riboswitches: the *GGAM* RNA motifs. (**A**) Diagram showing conserved features of RNA sequences in *GGAM-1* to *GGAM-6*. 'Terminator': this stem exhibits the properties typical of Rho-independent transcription terminator hairpins. A legend (gray box) explains other symbols. (**B**) Genes frequently associated with *GGAM-1* RNAs. The six conserved protein domains most commonly encoded by genes that are immediately downstream of *GGAM-1* RNAs. Such genes are almost certainly regulated by the riboswitch, in comparison to genes that might be located in extended operons. Three domains function as transporters (green), while the other three do not (blue). Less common domains and domains that did not match the Conserved Domain Database were classified as 'other' (black).

4.2.Results & Discussion

4.2.1. Guanidine binds selectively to the GGAM-1 motif

Initial experiments showed that the motifs *GGAM-2* to *-6* did not exhibit binding to guanidine assayed by in-line and *in vitro* transcription experiments (data not shown; for sequences tested see Supplementary Table 2), and they were not further pursued.

Due to its properties, like its widespread occurrence, the presence of highly conserved nucleotide and a potential pseudoknot, the *GGAM-1* motif is likely a good riboswitch candidate. Based on its genetic context, we hypothesized that it to function as a guanidine sensing riboswitch. In subsequent sections, we demonstrate that the *GGAM-1* motif corresponds to a class of guanidine riboswitches. A 95-nucleotide-long RNA construct (*95 Lla*) (Figure 5A) from the 5'-UTR of the *sugE* gene of *Lactococcus lactis* was investigated in an in-line probing reaction. This method is often used to characterize riboswitches and their interaction with certain molecules. Indeed, guanidine causes a concentration-dependent structural modulation of the *95 Lla* construct (Figure 5B,C). Especially, the band intensity of the nucleotides within the two loop regions increased in the presence of guanidine. These results suggest that the loop regions are somehow involved in direct ligand binding or the formation of the binding pocket. By quantifying the extent of spontaneous cleavage at nucleotide position G 62 over increasing concentrations of guanidine hydrochloride, a mean apparent K_D value of 210 µM (+/− 20 µM) was determined (Figure 5C). To verify that the *GGAM-1* sequence from *L. lactis* is not an exceptional case and that all representatives identified for the *GGAM-1* motif bind the same ligand, additional motif sequences from another organism were characterized. In detail, the *GGAM-1* motif sequences found in *Raoultibacter timonensis*, *Brachyspira alvinipulli* and *Ruminococcus albus* (Supplementary Table 1) were tested for their ability to bind guanidine. Again, these constructs showed a concentration-dependent modulation of the cleavage pattern in response to guanidine (Supplementary Figure 1, Supplementary Figure 2). Specifically, a 92-nucleotide-long RNA (*92 Rti*) from the 5'-UTR of the *emrE* (i.e. *sugE*) gene of *R. timonensis* was used. Here a mean apparent K_D of 190 µM (+/− 30 µM) was determined (Supplementary Figure 1).

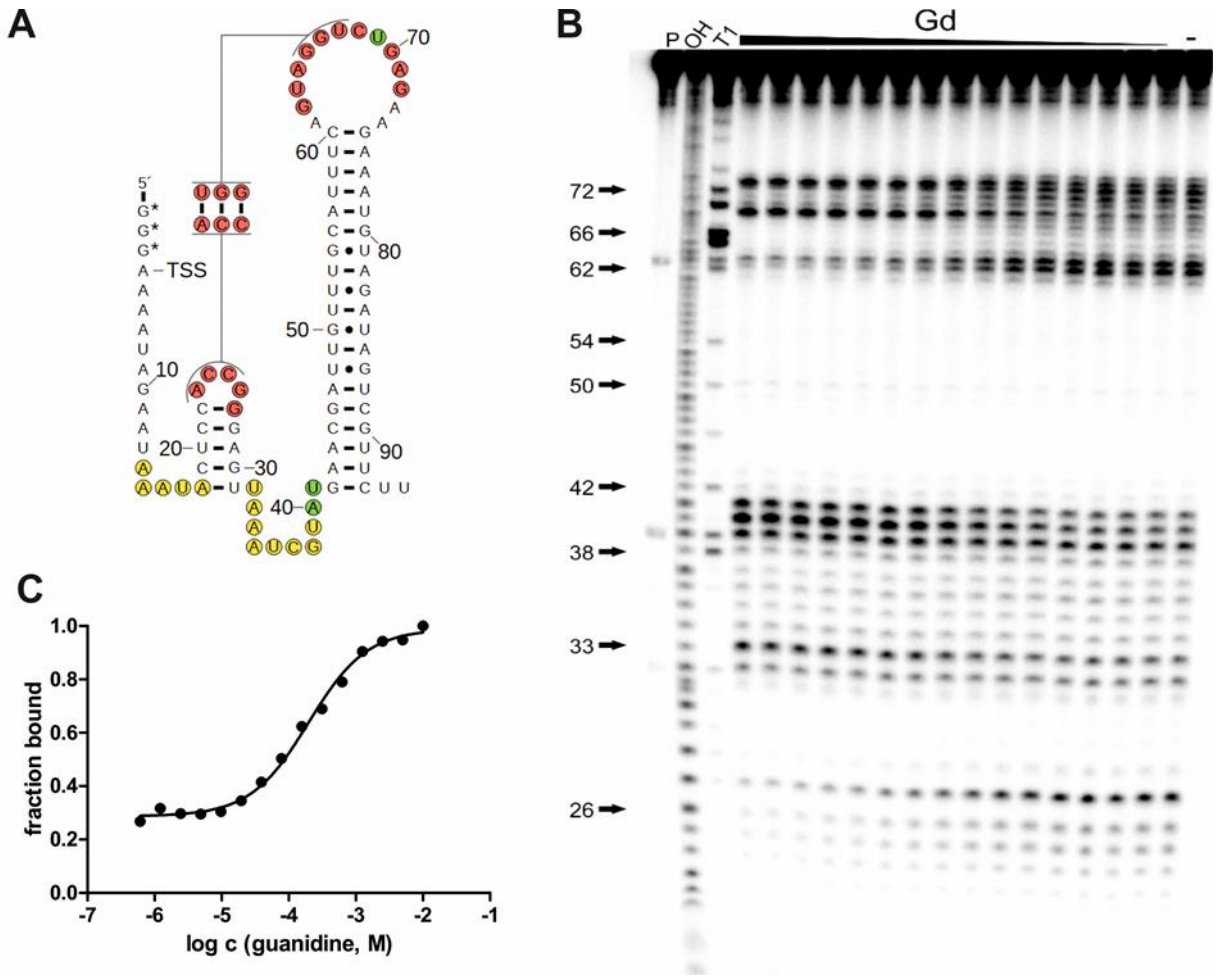


Figure 5: The *GGAM-1* RNA motif binds guanidine. (**A**) Sequence and secondary structure of *95 Lla* RNA construct from the 5'-UTR of the *sugE* gene of *L. lactis*. The 5' terminus of the construct includes three non-genomic guanosine nucleotides, indicated with asterisks, to improve *in vitro* transcription efficiency. Nucleotides are numbered from the 5' end. They are colored according to the changes in spontaneous cleavage rates upon addition of guanidine. Red: cleavage rate decreases with guanidine. Green: cleavage increases. Yellow: cleavage is high with and without guanidine. No color: cleavage is low in both conditions. (**B**) PAGE analysis of an in-line probing reaction of 5'-^{32}P-labeled *95 Lla* RNA without (−) or with guanidine hydrochloride in a range of 0.61 µM – 10 mM. P, OH and T1 represent 5'-^{32}P-labeled RNA undergoing no reaction, digest under alkaline conditions, or digest with RNase T1, respectively. Selected nucleotide numbers from part A are indicated. (**C**) Plot of the fraction of RNA bound to ligand as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined based on quantification of band intensity changes at U 63 and normalized by the intensity of the constant band U 33. A trendline was generated using a sigmoidal dose-response curve fit (maximum value equal to 1) to determine an apparent K_D value. A mean apparent K_D of 210 µM (±20 µM) was determined in three independent experiments (Supplementary Figure 3A,B).

Highly conserved nucleotides in the aptamer domain are often found to be directly involved in ligand binding and are crucial for the high selectivity of the binding domain. It has been already shown for many riboswitches, that pint mutations of these nucleotides lead to a decreased ligand binding affinity or even to a complete loss of the binding function of the aptamer. To further characterize the G*GAM-1* motif RNA and show that its binding pocket selectively binds guanidine we aimed for identification of nucleotides being essential for ligand binding. As the band intensity of nucleotides in both loops

decreases in the cleavage pattern when guanidine is present, we stated that they are involved in binding pocket formation. Since guanidine is a small molecule, it seems reasonable to speculate that both loops come together to form a binding pocket. Consequently, different mutant constructs of the *95 Lla* RNA were tested in in-line probing reactions (Figure 6A). The mutant constructs M1, M2 and M3, each carrying a point mutation at a highly conserved position (97% nucleotide identity) in the second loop, completely eliminated guanidine-dependent modulation (Figure 6B). The construct M4 carries a mutation at a less conserved nucleotide position (90% nucleotide identity) but still shows a greatly diminished structural modulation. However, the folding of this construct differs from the wildtype (wt) motif. These results demonstrate that binding of guanidine to the *GGAM-1* motif RNA is dependent on the presence of the highly conserved nucleotides in the loop regions that likely form a selective binding pocket. Additionally, mutant constructs carrying nucleotide changes in the first stem of the *GGAM-1* motif were tested (Supplementary Figure 4). Here, these mutations only lead to a small diminution of the structural change induced by guanidine. The highly conserved nucleotides C22 and G27 form a base pair on top of the first stem. In order to analyze the importance of these two nucleotides, the construct M5 was designed, carrying the two point mutations C22G and G27C. For this construct a structural change in the presence of guanidine, similar to the effect with the wt construct, was observed. This indicates, that both nucleotides are not directly involved in the binding of guanidine. However, the presence of the CG base pair might be important to maintain the stability of the first stem, which is essential for the formation of the binding pocket. Taken together, these mutation studies show that both loop regions, carrying highly conserved nucleotides, are involved in the selective binding of guanidine. Especially the positions G66, U67 and G72 are essential to sense guanidine as the absence of each of them completely eliminates guanidine binding.

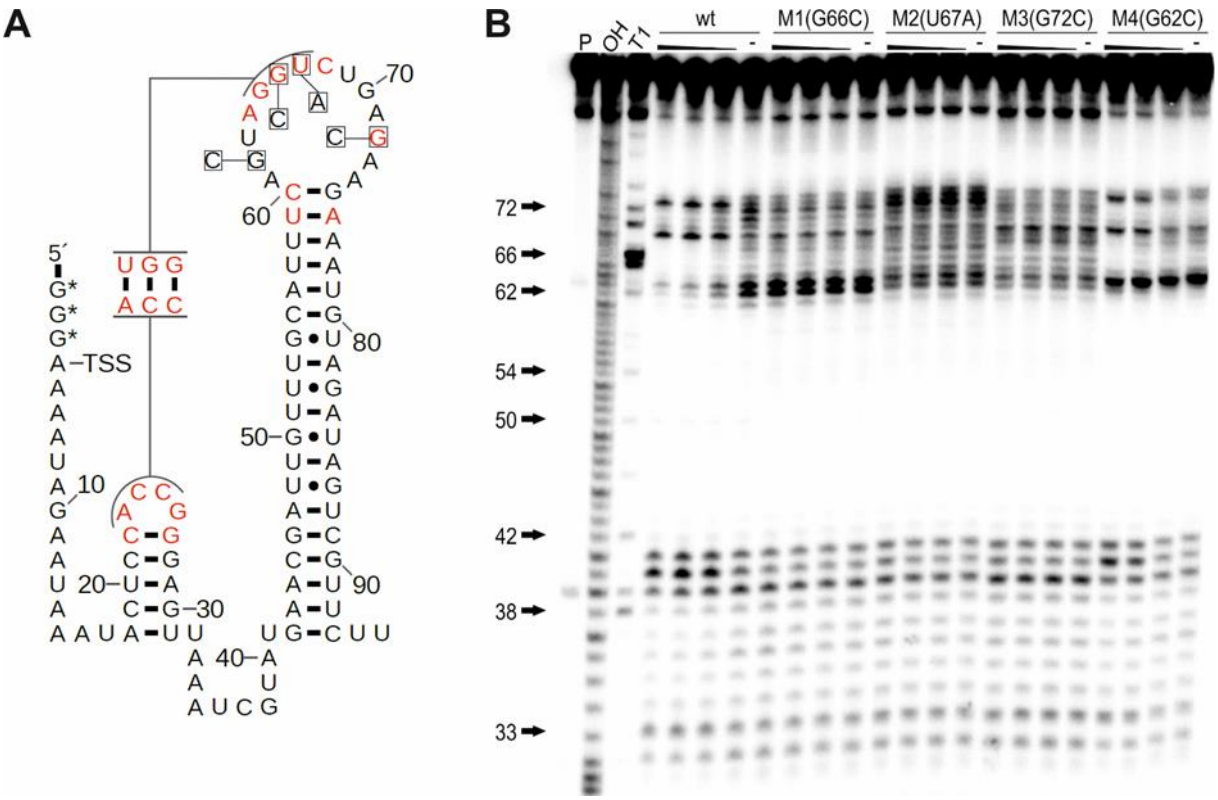


Figure 6: Point mutations of conserved nucleotides compromise guanidine binding. (**A**) Sequence and secondary structure of *95 Lla* RNA with the location of mutations in constructs M1(G66C), M2(U67A), M3(G72C) and M4(G62C). Highly conserved positions (97% nucleotide identity) are shown in red. (**B**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *95 Lla* wt RNA and mutants M1(G66C), M2(U67A), M3(G72C) and M4(G62C) without (−) or with guanidine hydrochloride with concentrations of 10 mM, 1 mM and 100 µM. P, OH and T1 represent 5′-^{32}P-labeled RNA undergoing no reaction, digest under alkaline conditions, or digest with RNase T1, respectively.

4.2.2. *GGAM-1* controls transcription termination in response to guanidine

We were able to show that the *GGAM-1* motif selectively binds guanidine and undergoes a structural modulation in a guanidine dose dependent manner. Based on this structural modulation and on the mutation studies, we hypothesized that both loop regions are involved in ligand binding and come together to form of a binding pocket. However, the mechanism of gene expression regulation remained unknown. In general, riboswitches can control the expression of downstream genes either on the level of transcription or translation. Since almost all examples of the *GGAM-1* motif RNA are found to be associated with a Rho-independent transcription terminator, it seems reasonable that the guanidine binding observed in in-line probing assays would result in a modulation of transcription termination. To test this assumption, the transcription of a DNA template for a 147-nucleotide-long RNA construct (*147 Lla*) from *L. lactis* was monitored. This RNA construct carries the *GGAM-1* motif RNA and the following sequence context, including an intrinsic terminator stem sequence followed by 7 U residues, the start codon and 31 nucleotides of the *sugE* open reading frame. Using an *in vitro*

transcription assay, the DNA template was transcribed with *E. coli* RNA polymerase in the presence or absence of guanidine hydrochloride. In accordance with our hypothesis, the yield of detected full-length transcription product increases in a concentration-dependent manner, whereas the termination product decreases in response to guanidine hydrochloride (Figure 7A), with a half maximal effective concentration EC_{50} of 260 µM (±30 µM) (Figure 7B). This is consistent with the proposed riboswitch mechanism, in which binding of guanidine stabilizes a structure that prevents transcription termination, whereas the structure of the non-bound RNA enables formation of the intrinsic terminator and thus promotes transcription termination. These data show that the *GGAM-1* motif acts as an ON-switch. Ligand binding allows the polymerase to continue and transcribe a full-length transcription product of the downstream gene to activate gene expression.

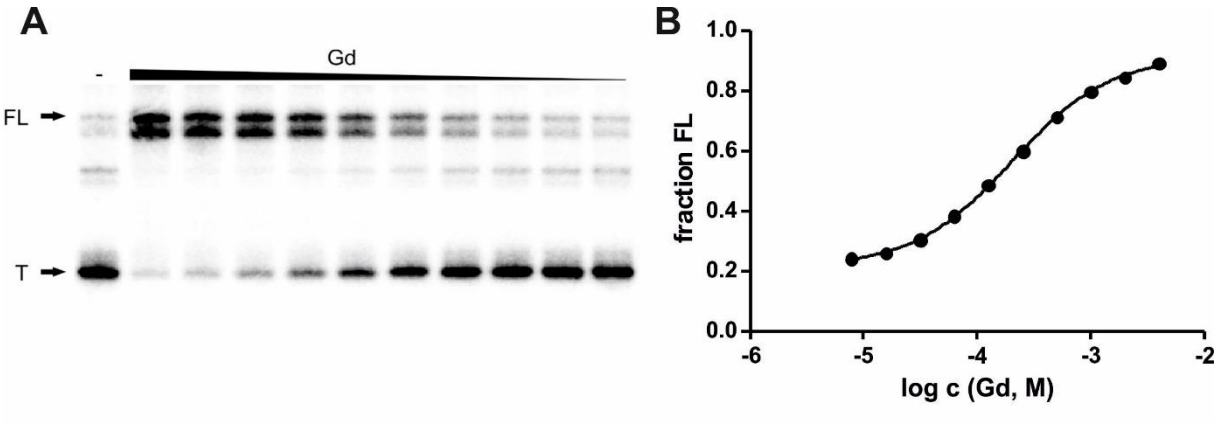


Figure 7: The *GGAM-1* motif RNA controls gene expression via transcription. (**A**) PAGE analysis of a transcription termination assay of *147 Lla* RNA without (−) or with guanidine hydrochloride ranging from 7.8 µM - 4 mM. FL and T denote full length product at 147 nucleotides and termination product at 93 nucleotides, respectively. (**B**) Plot of the fraction of full length *147 Lla* product relative to the total number of transcripts (FL plus T) as a function of the guanidine hydrochloride concentration. A mean EC50 of 260 µM (±30 µM) was determined applying a sigmoidal dose-response curve fit in three independent experiments (Supplementary Figure 4A,B).

4.2.3. *GGAM-1* regulates *in vivo* gene expression

We demonstrated that the *GGAM-1* motif RNA specifically binds guanidine, causing a concentration-dependent structural modulation to prevent transcription termination, leading to an increased full-length transcription product. With these results, we show that the *GGAM-1* motif is a riboswitch, promoting expression of the associated downstream gene when guanidine is present. To fully characterize the motif and show that it functions as riboswitch in living cells, the ability of guanidine to regulate gene expression *in vivo* was assessed. Therefore, a reporter plasmid carrying the *GGAM-1* motif of *L. lactis* in the 5'-UTR of an *egfp* reporter gene was designed to conditionally control gene expression in *S. aureus* RN4220 strain. Assuming that the *GGAM-1* RNA motif is an ON-switch, *egfp* should not be expressed in the absence of guanidine, since gene expression is inhibited by formation of an intrinsic terminator. Due to guanidine addition, the expression of *egfp* should be increased, as transcription termination is prevented and the full-length transcription product favored. The reporter strain was grown in Brain Heart Infusion (BHI) medium in the presence or absence of guanidine hydrochloride. To monitor reporter gene expression, the fluorescence intensity was measured and normalized by the optical density (OD_{600}). As expected, very low eGFP fluorescence was detected when guanidine was not present in the medium. Addition of 5 mM guanidine hydrochloride resulted in an 80-fold increase in fluorescence intensity (Figure 8A). Varying the amount of added guanidine showed a concentration-dependent change of gene expression (Figure 8B). To exclude the possibility that an unspecific effect of guanidine to the cells changes expression of the *egfp* gene, we used a plasmid that lacks the *GGAM-1* motif, but still carrying the reporter gene, as control. Here, we did not observe any influence of guanidine on eGFP fluorescence intensity. Our *in vitro* binding assay results revealed importance of the high conserved nucleotides for the selective binding of guanidine. This was confirmed *in vivo*. Plasmid were used carrying point mutations at positions that have also been investigated in binding assays. Consistent with the in-line probing results, where the point mutations M1, M2 and M3 fully eliminated guanidine binging (Figure 6B), no change in fluorescence intensity for these mutants in response to guanidine was observed *in vivo* (Figure 8A). These data are consistent with our previous results showing the necessity of these nucleotides for binding of guanidine and the switching activity of the *GGAM-1* motif. *In vitro*, we observed that the M4 mutation does not completely eliminate binding to guanidine, but does reduce binding affinity. *In vivo*, the M4 mutation causes a 20-fold lower fluorescence intensity compared to the wt sequence (Figure 8A).

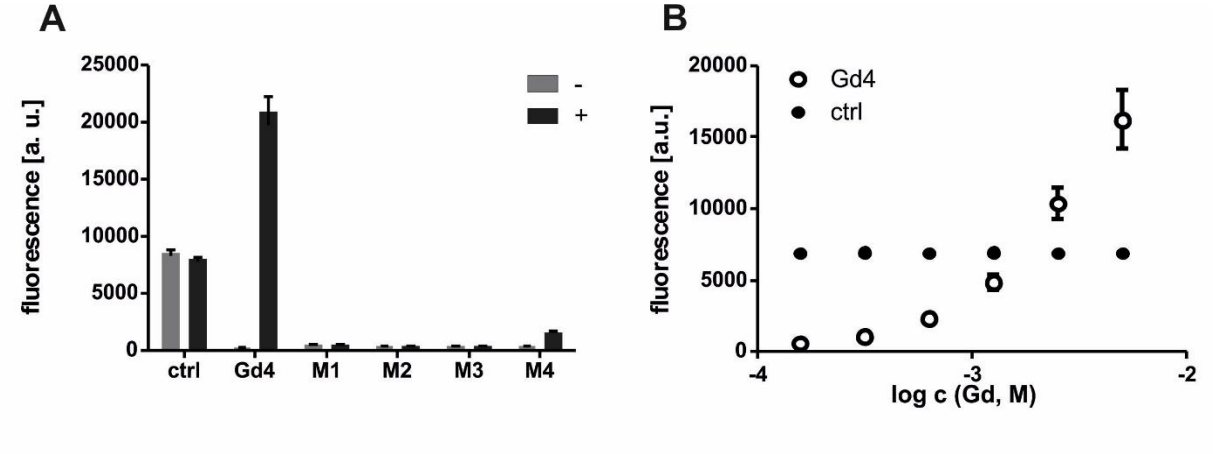


Figure 8: *GGAM-1* controls gene expression of the downstream gene *in vivo*. (**A**) Bars show the eGFP fluorescence intensity in *S. aureus* without (−) and with (+) 5 mM guanidine hydrochloride, normalized to the OD_{600}. Error bars represent the standard deviation from three independent experiments. A construct lacking the *GGAM-1* motif, but constantly expressing *egfp* (ctrl) served as a control. M1, M2, M3 and M4 represent constructs carrying point mutations at positions hypothesized to be important for the formation of a selective binding pocket. Location of these point mutations are shown in Figure 6A. (**B**) Plot of dose-dependent eGFP fluorescence intensity in *S. aureus*. Strains were grown in media containing guanidine hydrochloride in the range 156 µM–5 mM. Fluorescence was measured and normalized to OD_{600}. Error bars represent the standard deviation from three independent experiments. A construct with constitutively active *egfp* (ctrl) served as control.

For the previously identified guanidine riboswitch classes, it has been shown that they strongly discriminate against similar compounds. For the *guanidine-I* riboswitch affinities for guanidine analogs that differ 100-fold compared to the parent ligand were determined[85]. Although, the *guanidine-II* and *-III* riboswitches showed a less decline in affinity for analogs with small addition to the guanidine moiety core, far poorer affinities for analogs with larger substitutes were observed[87,86]. This suggests that these motifs use steric effects to exclude larger compounds.

To examine the ligand-binding selectivity of the *GGAM-1* motif and its ability to proficient discriminate against small molecules similar to the natural ligand, guanidine analogs were tested in the binding assay as well as in the transcription termination assay. In both assays, only methyl-guanidine and amino-guanidine, both of which carry only small substitutes to the guanidine moiety core, were observed to bind and regulate transcription of the *GGAM-1* motif RNA, respectively (Figure 9). In in-line probing reactions, they induced structural modulation of the *95 Lla* RNA similar to that induced by guanidine hydrochloride (Figure 9A). However, the affinities to bind these analogs is 20- to 40-fold poorer than with the parent compound. Methyl- and amino-guanidine show an increased (poorer) apparent K_D of 4.1 mM (±0.5 mM) and 7.5 mM (±0.8 mM), respectively (Figure 9C and Figure 10). In transcription termination assays, addition of both analogs led to an increase in full-length product in a concentration-dependent manner (Figure 9D). Both compounds showed a 10-fold poorer effect on the

in vitro transcription compared to the natural ligand guanidine. For methyl-guanidine and amino-guanidine EC50 of 2.3 and 3.5 mM was determined, respectively (Figure 11).

In contrast, with urea and arginine, no binding to the RNA in in-line probing reactions was observed (Figure 9A). Also, no regulation of transcription termination was observed (Figure 9D). It seems likely that the aptamer binding pocket sterically excludes larger compounds. Moreover, urea is a relatively small molecule that carries an oxo group instead of the imine nitrogen atom of guanidine. Thus, replacing a hydrogen bond donator by a hydrogen bond acceptor. Additionally, urea is neutral, whereas guanidine is positively charged under physiological pH conditions. These differences might cause exclusion of urea. Concluding, our data indicate that the *GGAM-1* motif binds guanidine with high selectivity and that the binding pocket proficient discriminates against analogs with small and larger substitutions or an oxo group such as in urea. Additionally, it has been previously shown that guanidine does not bind riboswitch classes that are already known to bind other molecules[85], providing another reason to assume that the guanidine binding of the *GGAM-1* motif is specific to the properties of this RNA.

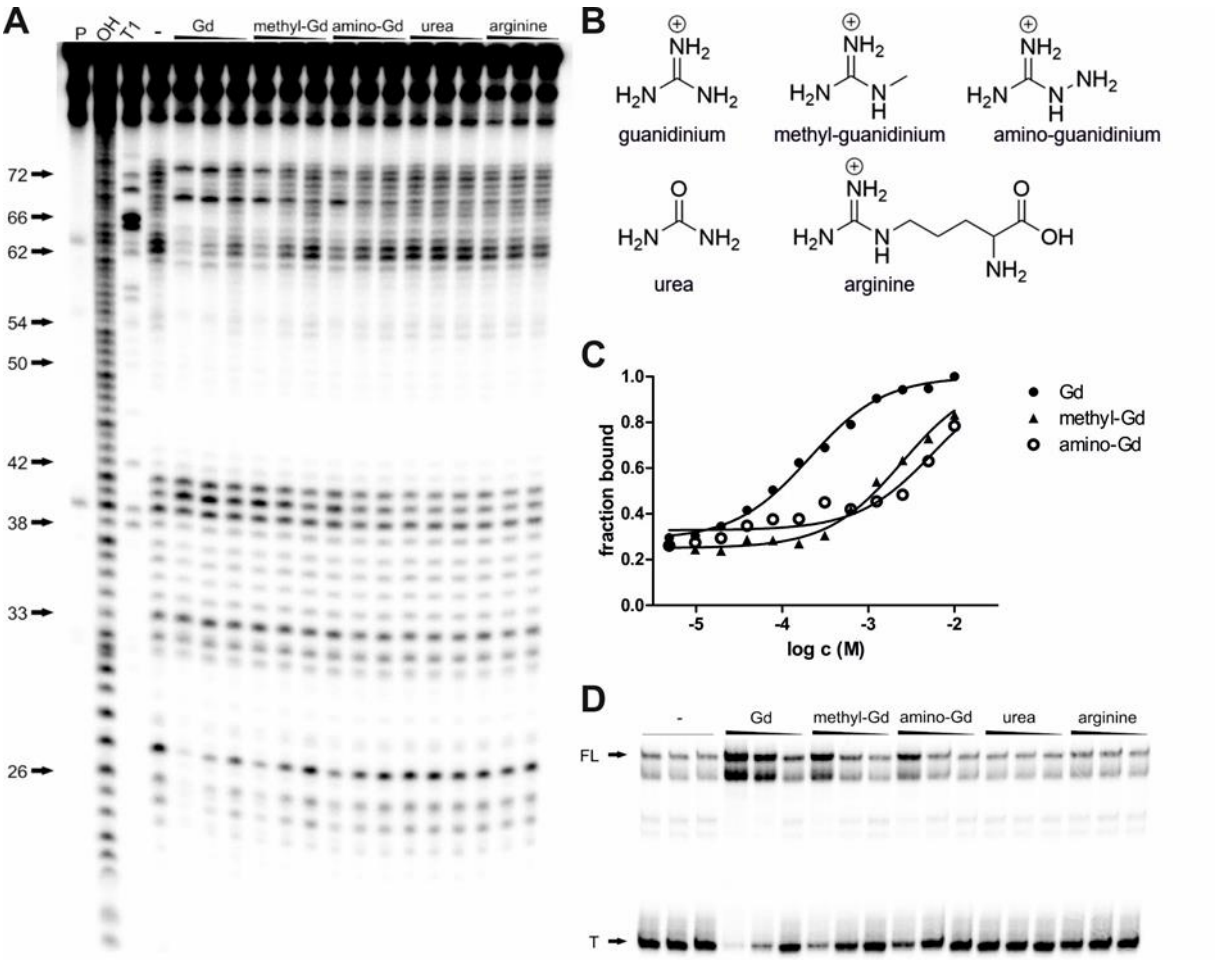


Figure 9: Binding and transcription control by guanidine derivatives. (**A**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *95 Lla* RNA without (−) or with guanidine hydrochloride, methyl-guanidine hydrochloride, amino-guanidine hydrochloride, urea or arginine with concentrations of 10 mM, 1 mM and 100 µM. P, OH and T1 have the same meaning as in Figure 5B. (**B**) Chemical structures of guanidine, methyl-guanidine, amino-guanidine, urea and arginine. (**C**) Plot of the fraction of RNA bound to methyl-guanidine and amino guanidine, respectively, as a function of the logarithm (base 10) of the molar concentration. Fraction of RNA bound was determined as in Figure 5C. The corresponding PAGE analysis of in-line probing reactions with methyl-guanidine and amino-guanidine are shown in Figure 10. For methyl-guanidine and amino-guanidine, a mean apparent K_D of 4.1 mM (±0.6 mM) and 7.5 mM (±0.8 mM), respectively, was determined in three independent experiments (Supplementary Figure 3C-F). (**D**) PAGE analysis of a transcription termination assay of *147 Lla* RNA without (−) or with guanidine, methyl-guanidine, amino-guanidine, urea or arginine with concentrations of 10 mM, 1 mM and 100 µM. FL and T denote full length product at 147 nucleotides and termination product at 93 nucleotides, respectively.

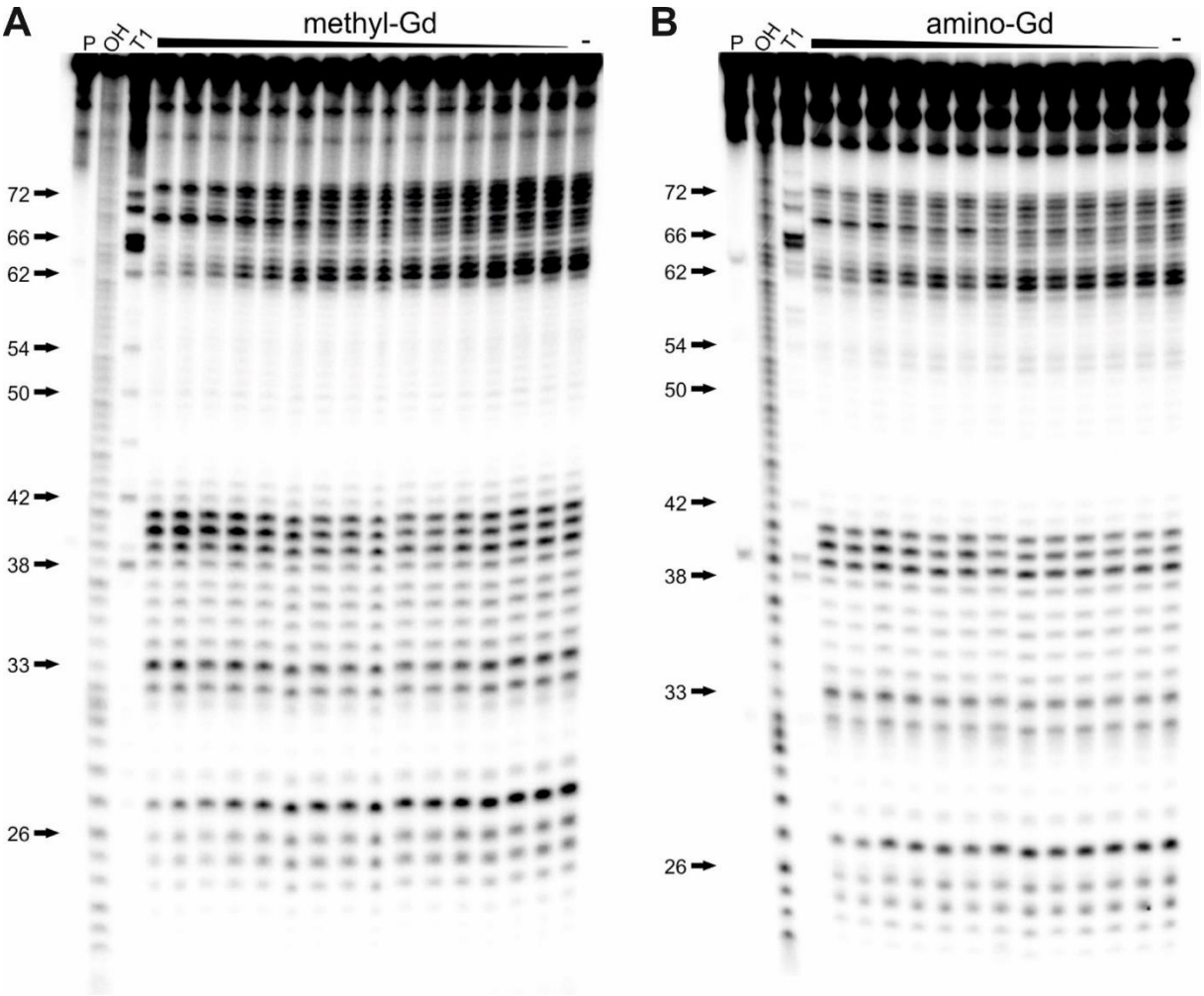


Figure 10: Methyl-guanidine and amino-guanidine binding by the *95 Lla* motif from *L. lactis*. (**A**) PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *95 Lla* RNA without (-) or with methyl-guanidine hydrohydrochloride in a range of 0.61 µM – 10 mM. P, OH and T1 have the same meaning as in Figure 5B. The plot of the fraction of RNA bound to ligand-derivate as a function of the logarithm (base 10) of the molar concentration is shown in **Figure** 9C. (**B**) PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *95 Lla* RNA without (-) or with amino-guanidine hydrohydrochloride in a range of 4.88 µM - 10 mM. P, OH and T1 represent 5´-^{32}P-labeled RNA undergoing no reaction, digest with RNase T1, or digest under alkaline conditions, respectively. The plot of the fraction of RNA bound to ligand-derivate as a function of the logarithm (base 10) of the molar concentration is shown in **Figure** 9C.

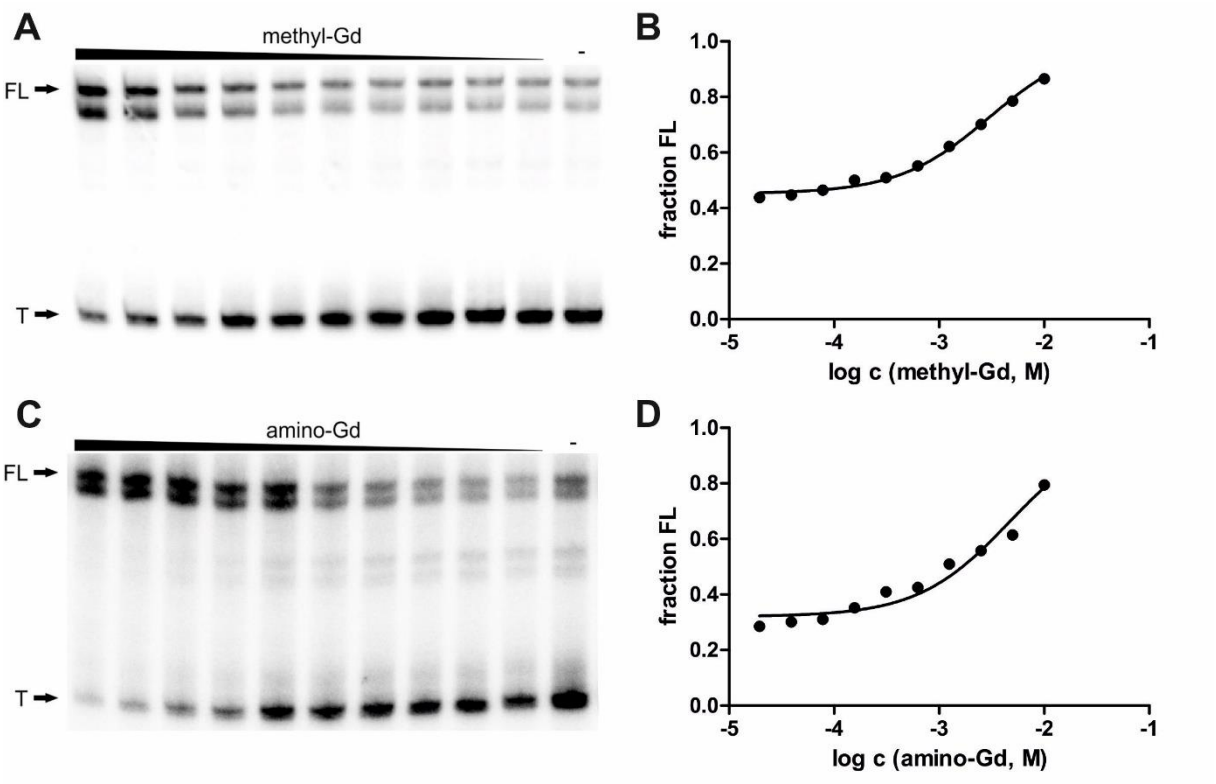


Figure 11: Guanidine derivatives prevent transcription termination less effectively than guanidine. (**A**) PAGE analysis of a transcription termination assay of *147 Lla* RNA without (-) or with methyl-guanidine hydrohydrochloride ranging from 19.5 µM - 4 mM. FL and T denote full length product at 147 nt and termination product at 93 nt, respectively. (**B**) Plot of the fraction of full length *147 Lla* productcontributing to the total number of transcripts (FL plus T) as a function of the methyl-guanidine hydrochloride concentration. An EC_{50} of 2.3 mM was determined applying a sigmoidal dose-response curve fit in three independent experiments (Supplementary Figure 4C,D). (**C**) PAGE analysis of a transcription termination assay of *147 Lla* RNA without (-) or with amino-guanidine hydrohydrochloride ranging from 19.5 µM - 4 mM. FL and T denote full length product at 147 nt and termination product at 93 nt, respectively. (**D**) Plot of the fraction of full length *147 Lla* product contributing to the total number of transcripts (FL plus T) as a function of the amino-guanidine hydrochloride concentration. An EC_{50} of 3.5 mM was determined applying a sigmoidal dose-response curve fit in three independent experiments (Supplementary Figure 4E,F).

4.3.Conclusion

The bioinformatics analysis of the *GGAM-1* motif indicated that its characteristics fit the typical properties of a riboswitch. These findings and the RNAs association with guanidine-related genes provided the basis for our hypothesis, that the *GGAM-1* motif represents a new class of guanidine riboswitches. In this chapter, were able to show that *GGAM-1* RNAs indeed selectively bind guanidine and in response, efficiently regulate the downstream genes *in vivo*. Comparing the *GGAM-1* motif with previously characterized guanidine riboswitch classes, there are no meaningful similarities. We therefore propose the name *guanidine-IV* riboswitches for *GGAM-1* RNAs.

Computational alignments showed that all representatives of the *GGAM-1* motif contain a predicted Rho-independent transcription terminator. In *in vitro* transcription assays we were able to demonstrate that the *GGAM-1* motif regulates on the level of transcription termination in response to guanidine. Hereby, addition of guanidine lead to a decrease of termination product, indicating that the induced modulation prevents formation of the terminator hairpin. Binding of guanidine depends on the presence of highly conserved nucleotides within the loop region. These nucleotides likely form a binding pocket that selectively binds guanidine and discriminates against other, similar compounds. Mutation of one of these highly conserved nucleotides lead to the loss of *in vitro* modulation and *in vivo* regulation of the downstream gene. Taken this *in vivo* and *in vitro* results together, we verified that addition of guanidine induces a structural modulation that results in the upregulation of the downstream gene via transcription termination control. This indicates that the *GGAM-1* motif function as genetic 'ON'-switch.

Multiple riboswitches that use Rho-independent terminators to function as ON-switches have been previously identified[80,99,25]. The typical mechanism for these riboswitches is that ligand-binding involves nucleotides that otherwise would form as part of the terminator stem. Ligand binding leads to the formation of a binding pocket and thus inhibits the formation of a terminator stem. However, the *guanidine-IV* riboswitch aptamer includes the terminator stem. Since the presented *in vitro* transcription and *in vivo* reporter expression experiments establish these riboswitches as ON-switches, they seem to use a new regulatory mechanism based on transcription termination. One potential mechanism is that ligand binding leads to a steric exclusion of a full terminator stem formation. Guanidine binding might stabilize a conformation of two kissing loops, possibly similar to the structure of the *guanidine-II* riboswitch[100,101]. This structure could stretch the linking region between the two stems to the point that the outer base pairs of the terminator stem do not form, hence resulting in antitermination.

The newly discovered *guanidine-IV* riboswitch shows similar ligand-binding characteristics to previously described classes. Some examples of the *guanidine-I*, *-II*, and *-III* riboswitches have been

reported to bind to guanidine with K_D's ranging from 25 to 300 µM[85–87], whereas the sequences tested in this work bind with dissociation constants of ~150–250 µM. With regards to the selectivity, the *guanidine-IV* riboswitch more closely resembles the classes II and III, since these also bind to guanidine derivatives with small substitutions such as amino- and methyl-guanidine. The strong discrimination against urea and arginine is shared with all three known classes of guanidine riboswitches[85–87].

Being the fourth class of guanidine sensing riboswitches, the *guanidine-IV* riboswitch class expands the list of structurally unrelated riboswitch classes that bind the same molecule. Apart from guanidine, already six riboswitch classes were identified that bind SAM[18,19,96,102–104], three for $preQ_1$[105,106,94] and two for cyclic di-GMP[32,95,107]. These observations raise the question of what factors lead to multiple structural solutions to bind a given molecule. The answer could relate to the biochemistry of the ligand and RNA, cellular metabolism or other issues. Regardless of the cause, it seems reasonable to speculate that further structural classes will be found for guanidine, SAM, $preQ_1$ and cyclic-di-GMP. Also, the identification of another guanidine riboswitch, again highlights the widespread occurrence of guanidine riboswitches. This also supports the idea that guanidine might have a prominent role in natural physiology or metabolism that has not been identified to date[108].

The *guanidine-IV* riboswitches are most commonly found upstream of the *sugE* gene which is common for guanidine sensing riboswitches. But the *guanidine-IV* riboswitches are also associated with genes that have not been found to be associated with other guanidine riboswitch classes. Two of these gene classes encode for the transporters PnuC and members of the MATE family. The identification and characterization of the *guanidine-IV* riboswitch class gives insights into the biological function of these uncharacterized proteins. Being under control of guanidine sensing riboswitches, it seems reasonable to speculate that they are guanidine exporters. The discovery of the *guanidine-IV* riboswitches does not only help to understand functions of associated genes and their encoded proteins, but also could be used as a detecting tool. The *in vivo* results show that the system works very efficient with a high-fold change in reporter protein expression, making it a promising candidate for the use as biosensor, where binding of a specific ligand leads to a detectable signal production.

5. Chapter III: Controllable gene expression in *Staphylococcus aureus*

5.1. Introduction

Staphylococcus aureus has been discovered in 1880 by the surgeon Sir Alexander Ogston. He analyzed post-operative wound suppuration and abscesses and observed grape-like clusters of bacteria which he described as "staphylococcal disease"[109]. Until now, *S. aureus* was demonstrated to be a dangerous human pathogen causing a wide range of infections including sepsis, abscesses, pneumonia, osteomyelitis, and infective endocarditis[110–112]. Today, *S. aureus* is one of the major pathogens and the cause of considerable human mortality, but also affects livestock animals like cows and sheep[113,114]. The spread of antibiotic-resistant strains, like methicillin-resistant *S. aureus* (MRSA)[115,116], make the treatment increasingly difficult and raises the need for investment in research.

With the growing need to understand *S. aureus* pathogenesis mechanisms, the demand for tools to control and monitor gene expression, protein localization and protein interactions increases. Overexpression systems can be helpful to study physiology and pathogenesis of *S. aureus*. The overproduction of a certain protein can lead to a distinct phenotype that gives insights in protein functions and potential protein interactions. However, some proteins that are interesting to be studied, are toxic to cells when overexpressed, raising the need for systems that enable the tight regulating of overproduction. One possibility of controllable gene expression is the integration of a genomic copy of the gene of interest under the control of a regulatable promotor. However, the establishment of such a system, with the need of integrating this construct in the genomic DNA, is very laborious, making plasmid-based systems more attractive. A plasmid with a riboswitch-based expression system allows controlling gene expression of the downstream gene through direct ligand binding and without any protein intermediates. A protein-based system is limited to the permeability of the regulating protein into the cell. Thus, these systems are dependent on the expression of the regulating protein at a stable level.

For the analysis of a controllable gene expression system, we decided to regulate the gene *egfp*, as it is one of the most used reporter genes. Its product, the fluorescent protein eGFP is very sensitive and has been shown to be suitable to observe processes in living cells and tissues[117,118]. In Chapter 2, we have already shown that the *guanidine-IV* riboswitch can be used in *S. aureus* to control the expression of the reporter gene *egfp*. Here, we optimize this easy and quick plasmid-based system to conditionally control reporter expression with a riboswitch in *S. aureus* RN4220.

5.2.Results & Discussion

5.2.1. Optimization of the *guanidine-IV* riboswitch *in vivo* reporter system

To assess a potential toxicity and impact of guanidine on growth, a growth curve of *wt S. aureus* RN4220 was measured in the absence and presence of 5 mM and 10 mM guanidine hydrochloride, respectively (Figure 12C). In the presence of 10 mM guanidine, bacterial growth was reduced that was not observed for 5 mM guanidine. We considered 5 mM guanidine as well-suited concentration for the following experiments.

To characterize the *guanidine-IV* riboswitch *in vivo*, we used the pCN-pBlaZ plasmid (kindly provided by the group of P. Romby, University of Strasbourg). To control expression of the reporter gene *egfp*, the *GGAM-1* motif sequence of *L. lactis* was placed in the *hly* 5´-UTR encoded on the plasmid. This plasmid-based system was shown to enable the guanidine concentration-dependent control of the reporter gene *egfp* (Chapter 2, Figure 8).

It is known that the dynamic range of a riboswitch can vary depending on its surrounding and genetic context. Thus, it can affect the switching activity when the riboswitch motif sequence is placed at different positions in the 5`-UTR. In the construct studied in Chapter 2, the motif sequence, including the terminator and the following nucleotides until the natural SD sequence, was placed at the position -33 of the hyl 5´-UTR. Here, the 16 nucleotides of the *hyl* sequence between the position -33 and the SD sequence (-33 until -17) were excluded (Figure 12A). We call this construct “33 (1)”. To further optimize this controllable expression system in *S. aureus*, additional constructs with the riboswitch sequence inserted at different positions were tested (sequences shown in Supplementary Table 3). The riboswitch sequence, including the terminator stem, was placed at the position -18, -33 and -47 in the 5´-UTR, respectively (Figure 12A). The positions -18 and – 47 were chosen, based on data from Monika Finke that show that the *guanidine-II* aptamer motif can control the downstream gene expression at both of this positions[119]. However, she found a higher switching activity at the position -18. Position -33 was chosen because the predicted secondary structure of the *hly* 5´-UTR appears to be a long stem, including a bulge and a loop on top (Figure 12A). We speculated, that inserting the riboswitch sequence at this position, would enable more flexibility likely needed for the *guanidine-IV* riboswitch to form a binding pocket. The reporter strains were grown in BHI medium in the presence or absence of guanidine hydrochloride. To monitor the expression of *egfp*, the eGFP fluorescence intensity was measured and normalized by the optical density (OD_{600}). For the constructs carrying the riboswitch motif at the position -47 and -18, addition of 5 mM guanidine hydrochloride resulted in a 10-fold and 12-fold increase of eGFP fluorescence, respectively (Figure 12B). The construct with the motif inserted at position -33 enabled an even 80-fold increase in fluorescence intensity. Thus, the

latter construct was further optimized trough additional adjustments of the riboswitch positioning, yielding the construct “33 (1)-(3)”.

The construct “33 (1)” excludes the *hly* 5´-UTR nucleotides from position -33 to position -17, placing the SD sequence directly behind the poly-U sequence of the terminator stem of the riboswitch. For the construct “33 (2)” the riboswitch motif sequence including the terminator stem was inserted at the position -33 in the 5`-UTR without replacing off any nucleotides from the *hly* sequence. Additionally, the construct “33 (3)” was designed, also carrying the riboswitch motif sequence at position -33, but including the natural sequence of the riboswitch motif from *L. lactis* with the terminator stem and the SD sequence and therefore lacking the rest of the *hly* 5`-UTR. The constructs “33 (2)” and “33 (3)” showed a 120-fold and 23-fold change in fluorescence intensity in response to 5 mM guanidine hydrochloride, respectively. All tested constructs showed a concentration-dependent change in eGFP fluorescence (Figure 12B). However, the construct “33 (2)” showed the best switching activity. With this result we can conclude, that we were able to optimize and enhance the controllable switching activity of the already established construct in Chapter 2.

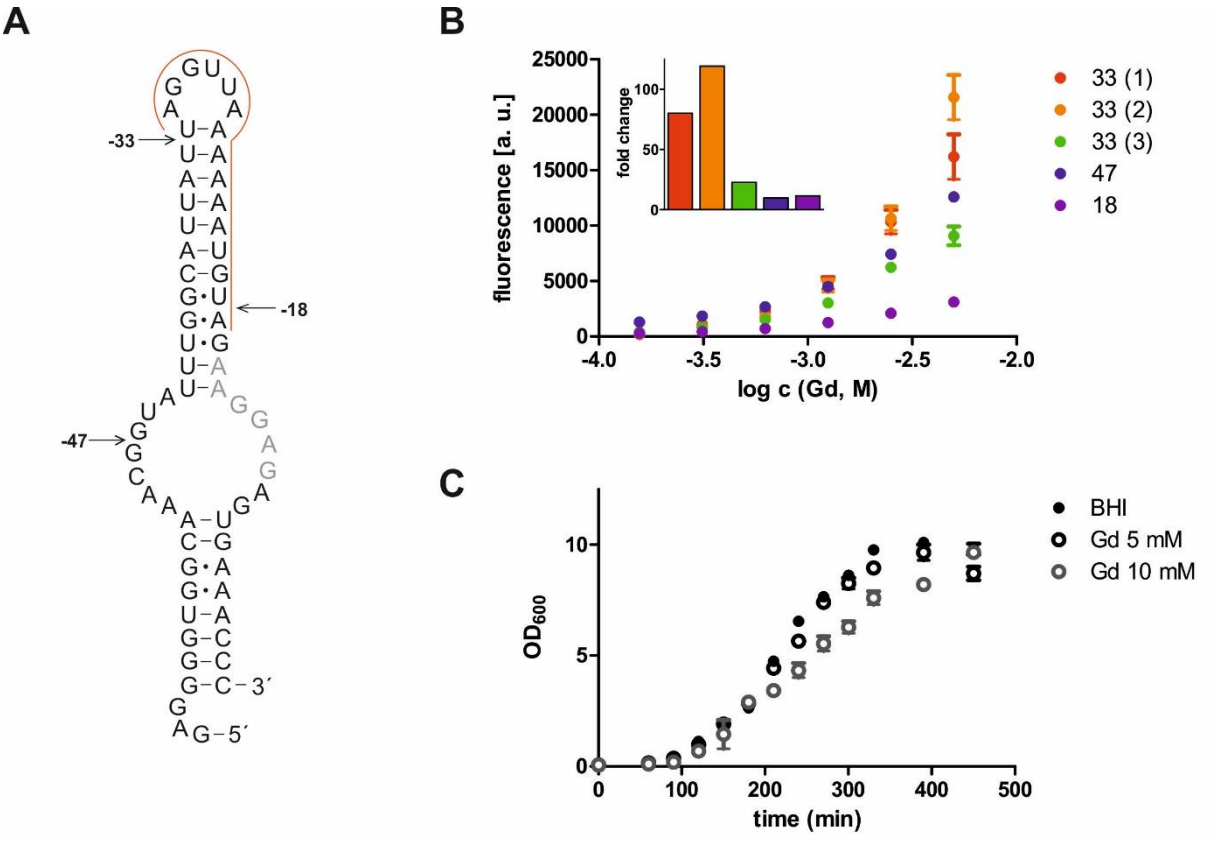


Figure 12: Optimization of *in vivo guanidine-IV* riboswitch constructs using the *hly* 5´-UTR in *S. aureus*. (**A**) Sequence and secondary structure of the *hly* 5´-UTR. Secondary structure is shown as predicted from the mfold server[120]. The SD sequence is shown in grey. For optimization, the *guanidine-IV* motif has been inserted at the positions -47, -33 and -18, that are indicated by arrows, respectively. Nucleotides that have been replaced in the construct "33 (1)" are indicated by a red line. (**B**) Plot of dose-dependent eGFP fluorescence in *S. aureus*. Strains were grown in media containing guanidine hydrochloride in the range 156 µM – 5 mM. Fluorescence was measured and normalized to OD_{600}. Error bars represent the standard deviation from three independent experiments. Fold changes observed in response to 5 mM are shown in the bar chart. (**C**) Growth curves of *wt S. aureus* RN4220. OD_{600} was measured over time in BHI medium without (BHI) and with 5 mM and 10 mM guanidine, respectively.

5.2.2. Comparison of the *guanidine-II* and *–IV* riboswitch for controllable gene expression

The plasmid-based system, established and optimized in this work, uses the guanidine-IV riboswitch to control gene expression via transcription termination. For a controllable gene expression system, it is advantageous when the signal can be detected early after induction. Thus, we compared inducible gene expression between systems controlling gene expression on the level of transcription and systems controlling translation. Monika Finke from our group has already established an inducible expression system for *S. aureus* using the *guanidine-II* aptamer, action on translational level[119]. Here, the aptamer sequence was placed at the position -18 in the *hyl* 5`-UTR upstream of the reporter gene *egfp*. With addition of 5 mM guanidine hydrochloride she observed an 8-fold increase in fluorescence activity. However, the construct established and optimized in the present work, using the *guanidine-IV* riboswitch sequence to control reporter gene expression, showed an increase of up to 120-fold under the same conditions. The *guanidine-II* riboswitch is functioning on the level of translation control. When guanidine is not present, the SD sequence is masked and the ribosome cannot bind to initiate translation. In the presence of guanidine however, structural rearrangement leads to the liberation of the SD sequence and translation takes place. By comparison, the *guanidine-IV* riboswitch is acting on the level of transcription control. Its ligand-unbound conformation favors formation of an intrinsic terminator to stop transcription. In the presence of guanidine, a binding pocket is formed preventing the full formation of the terminator and thereby enabling production of the full transcription product leading to an increase in gene expression. Thus, both riboswitches, *guanidine-II* and *guanidine-IV*, act as ON-switches but on different levels to control expression of the downstream gene. For an expression-system that can be used to control and monitor the expression, localization, and interactions of proteins, it is important that the expression induction is possible at different growth stages. Therefore, *S. aureus* cultures were induced in early, middle and late exponential phase and monitored the *egfp* expression by measuring every 5 min the eGFP fluorescence intensity and normalizing it by the optical density (OD_{600}). To compare the systems based on the *guanidine-II* and *guanidine-IV* riboswitch, experiments were performed simultaneously with *S. aureus* strains carrying the respective plasmid (Figure 13A-C). For both systems, an induction of the reporter gene expression is possible at all growth stages. An increase of eGFP fluorescence intensity could always already be measured 20 min after the induction, reaching maximal intensity after about 90 min. However, induction at the early exponential phase leads to the strongest increase in fluorescence intensity. For the samples that have not been induced with guanidine hydrochloride, no significant change was observed.

As described, the *guanidine-II* riboswitch acts via SD sequence masking. Guanidine binding directly leads to a liberation of the SD sequence and translation can occur. Instead, the *guanidine-IV* riboswitch uses a Rho-independent terminator stem to control gene expression. When guanidine is present, it binds to the emerging RNA, preventing transcription termination and thus enabling the formation of the full transcription product. While with the *guanidine-IV* riboswitch, the gene still needs to be transcribed, the full transcription product is already present in the *guanidine-II* based system. Since an increase in fluorescence was observed 20 min after induction for both systems, not the transcription of the gene, but the translation process and maturation of eGFP seem to be the crucial steps. The maturing time for eGFP was shown to take less than 10 min in *E. coli*, but can greatly vary between even closely related strains[121,122].

In the early as well as the late exponential phase, the *S. aureus* strain carrying the *guanidine-II* riboswitch-based expression system showed a higher background expression. In the absence of the ligand guanidine, masking of the SD sequence might be less efficient in this genetic context than using the terminator stem to prevent gene expression. The low background expression makes the *guanidine-IV*-based system attractive for expression studies.

In our experiments, expression of the reporter gene *egfp* was monitored by measuring the eGFP fluorescence intensity and normalizing it by the optical density (OD_{600}). For overexpression studies it is important to exclude to possibility that low fluorescence levels refer to misfolded eGFP and not to the absence of the protein. Therefore, the *S. aureus* strain carrying the "33 (2)" plasmid was cultivated in BHI medium and induced it at an OD_{600}=0.3 with 5 mM guanidine hydrochloride. 40, 60 and 80 minutes after the induction, samples were taken from an induced (+) and uninduced (-) culture, respectively. The samples were analyzed using SDS page and silver staining (Figure 13D). According with our fluorescence measurements, almost no eGFP was detected in uninduced samples. In contrast, already 40 min after induction with guanidine hydrochloride, eGFP can be found in the sample. The amount of produced eGFP increases over time.

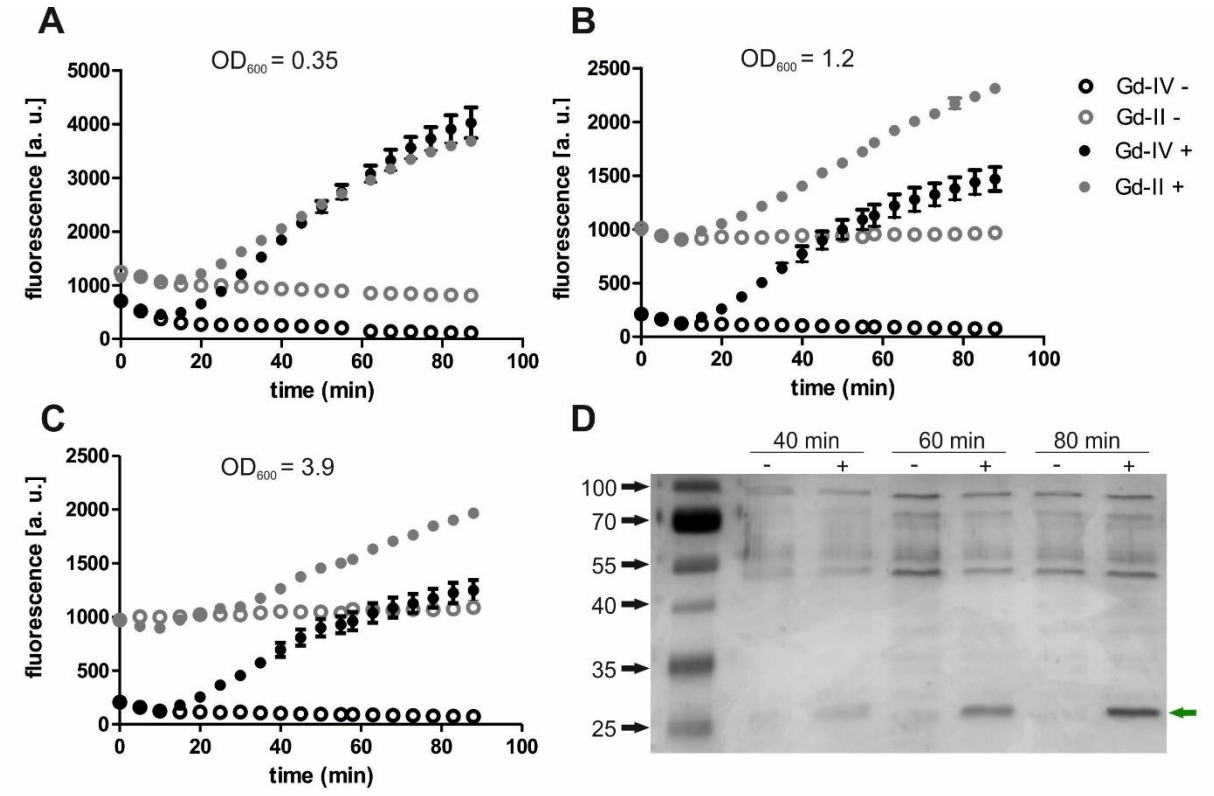


Figure 13: Inducible *egfp* expression using the *guanidine-II* and *–IV* riboswitch. eGFP fluorescence intensity in *S. aureus* carrying *guanidine-II* and *-IV* riboswitch constructs, respectively. Plots of time-dependent eGFP fluorescence intensity in *S. aureus*. At an OD_{600} = 0.35 (**A**), OD_{600} = 1.2 (**B**) and OD_{600} = 3.9 (**C**) one sample was induced with 5 mM of guanidine hydrochloride (Gd-II + and Gd-IV +) and one sample was grown in the absence of guanidine ("Gd-II -" and "Gd-IV -"). Fluorescence was measured every 5 min for 90 min and normalized to OD_{600}. Error bars represent the standard deviation from three independent experiments. (**D**) Silver stained SDS page showing different levels of eGFP overexpression in *S. aureus*. A strain carrying a reporter plasmid with the *guanidine-IV* riboswitch motif in the 5`-UTR of an *egfp* gene, was cultivated in BHI medium and induced at an OD_{600} = 0.3 with 5 mM guanidine hydrochloride. Samples were taken from an induced (+) and uninduced (-) culture 40, 60 and 80 minutes after induction. A prestained protein ladder (left lane) was used to identify the size of the proteins. Black arrows indicate protein size in kDa. Band referring to eGFP is indicated with a green arrow (~27 kDa).

5.2.3. The *guanidine-IV* system can be used as *in vivo* biosensor

The discovery of guanidine sensing riboswitches and their widespread occurrence in different phyla suggested that guanidine occurs in nature and might be involved in processes that remain to be identified. The identification of a forth class of guanidine riboswitches, presented in Chapter 2, supports the hypothesis that guanidine plays an important role in natural physiology and metabolism. To identify processes guanidine might be involved in and its origin in nature, it would be beneficial to find metabolic reactions that result in the accumulation of guanidine. In this context, to monitor the occurrence and concentration of a metabolite like guanidine, a biosensor could be very useful.

Riboswitches provide molecular tools for such *in vivo* biosensors. Exploiting biological reporters, such as eGFP, β-galactosidase and luciferase, the presence of the cognate ligand to the used riboswitch or aptamer can be monitored[123,124]. For example, using G-quadruplex formation of ATP-binding aptamers, the sensitive and selective detection and concentration determination of K^+ ions was possible[125]. Aptamers have also been used for the detection of small molecules. In a study from 2008, synthetically generated specific aptamers were shown to bind the explosive molecule TNT and provide the opportunity to determine the concentration directly in the field[126]. In 2016, aptamers were used to monitor the antibiotic neomycin in biological samples[127]. The controllable expression system that we presented in this chapter, is a promising biosensor, using *guanidine-IV* riboswitch motif as sensor unit and the reporter gene *egfp* as signal unit. The system works very efficient with a high-fold change (120-fold) in reporter induction, where binding of guanidine leads to a detectable signal production. We have already shown that the *guanidine-IV* riboswitch motif binds not only guanidine, but also its derivatives like methyl-guanidine and amino-guanidine, although with a lower affinity (Chapter 2, Figure 9). Thus, we hypothesized that the *in vivo* system would also lead to a detectable signal in the presence of this derivatives.

The *S. aureus* strain carrying the plasmid-based reporter system was grown in the presence of several guanidine derivatives and measured the eGFP fluorescence intensities. As expected, guanidine hydrochloride induced the highest increase in fluorescence intensity (Figure 14A). A significant difference in fluorescence intensity was detected from a concentration of 300 µM. However, hydroxy-guanidine also leads to an increase in fluorescence, with concentration higher than 1 mM. For methyl-guanidine a detectable increase at a concentration of 6 mM was observed. In addition to amino-guanidine, arginine and urea, eGFP fluorescence intensity was unaltered. With these results, the system is a promising *in vivo* biosensor for guanidine and some of its derivatives like hydroxy-guanidine. To demonstrate its applicability, the bacterium *Pseudomonas canavaninivorans* sp. nov., recently isolated by Franziskus Hauth from rhizophere soil and shown to degrade canavanine (manuscript in preparation), was used. Franziskus Hauth demonstrated that this novel strain can use

canavanine as nitrogen and carbon source, degrading it to homoserine and hydroxy-guanidine. Hydroxy-guanidine is only formed when canavanine is provided as carbon source. The strain was grown overnight in M9 medium with canavanine and glycerol as carbon source, respectively. Both samples were centrifuged for 5 min at 15,000 rpm at 4 °C. 200 µl of the supernatant were added to a *S. aureus* culture (OD_{600} = 0.3), carrying the *guanidine-IV* reporter plasmid in BHI medium containing erythromycin. The culture was incubated overnight at 37°C. eGFP fluorescence was measured and normalized by the optical density (Figure 14B). As a control, the *S. aureus* strain was cultivated in medium supplemented with canavanine to assure that it does not affect reporter activity. No change in fluorescence intensity occurred in presence of canavanine (data not shown). Samples grown with the supernatant of the *P. canavaninicorans* culture with glycerol as carbon source, did not show eGFP fluorescence intensity above background level. On the other hand, samples grown with the supernatant of the *P. canavaninicorans* culture with canavanine as carbon source, showed a 9-fold increase in reporter gene activity.

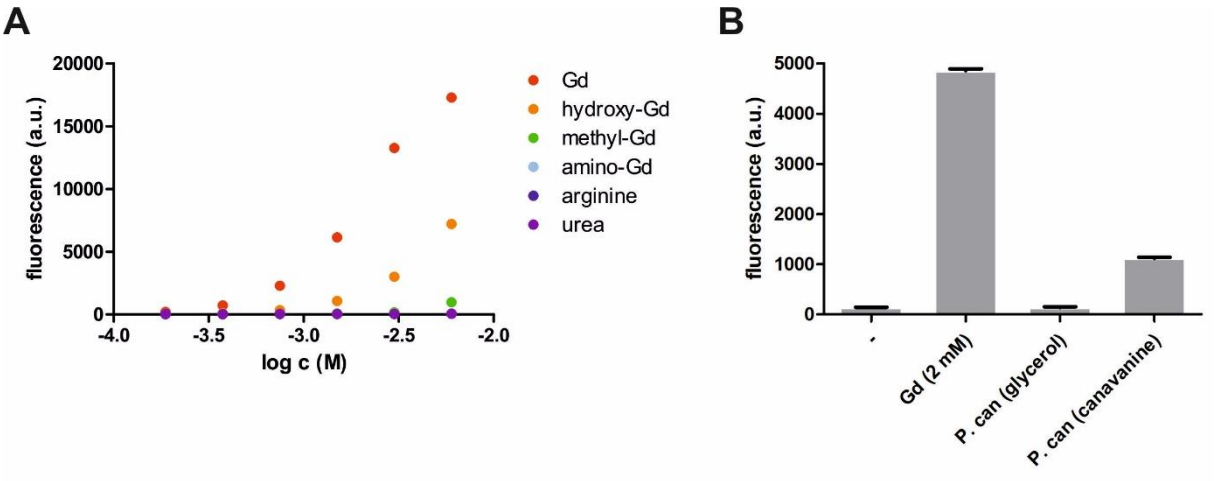


Figure 14: The *guanidine-IV* riboswitch construct can be used as biosensor for guanidine and guanidine derivatives *in vivo*. (**A**) Plot of dose-dependent eGFP fluorescence in *S. aureus* in response to guanidine and guandine derivatives. Strain was grown in media containing guanidine hydrochloride, hydroxy-guanidine, methyl-guanidine, amino-guanidine, arginine and urea in the range 187 µM–6 mM, respectively. Fluorescence was measured and normalized to OD_{600}. (**B**) Bars show eGFP fluorescence intensity of *S. aureus* carrying the *guanidine-IV* riboswitch plasmid, normalized to the OD_{600}. The strain was grown in the absence (-) or presence of 2 mM guanidine hydrochloride or in addition of *P. canavaninivorans* culture supernatants. Beforehand, *P. canavaninivorans* (P. can) was grown with glycerol and canavanine as carbon source, respectively. Error bars represent the standard deviation from three independent experiments.

5.3.Conclusion

In the first part of this chapter, the controllable expression system using the *guanidine-IV* riboswitch motif, presented in Chapter 2, was optimized in *S. aureus*. We showed that the motif sequence can be inserted at different positions in the *hly* 5´-UTR, but has the best switching performance at the position -33. We were able to increase the switching activity from 80-fold to 120-fold by optimizing the position of the riboswitch. Guanidine activates the expression of the reporter gene *egfp* in a dose-dependent manner. Also, the expression can be induced independently from the bacterial growth state. However, induction at the early exponential phase leads to the strongest induction of *egfp* expression. We compared our system with another expression system designed by Monika Finke, based the *guanidine-II* riboswitch. Although the *guanidine-II* motif acts as translational and the *guanidine-IV* as transcriptional switch, an increase in the eGFP fluorescence signal could be observed 20 min after induction for both systems. This shows, that not the transcription of the gene, but the translational process and the maturation of eGFP are the crucial steps. However, using the *guanidine-II* riboswitch to control gene expression, the background expression is higher in comparison. One could speculate that the SD sequence masking is not as efficient as the use of an intrinsic terminator to prevent reporter gene expression in this context in *S. aureus*. The expression system based on the *guanidine-IV* riboswitch, showed almost no background expression.

In the last part of this chapter, we showed that our *guanidine-IV* riboswitch-dependent expression system can also be used as biosensor for guanidine and analogs with only small additions like hydroxy-guanidine. Here, the riboswitch functions as the sensor unit and the reporter gene *egfp* as the signal unit. We were able to verify that the system monitors the occurrence and concentration of guanidine when incubated with the supernatant of cultures from other bacterial strains. This gives the opportunity to detect guanidine accumulation and thus it is a promising tool for the identification of processes in which guanidine is involved in and its natural source.

Taken together, the controllable expression system that we established in this work, is a promising tool for overexpression studies in *S. aureus* to control and monitor expression, localization and interaction of certain proteins. It is easy to handle, needs only minimal coding space and is not depending on any protein intermediates. Based on the data shown, it seems reasonable to test additional natural riboswitches that bind other ligands in this context in *S. aureus*. This could further expand the toolbox of controllable expression systems in *S. aureus* to study physiology and pathogenesis.

6. Chapter IV: Characterization of further riboswitch candidates

6.1. Introduction

Many riboswitch candidates have been identified using comparative sequence analysis. The main strategy to characterize and verify cognate ligands is still testing potential ligands in suitable assays. These potential ligands are selected by educated guesses based on the genetic context. For already verified riboswitches this was possible, since most of them regulate biosynthesis and utilization of the respective metabolite using a feedback mechanism. However, also riboswitches that do not act on a simple feedback mechanism have been discovered. For example, the second messenger cyclic-di-GMP is sensed by a class of riboswitches to control expression of diverse genes involved in unrelated cellular processes[95]. However, it is likely that many of the riboswitches that remain to be identified and characterized control not only processes to adjust and maintain intracellular metabolite concentration. In fact, many riboswitches are not expected to regulate on a simple feedback mechanism, making ligand prediction increasingly challenging. Moreover, there are also many orphan riboswitches that are hard to be matched with their natural target because the functions of biological pathways and processes they regulate are poorly understood or so-far even completely unknown[128]. One example are the guanidine sensing riboswitches. At the time of their discovery, guanidine was not known to play a role in biology and their associated genes were of unknown function. It was only after the ligand identification that the most widely associated genes were shown to encode for guanidine transporters. In cases like this, identification of the cognate ligand is of high interest since it gives insights into functions of the associated networks. However, here methods based on educated guesses limit the characterization of the riboswitch candidates. This shows the need of a systematic, unbiased pipeline for matching orphan riboswitches with their cognate ligand.

For this purpose, we developed a robust bacterial extract screening. The general workflow of this method is shown in Figure 15. *E. Coli* and *B. subtilis* cells are cultivated in minimal media under varying cultivation conditions, followed by a small molecule extraction. These extracts are then fractionated using HPLC and further tested in suitable assays like *in vitro* transcription termination assay as well as in-line probing reactions. Respective screening methods were chosen depending on the predicted mechanism of the studied riboswitch candidate. In the following chapter we present the characterization of three orphan riboswitches. Ligand predictions based on genetic context will be discussed as well as bacterial extract screenings.

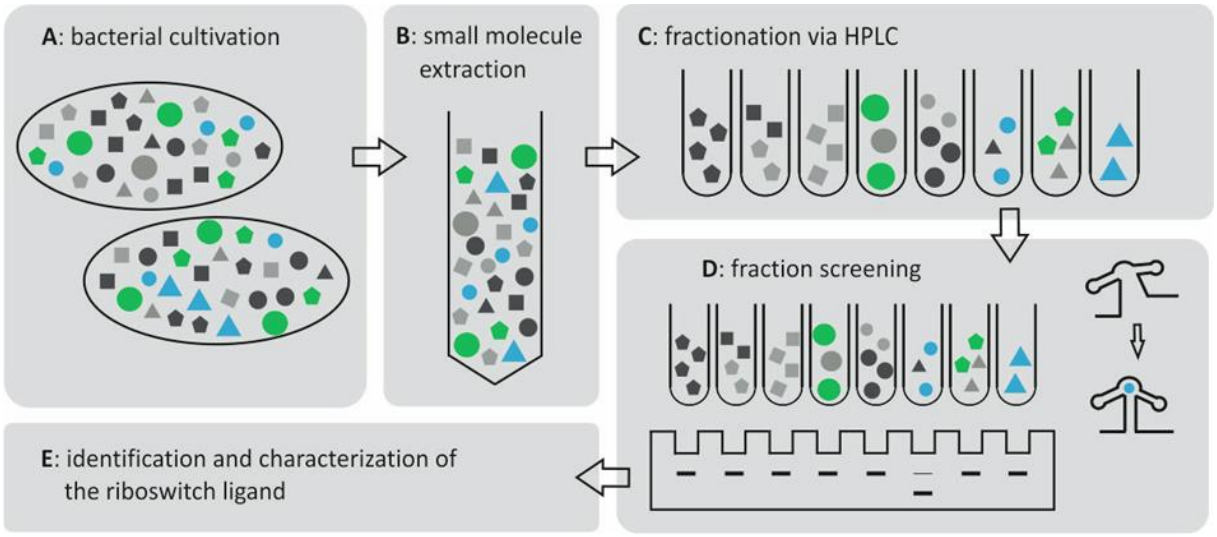


Figure 15: General workflow of bacterial extract screening. Grey, blue and green geometrical shaps represent small moecules. Bacterial strains are cultivated (**A**) and intracellular, small molecules extracted (**B**). The small molecule extract gehts fractionated using HPLC (**C**). These fractions can than be used for a screening (**D**) in suitable assays.

6.2.The *guanidine-IV variant* motif

Using discovery methods based on comparative genomics Zasha Weinberg has identified six guanidine riboswitch candidates, *GGAM-1-6*. In Chapter 2 we demonstrated that the *GGAM-1* motif represents a fourth class of guanidine riboswitches that act via transcription termination control[97,48]. Within a further detailed analysis of computationally predicted alignments, an RNA motif that is extremely similar to the *GGAM-1* motif has been identified. We call this motif the *guanidine-IV-variant* (*Gd4v*). Except for a few differences, the *guanidine-IV* riboswitch motif and the *Gd4v* motif seem to be almost identical in their primary and secondary structure. The *guanidine-IV* riboswitch includes a terminator hairpin that has a terminal loop with several conserved nucleotides. In the *Gd4v* motif, the highly conserved nucleotides that are likely to form a pseudoknot with nucleotides of the first hairpin, are found in an internal loop within the second hairpin (Figure 16). By mutating conserved nucleotides in the second loop of the *guanidine-IV* riboswitch, we were able to show that they are essential for the binding of guanidine. The structural differences at a position that is required for the efficient binding of the ligand is one hint that the *Gd4v* might differ in its ligand specify. Furthermore, the variant is most strongly (over 50 %) associated with a transaminase gene which is never associated with the *guanidine-IV* riboswitch (Figure 16). This transaminase is a member of the aspartate aminotransferase (AAT) superfamily of pyridoxal phosphate (PLP)-dependent enzymes, but its precise reaction is still unknown. Furthermore, the *sugE* gene, that is most commonly found downstream of the *guanidine-IV* riboswitch, is never found associated with the variant. However, there are also genes that are found upstream of both motifs, like members of the GNAT family N-acetyltransferase that catalyzes the transfer of an acetyl group from acetyl-CoA to a substrate. Also, both motifs are associated with multidrug and toxic compound extrusion (MATE)-like proteins (Figure 16).

About 390 unique representatives of the *Gd4v* motif were found, represented in the phyla Actionobacteria, Firmicutes and Bacteroidetes. Many of the representatives are also found in environmental sequences from which the environments with highest frequencies are mammal guts. All identified *Gd4v* RNAs contain a predicted Rho-independent transcription terminator. Having that and the structural similarities to the *guanidine-IV* riboswitch in mind, we hypothesize that the motif acts as an ON-switch on the level of transcription regulation, using the same mechanism as the *guanidine-IV* riboswitch. The differences in the genetic surrounding and in conserved nucleotides however, lead to the assumption that the cognate ligand might be structurally related to guanidine.

Gd4v number	Percent % (389)	Genes	Percent % (1262)	*Gd4* number
		EmrE/SugE	31	391
		NMN_transporter/PnuC/ NMN_trans_PnuC	25.2	318
		Pyridox_ox_2/NimA	10.5	133
		B3/B4	5	63
		UgpQ	0.87	11
15	3.9	MepA/matE/NorM/vmrA/ Polysacc_synt_C	22.6	285
42	10.8	other genes	2.3	29
78	20	GNAT_1 or GNAT_7	2.5	32
11	2.8	alanine dehydrogenase		
15	3.9	HDc		
22	5.6	YitT		
206	53	AAT_like/AspB/ viomycin_VioD		

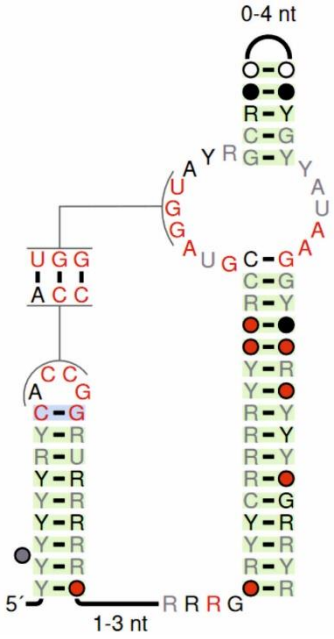


Figure 16: Genetic context and consensus sequence of the *Gd4v* motif. Table (left side) shows comparison between the genetic context of the *Gd4v* motif and the *guanidine-IV* riboswitch. Genes that are found to be associated with both motifs are shaded yellow. Diagram (right side) showing conserved features of RNA sequence in *Gd4v.* Symbols are explained in a legend (gray box) in Figure 4.

6.2.1. Results and Discussion

The *Gd4v* motif, identified by Zasha Weinberg, and the *guanidine-IV* riboswitch motif are almost identical in their primary and secondary structure. However, structural differences at positions with highly conserved nucleotides and the variation in genetic context between these two motifs, indicates that they might differ in their ligand specify. To verify that the *Gd4v* motif senses and binds a ligand structurally related to guanidine, a 95-nucleotide-long RNA construct (*95 Csp*) (Figure 17A) from the 5'-UTR of an aminotransferase gene of *Cloacibacillus sp.* was investigated in an in-line probing reaction. Guanidine causes a concentration-dependent structural modulation to this construct (Figure 17B). However, guanidine seems to bind this construct with a very low affinity compared to the *guanidine-IV* riboswitch. By quantifying the extent of spontaneous cleavage at nucleotide position U54 over increasing concentrations of guanidine hydrochloride, it was not possible to determine an apparent K_D value since no saturation was achieved with concentrations still suitable for the assay (Figure 17C). However, structural modulation induced by binding of guanidine to the RNA can be observed with concentrations higher than 1 mM. In comparison, the *GGAM-1* motif showed changes in modulation pattern at concentration from ~20 µM. Constructs of the *Gd4v* motif found in *Dialister succinatiphilus* (*95 Dsu*) and *Mitsuokella jalaludinii* (*105 Mja*) were tested. With these constructs no modulation in response to guanidine was observed, even at high concentrations.

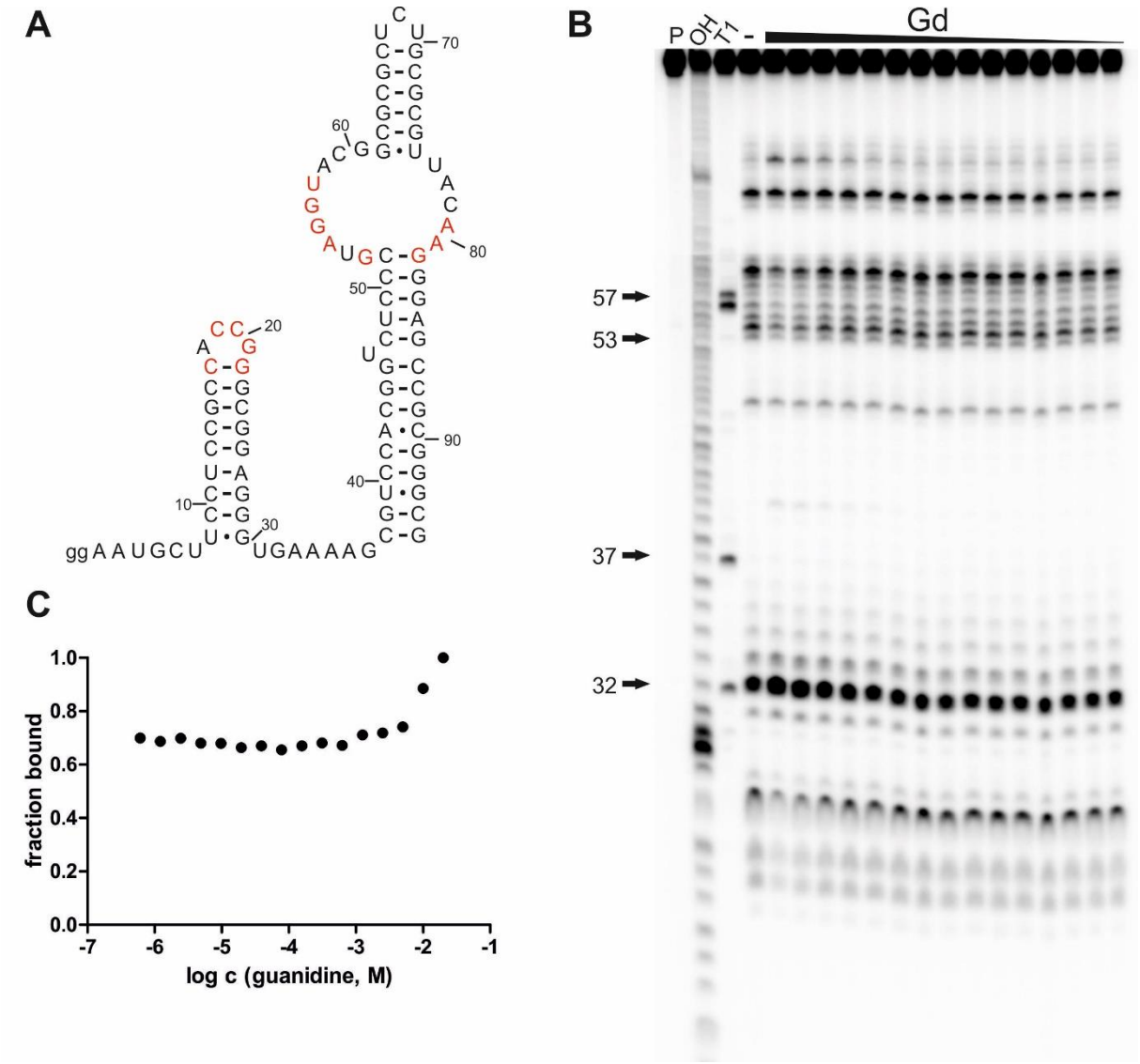


Figure 17: *Gd4v* motif RNA binds guanidine with low affinity. (**B**) PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *95 Csp* RNA (sequence from *Cloabacillus sp.*) without (-) or with guanidine hydrohydrochloride in a range of 0.61 µM – 10 mM. (**C**) Plot of the fraction of RNA bound to ligand as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined based on quantification of band intensity changes at U 54, normalized by the intensity of the constant band U 47.

Although we do see a guanidine-dependent structural modulation with the construct *95 Csp*, the low affinity indicates that either the construct is not chosen correctly or guanidine might not be the natural ligand of the *Gd4v* motif. It has already been shown for the *guanidine-IV* riboswitch, that the length of the construct tested in the in-line probing can be crucial for the observed binding affinity to the ligand[48]. Salvail et al. showed that binding to guanidine is significantly lower when the investigated in-line construct carries the full length of the terminator stem, including the following U residues. Formation of the terminator stem leads to a shorter single stranded linker between the two stems, which might permit contact between the two loops to prevent the binding pocket formation. As the *guanidine-IV* riboswitch and the *Gd4v* are structurally almost identical, it seems reasonable to

speculate that they are acting with the same mechanism to bind their cognate ligand. Thus, shorter constructs, that do not include the full terminator stem sequence, were tested (*81 Csp*, Supplementary Table 4). The guanidine dose-dependent structural change for this construct was even lower than observed with the long construct *95 Csp* (Supplementary Figure 6). Since changes in the length of the construct did not increase biding affinity to guanidine, we speculate that guanidine is not the natural ligand for the *Gd4v* riboswitch. However, as hypothesized in the beginning, the natural ligand might be structurally related to guanidine. A lot of characterized riboswitches, including the *guanidine-IV* riboswitch, have been shown to bind derivatives of their cognate ligands with a lower affinity. With respect to this it would make sense that the variant is still found to bind guanidine with a low affinity if a guanidine-derivate is its natural ligand. With this hypothesis, we have tested several guanidine derivatives, compounds carrying primary amino groups and signaling molecules (Table 1). Besides guanidine, only addition of methyl-guanidine and hydroxy-guanidine lead to a structural change of the *95 Csp* construct. However, the affinity for these two guanidine derivatives was even lower than observed with guanidine (data not shown).

Highly conserved nucleotides in the sequence of riboswitches are often essential for the formation of a binding pocket and are directly involved in interactions with the cognate ligand to enable highly selective binding. For many identified riboswitches it has been shown that point mutations at highly conserved positions in the sequence lead to a diminished or fully eliminated ligand interaction, proving that these nucleotides are involved in e binding. This was also shown for the *guanidine-IV* riboswitch. Guanidine-dependent *in vitro* modulation and *in vivo* gene expression control was fully eliminated by point mutations in the second loop. It seems reasonable that the highly conserved loop regions of stem 1 and stem 2 come together to form a selective binding pocket. For the *Gd4v* we speculated that it functions as riboswitch using the same mechanism. Thus, we hypothesized that point mutations in the second loop of the variant would also diminish guanidine-dependent modulation. Therefore, two different mutant constructs of the *95 Csp* RNA were tested in in-line probing reactions (Figure 18A). The tested mutant constructs, M1 (C57G) and M2 (U58A), were analyzed in in-line probing reactions in the absence or presence of guanidine hydrochloride with concentrations of 10 mM, 1 mM and 100 µM (Figure 18B). When determining the fraction of RNA bound to guanidine, it can be observed that the guanidine-dependent modulation effect is extremely diminished, or almost eliminated (Figure 18C). These results show, that the binding of guanidine to the *95 Csp* construct is not due to an unspecific effect in response to high guanidine concentration, but relays on the presence of highly conserved nucleotides in its primary sequence.

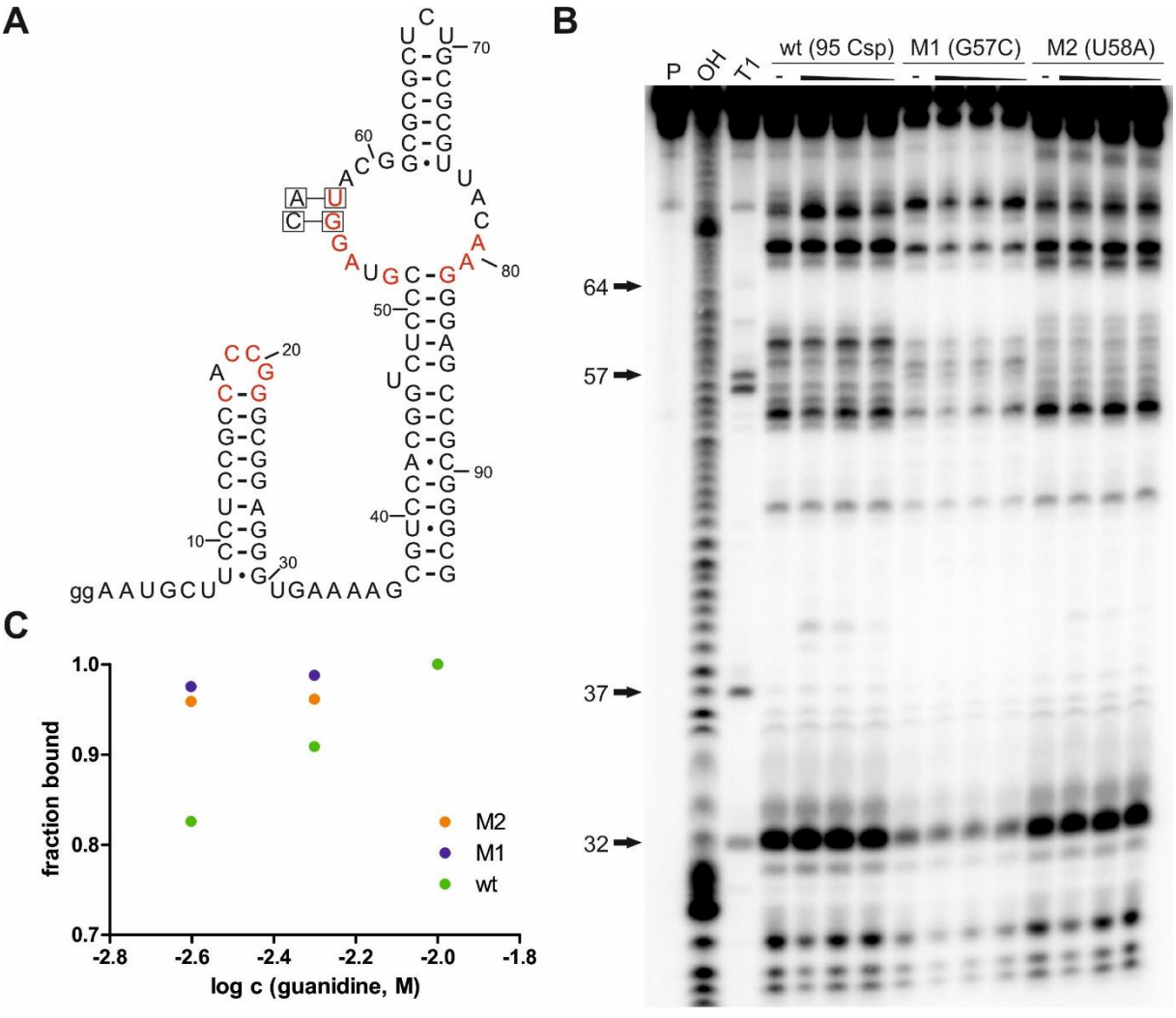


Figure 18: Point mutations of conserved nucleotides compromise guanidine binding to *Gd4v* motif RNA. (**A**) Sequence and secondary structure of *95 Csp* RNA with the location of point mutations in constructs M1 (G57C) and M2 (U58A). (**B**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *95 Csp* wt RNA and the mutants M1(G57C) and M2(U58A) without (−) or with guanidine hydrochloride with concentrations of 10 mM, 1 mM and 100 µM. P, OH and T1 represent 5′-^{32}P-labeled RNA undergoing no reaction, digest under alkaline conditions, or digest with RNase T1, respectively. (**C**) Plot of the fraction of RNA bound to guanidine for the constrcts *95 Csp* wt, M1 and M2 as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined based on quantification of band intensity changes at position U 54 and normalized by the intensity of the constant band U 47.

Almost all identified representatives of the *Gd4v* motif are found to be associated with an intrinsic transcription terminator. It seems reasonable that the variant would function as a riboswitch controlling the downstream gene expression on the level of transcription termination. Giving that, we also used the transcription termination assay to test potential ligands of this orphan riboswitch. Transcription of a DNA template for a 170-nucleotide-long RNA construct (*170 Csp*) from *Cloacibacillus sp.* and a 162-nucleotide-long RNA construct (*162 Dsu*) from *Dialister succinatiphilus* was monitored. Both constructs carry the *Gd4v* motif sequence and the following sequence context, including a

Rho-independent terminator stem followed by 7 U residues, the start codon and 47 and 43 nucleotides of the downstream gene, respectively. As we speculate the *Gd4v* to act as an ON-witch on transcription level, like the *guanidine-IV* riboswitch, we expect an increase of the full-length product and a decrease of the termination product in response to the natural ligand. For both constructs used in the transcription termination assay, only formation of termination product was observed and no full-length product was detected in the absence of potential ligands. This supports our hypothesis that the *Gd4v* appears to be an ON-switch. All compounds listed in Table 1 were tested on both constructs, *162 Dsu* and *170 Csp*. For the construct *162 Dsu* no changes in the transcription product ratio in response to any compound tested was observed. For the construct *170 Csp*, a decrease of the termination product in a concentration-dependent manner in response to guanidine hydrochloride, as expected for an ON-switch, was observed. However, no full-length transcription product could be detected (data not shown). Hence, it seems likely that guanidine is not the natural ligand to this motif or that we did not meet the right conditions for this construct. This again supports our hypothesis that the *Gd4v* motif evolved from the *guanidine-IV* riboswitch, resulting in a riboswitch whose natural ligand is structurally related guanidine. Thus, guanidine is still able to bind the variant motif with a low affinity but does not trigger the change of transcription product formation in the way it is expected with the natural ligand itself.

For ligand identification, also bacterial extract screenings were performed. Therefore, *B. subtilis* and *E. coli* was cultivated in M9 medium. In the exponential phase and stationary phase, a small molecule extraction was done, respectively. In the following, the extract was lyophilized, solved in water and subsequently fractionated using HPLC. The fractions were further tested in in-line probing reactions and transcription termination assays (data not shown). However, no fraction was identified to induce a change in cleavage or transcription pattern.

To further verify our findings and to learn about the mechanism of the *Gd4v,* its regulation of the expression of a downstream gene was monitored *in vivo*. A reporter plasmid carrying the *Gd4v* motif of *Cloacibacillus sp.* in the 5´-UTR of an *egfp* reporter gene was used. The same reporter plasmid, only varying in the motif sequence, has already been studied to characterize the *guanidine-IV* riboswitch motif. Here, we were able to show that the motif can control the expression of the *egfp* reporter gene in response to guanidine. Based on the high structural similarity between the *guanidine-IV* riboswitch and the variant, we hypothesized that both function with the same mechanism leading to the assumption that the variant motif also works within the context of this plasmid. This reporter plasmid was transformed in *S. aureus* and the expression of *egfp* monitored by measuring the fluorescence intensity and normalizing it by the optical density (OD_{600}). When cultivating the strain in rich medium (BHI medium), fluorescence only slightly above background level was detected. Assuming the variant

to be an ON-switch, these *in vivo* results indicate that the natural ligand is not expressed in *S. aureus* growing under these conditions. When the *S. aureus* strain was cultivated in the presence of 5 mM guanidine hydrochloride, no change in fluorescence intensity was observed. Although a change in the modulation pattern in in-line probing reactions was detected in addition to guanidine hydrochloride, the binding affinity was quite low. *In vivo* it was not possible to use guanidine hydrochloride concentrations higher than 5 mM in the medium without effecting bacterial growth. It is possible that with this amount of guanidine hydrochloride in the medium, the intracellular concentration is not high enough to induce a detectable change in gene expression.

Table 1: Compounds tested as potential ligand of the *Gd4v* motif. Highest concentration that has been used are listed. All compounds have been tested in-line probing reactions with the construct *95 Csp* and *95 Dsu* and in transcription termination assay with the constructs *170 Csp* and *162 Dsu*. Compounds that induced a structural modulation in in-line reaction with the *95 Csp* construct are highlighted in grey.

compound	highest concentration tested
guanidine hydrochloride	10 mM
methyl-guanidine	10 mM
amino-guanidine	10 mM
hydroxyguanidine sulfate	10 mM
arginine	10 mM
urea	10 mM
ppGpp	100 µM
pppGpp	100 µM
NTPs	2.5 mM
dNTPs	1 mM
cAMP	1 mM
cGMP	1 mM
c-di-GMP	1 mM
SAM	1 mM
Biotin	100 µM
Isoleucine	1 mM
Glutamate	100 µM
Lysine	1 mM
Valine	1 mM
Alanine	1 mM
Threonine	1 mM
Agmatine	1 mM
Proline	1 mM
Leucin	1 mM
Ornithine	1 mM
Hypoxanthine	5 mM
Xanthine	1 mM
Creatin	1 mM
Creatinine	1 mM

6.3.The *ablB* motif

Zasha Weinberg and his group are using comparative sequence analysis algorithms to screen bacterial genome databases and identify potential riboswitch candidates. Using this method, they also identified the *ablB* motif RNA. This motif is found in the phylum fusobacterium, in many clostridia species, but also in two bacilli species. It is often located upstream of two genes that might be involved in the biosynthesis of N^{ε}-acetyl-β-lysine. N^{ε}-acetyl-β-lysine is a zwitterionic, uncharged and highly water-soluble molecule that can be accumulated under salt stress in some bacterial species[129]. Although β-amino acids are rarely used for coping with osmotic stress[130], N^{ε}-acetyl-β-lysine can be used intracellular for osmoadaption to prevent dehydration. When microorganisms have to adopt to medium or high salt concentration in their surroundings, there are two major strategies: the salt-in-cytoplasm strategy and the organic-osmolyte strategy[131]. Using the salt-in-cytoplasm mechanism, the osmotic equilibrium is achieved by maintaining a salt concentration (KCl) in the cytoplasm that is similar to the salt concentration in the environment[132]. The organic-osmolyte strategy means the accumulation of a variety of solutes in the cell to counter the external osmotic pressure[133,134]. These low-molecular-weight, water soluble, organic compounds that are uptaken or synthesized in response to osmotic stress, are knows as compatible solutes[135,131]. N^{ε}-acetyl-β-lysine, product of the genes likely to be regulated by the *ablB* motif, is such a compatible solute. It has been shown that the synthesis of N^{ε}-acetyl-β-lysine is performed in two steps from α-lysine. The first enzymatic reaction is the catalysis α-lysine to β-lysine performed by a lysine-2,3-aminomutase. The second step is catalyzed by a β-lysine-*N*-acetyltransferase, acetylating β-lysine to form N^{ε}-acetyl-β-lysine[136]. In addition to various transporters, the *ablB* motif is found to be associated with a lysine-2,3-aminomutase and a β-lysine-*N*-acetyltransferase. For the species identified to carry the *ablB* motif, it has not been shown yet that they accumulate N^{ε}-acetyl-β-lysine to overcome osmotic stress, but it seems reasonable to speculate that here N^{ε}-acetyl-β-lysine is also involved osmoadaption. Compatible solutes are accumulated intracellularly by de novo synthesis or by uptake from the medium. Thus, production of proteins that synthesize or uptake these metabolites is under osmotic control[137]. If the *ablB* motif indeed controls the expression of proteins that synthesize a compatible solute, it makes sense that the activity of this potential riboswitch is depending on the osmotic stress in the cell. Thus, we speculate that the *ablB* motif RNA is a riboswitch sensing a molecule involved in osmotic stress signaling.

6.3.1. Results and Discussion

The *ablB* motif is found to be associated with proteins that might be responsible for the biosynthesis of the compatible solute N^{ε}-acetyl-β-lysine. Based on that, we hypothesized that it functions as an ON-switch, activated by a ligand that is accumulated at high osmotic pressure. To name some examples, proline, glutamine, taurine, trimethylaminoxid (TMAO) and glutamate were shown to be accumulated in bacterial cells to function as osmoprotectants and compatible solutes[137–140]. A list of compounds (Table 2) was screened in in-line probing reactions on their ability to induce structural changes on the RNA. Therefore, constructs carrying the motif sequence from three different organisms, *Acetobacterium woodii*, *Psychrilyobacter atlanticus* and *Dethiobecter alkaliphilus*, were used (*130 Awo*, *109 Pat*, *92 Pat*, *98 Dal*, *111 Dal*, Supplementary Table 5). All constructs showed a difference in their cleavage pattern with high concentration of potassium glutamate. Potassium glutamate is a major osmolyte in *E. coli*[138]. However, to further investigate this effect also potassium chloride and sodium chloride were tested in in-line probing reactions. With these the same effect was observed, indicating that the modulation by potassium glutamate was not selectively induced by glutamate, but results from the high potassium concentrations. For all constructs tested a change in the cleavage pattern was observed with potassium and sodium chloride with concentrations higher than 50 mM (Figure 19 and Figure 20). It was not possible to determine an apparent K_D value since higher salt concentrations were not suitable for the assay. However, the structural modulation that was observed for this RNA could also be an unspecific effect due to this high salt concentration. To explore this possibility and further characterize the motif, nucleotides that might be involved in binding were investigated. Therefore, the constructs 130 Awo and 109 Pat were mutated within the internal bulge of the third loop (Supplementary Table 5). These highly conserved nucleotides are the ones found to decrease in their intensity in response to salt and thus thought to be involved in recognition. No diminished modulation effect of potassium and sodium chloride with the mutated constructs could be detected (data not shown). As a control the *GGAM-1* motif construct *95 Lla* (Chapter 1) was used. It was already shown that this motif functions as a guanidine binding riboswitch. If the effect of potassium and sodium chloride on the *ablB* motif is a specific effect, no modulation with high salt concentration should be observed. Indeed, using concentration of potassium and sodium chloride of up to 200 mM, no modulation on the *95 Lla* RNA was observed.

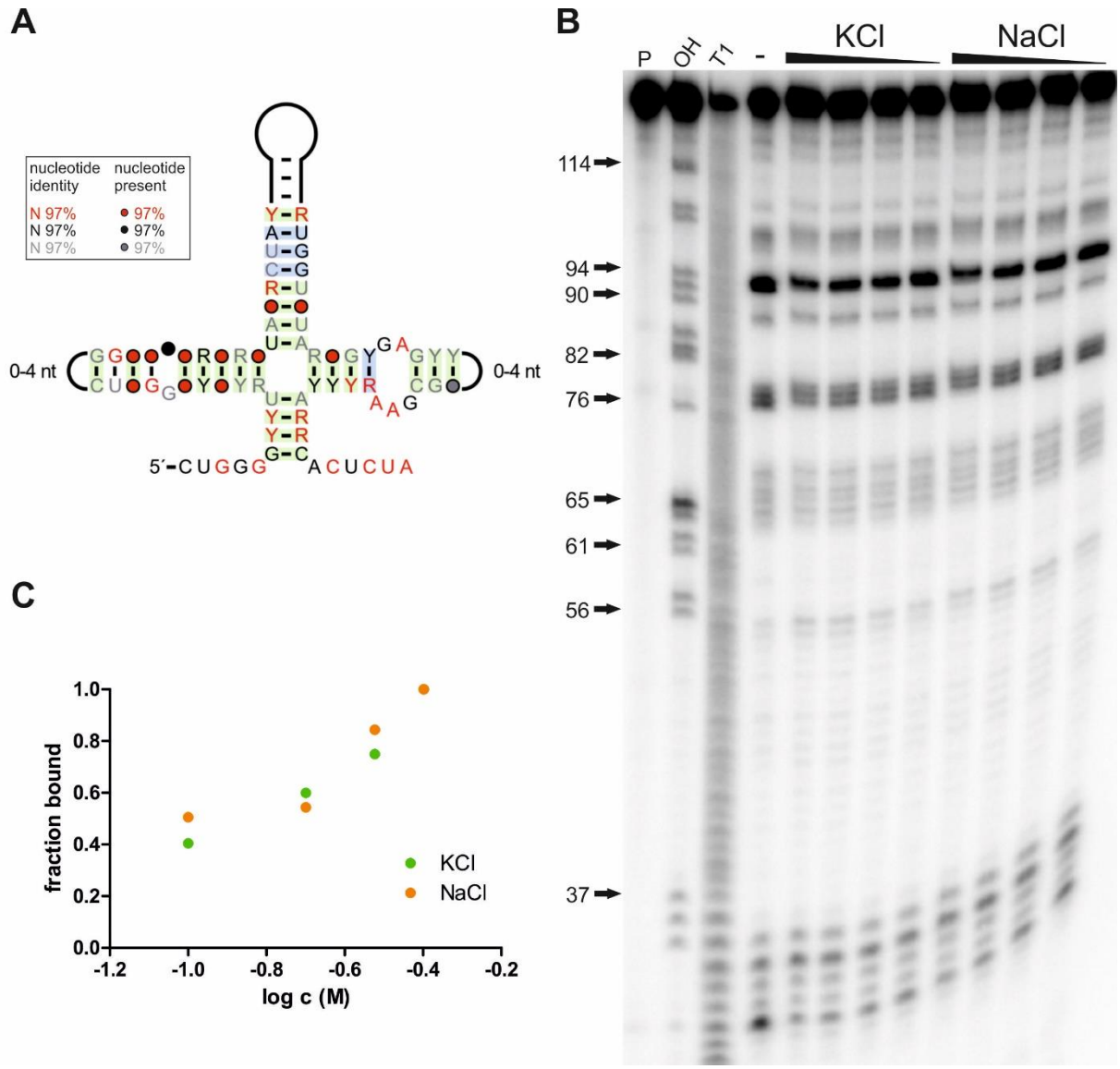


Figure 19: High salt concentration induce structural modulation at *ablB* motif RNA. (**A**) Conserved features of the *ablB* RNA motif. A legend (gray box) explains the symbols. (**B**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *130 Awo* RNA (sequence from *Acetobacterium woodii* without (-) and with KCl and NaCl in a range of 50 - 400 mM, respectively. (**C**) Plot of the fraction of RNA bound to ligand as a function of the logarithm (base 10) of the molar KCl and NaCl concentration, respectively. Fraction of RNA bound was determined based on quantification of band intensity changes at A 93, normalized by the intensity of the constant band G 56.

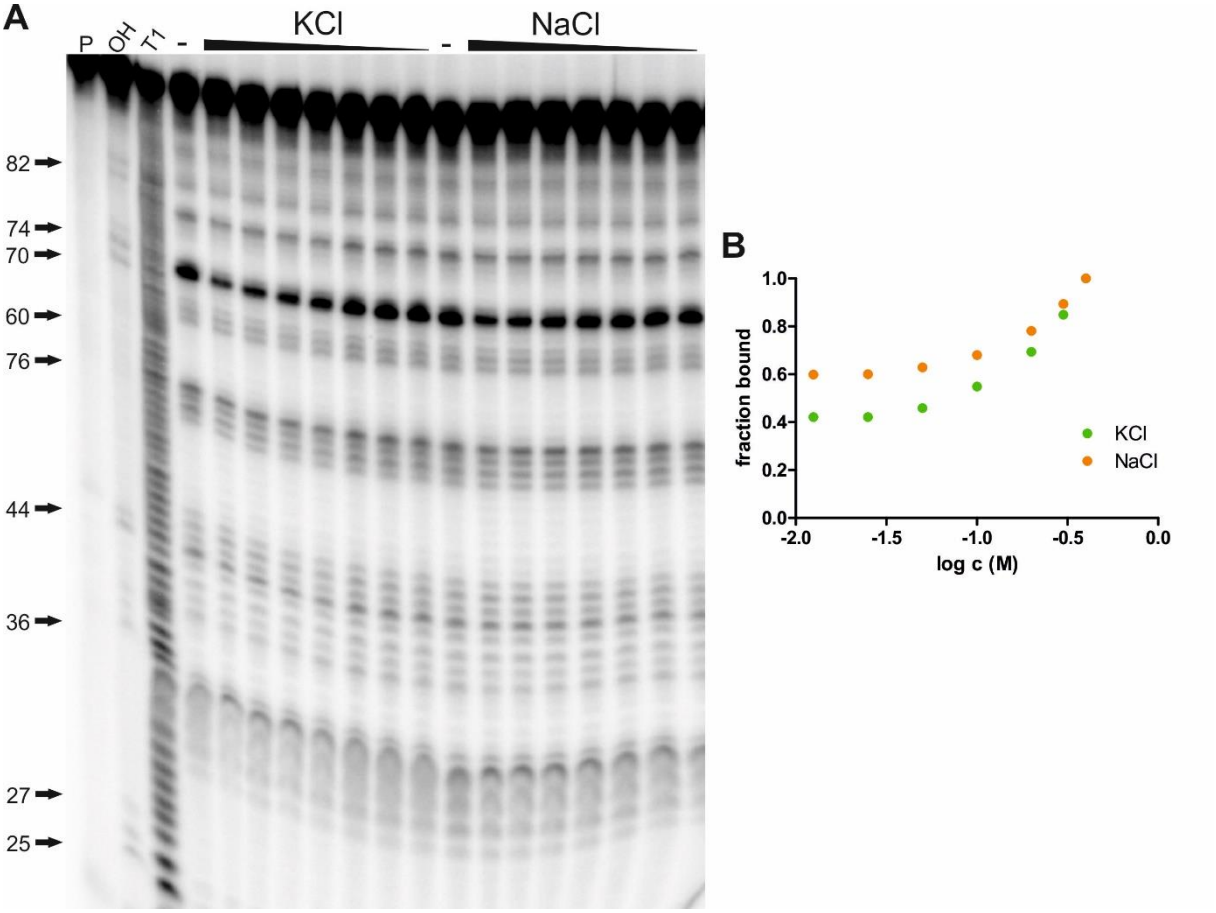


Figure 20: High salt concentration induce structural modulation at *ablB* motif RNA. (**A**) PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *109 Pat* RNA (sequence from *Psychrilyobacter atlanticus*) without (-) and with KCl and NaCl in a range of 6.25 mM - 400 mM, respectively. (**B**) Plot of the fraction of RNA bound to ligand as a function of the logarithm (base 10) of the molar KCl and NaCl concentration, respectively. Fraction of RNA bound was determined based on quantification of band intensity changes at A 73, normalized by the intensity of the constant band U 48.

Many of the identified examples for the *ablB* motif are found to be associated with an intrinsic terminator stem. Thus, all potential ligands (Table 2), already applied in in-line probing reactions, were tested in transcription termination assay. Transcription of the constructs *331 Awo* and *236 Pat* (Supplementary Table 5) was monitored, both carrying the *ablB* motif and the following natural sequence, including terminator stem, start codon and about 60 additional nucleotides. In the absence of any potential ligand, only termination product was formed and only little “full-length” product (data not shown). This finding supports our hypothesis that the *ablB* motif acts as an ON-switch. However, no changes of transcription in response to the tested molecules was observed.

Table 2: Compounds tested as potential ligand of the *ablB* motif. Highest concentration that has been used are listed. All compounds have been tested in in-line probing reactions with the constructs *130 Awo*, *92 Pat*, *109 Pat*, *98 Dal* and *111 Da*l and in transcription termination assay with the constructs *331 Awo* and *236 Pat*. Compounds that showed to induce a structural modulation in in-line reaction are highlighted in grey.

compound	highest concentration tested
KCl	400 mM
NaCl	400 mM
cGMP	1 mM
cAMP	1 mM
ppGpp	100 µM
pppGpp	100 µM
c-di-GMP	1 mM
c-di-AMP	1 mM
glutamate	400 mM
glutamine	1 mM
SAM	1 mM
Proline	10 mM
Lysine	10 mM
Glucose	100 mM
Mannitol	100 mM
Betaine	100 mM
Urea	100 mM
Choline	100 mM
Ectoine	100 mM
Taurine	100 mM
Glycine	100 mM
TMAO	100 mM
Trimethylamine chloride	100 mM

6.4.*folP* RNA Motif

Using comparative sequence analysis algorithms, bacterial genome databases are screened and riboswitch candidates identified. However, identification of cognate ligand and within that proof of functionality of the potential RNA motif is still a big challenge. In this context, the *folP* RNA motif was identified by Weinberg *et al.*[77] in 2017 via comparative sequence analysis of 5'-UTRs, but its ligand remains to be identified. The *folP* motif is found in *Dialister* bacteria of the phylum Firmicutes and in environmental samples. Since most of the *folP* motif RNAs are followed by a predicted terminator stem, making it likely that its regulation mechanism is based on intrinsic transcription termination. The RNA motif is predicted to form three hairpins. P1 and P3 exhibit a highly conserved nucleotide sequence, whereas P2 is supported by covariation. The *folP* motif is always found in the 5´-UTR of the *folP* gene encoding the dihydropteroate synthase[77]. This enzyme catalyzes the condensation of 4-amino-benzoate and 6-hydroxy-methyl-dihydropterin diphosphate to produce dihydropteroate. As a part of folate biosynthesis pathway dihydropteroate is further converted over dihydrofolate to tetrahydrofolate (THF) and 10-formyl-tetrahydrofolate (10f-THF) to function as carrier of one-carbon units. Methyl, methylene or formyl groups are transferred within the biosynthesis of methionine, thymidine and purine[141]. De novo purine biosynthesis is a high-flux pathway in replicating cell and thus, folate mediated one-carbon metabolism plays a fundamental role in metabolic cellular processes like DNA synthesis. This makes it important that the availability of 10f-TFH within the cell is robust and must therefore be tidily regulated[142,143]. A regulatory active ncRNA upstream of the *folP* gene would thus present an additional level of regulation for folate metabolism. therefore, we speculated that the *folP* motif RNA is a riboswitch involved in the regulation of folate biosynthesis.

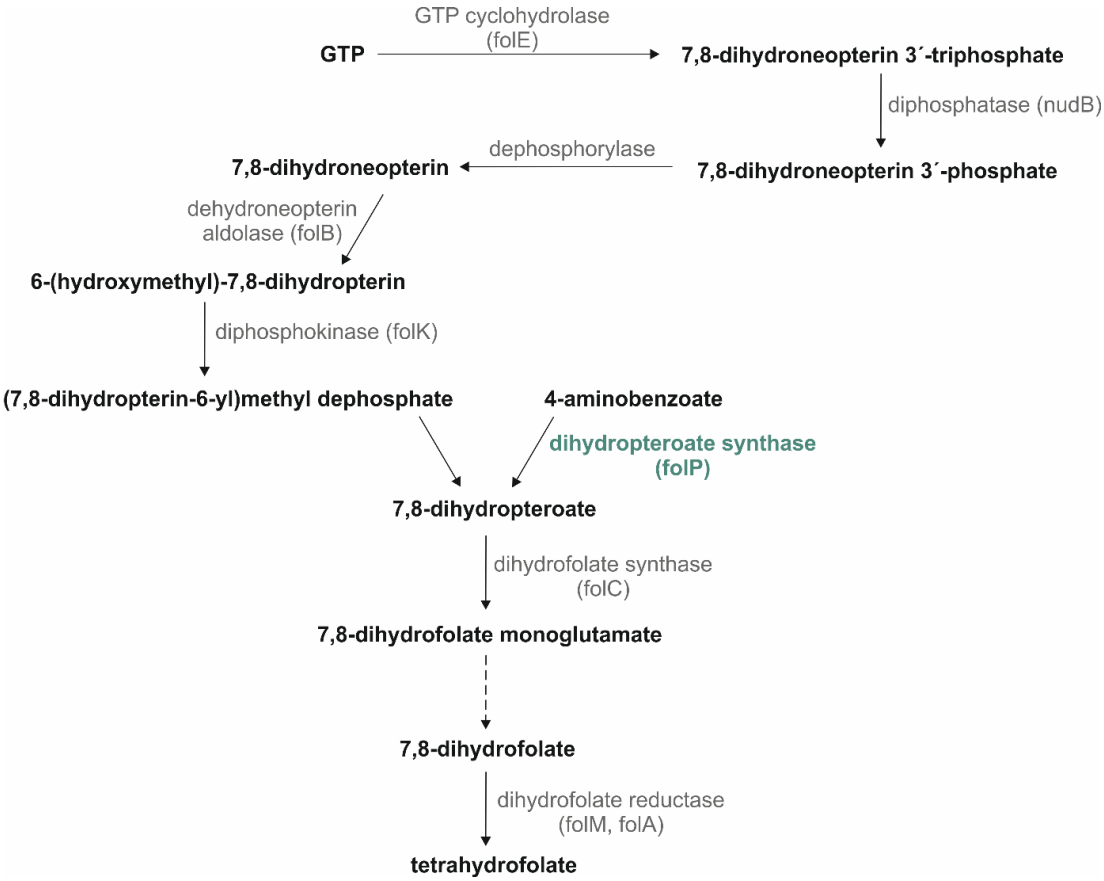


Figure 21: Simplified diagram of the tetrahydrofolate synthesis. Tetrahydrofolate, necessary for folate mediated one-carbon metabolism, plays a fundamental role in metabolic cellular processes. THF is synthesized by multiple enzymatically catalyzed reactions (involved enzymes shown in grey). The riboswitch candidate *folP* is found to be associated with the enzyme dihydropteroate synthase (green), that catalyzes the condensation of 4-amino-benzoate and 6-hydroxymethyl-dihydropterin diphosphate to produce dihydropteroate.

6.4.1. Results and Discussion

The dihydropteroate synthase folP, which is associated with the *folP* RNA motif, plays a role in the de novo folate synthesis. Folates like 10f-THF are required for formylation reactions within the de novo purine biosynthesis. We hypothezied that the *folP* RNA motif is involved in the regulation of the intracellular 10f-THF level and other THF derivatives. To explore this role, a reporter plasmid containing the motif sequence from *Dialister invisus* in frame with the reporter gene *lacZ* was designed. Thus, the construct carries a translational fusion between *folP* motif RNA including the predicted terminator stem, start codon and 8 codons of the native sequence and a *lacZ* reporter gene. We speculated that treatment with antifolates that target the de novo folate synthesis, should show an effect on the reporter gene expression in *E. coli* cultures carrying the *folP*-motif:*lacZ* reporter plasmid. Using agar diffusion assays and liquid-based assays, the *E. coli* reporter strain was incubated in the presence of the antifolates trimethoprim, sulfamethoxazole and methotrexate. Trimethoprim inhibits the dihydrofolate reductase und thus leads to a depletion of tetrahydrofolate[144]. Reporter gene expression was quantified using ONPG assay and normalized by reaction time and OD_{600} to obtain Miller units. Different concentrations of trimethoprim were tested and their effect on the bacterial growth monitored to find an optimal concentration where folate stress is as strong as possible but not lethal. With a concentration of more than 0.2 µg*ml^{-1} growth of *E. coli* was significantly reduced (Figure 22). Therefore, concentrations lower than 0.2 µg*ml^{-1} were used in the following experiments. When treated with trimethoprim, the *E. coli* reporter strain showed an increased gene expression in a dose-dependent manner (Figure 22). A plasmid carrying a constitutively translated *lacZ* reporter gene served as control. This could be a hint that the *folP* RNA motif regulates its downstream gene in response to a metabolite that is involved in the folate biosynthesis. To follow our hypothesis, the *E. coli* reporter strain was also cultivated in the presence of the antifolates sulfamethoxazole and methotrexate. Sulfamethoxazole is a structural analog of para-aminobenzoic acid and competes to bind the dihydropteroate synthase *folP*[145]. Methotrexate competitively inhibits the dihydrofolate reductase[146,147]. With both compounds, sulfamethoxazole and methotrexate, no effect on the reporter gene expression was observed (Figure 22). This shows that the effect observed with trimethoprim is not only an effect due to general cellular stress as response to antibiotic treatment. On the other hand, these results raise the question why treatment with other antifolates does not increase the reporter activity. One possible explanation is that the *E. coli* cells might undergo additional metabolic alterations under these conditions or otherwise suppress reporter gene expression.

When *E. coli* is grown in minimal medium (M9), the cells produce purines via their de novo purine biosynthesis and 10f-THF is needed for formylation steps. When the cells are grown in rich medium (LB), this pathway is not favored. Therefore, the expression of the reporter gene when the strain is

cultivated in M9 and LB medium was compared. In contrast to our expectations, the level of reporter gene expression measured in both mediums does not significantly differ from each other.

Based on our *in vivo* results of *E. coli* cultures carrying the *folP*-motif:*lacZ* reporter plasmid, educated guesses for possible ligands were made, representing metabolites that are intermediates within the THF synthesis and de novo purine biosynthesis. To test potential ligands in a binding assay, different *folP* constructs were analyzed regarding their folding. Therefore, constructs of different lengths from three organisms were used: *Dialister succinatiphilius*, *Dialister invisus* and *Veillonellaceae bacterium*. All constructs folded as predicted, forming three stems. However, we decided to use the constructs *105 Vba* and *87 Vba* in following in-line reactions since here the formation of stem P3 is most stable. The *folP* motif RNA is always found to be associated with a predicted terminator stem. We thus hypothesized, that the RNA motif regulates its downstream gene on the level of transcription termination. Addition of the cognate ligand in a transcription termination assay should lead to a change in the transcription product. Given that, all potential ligands were tested in in-line probing reactions as well as in transcription termination assays. However, all molecules tested showed neither a modulation of cleavage pattern in in-line probing reactions nor a change in transcription.

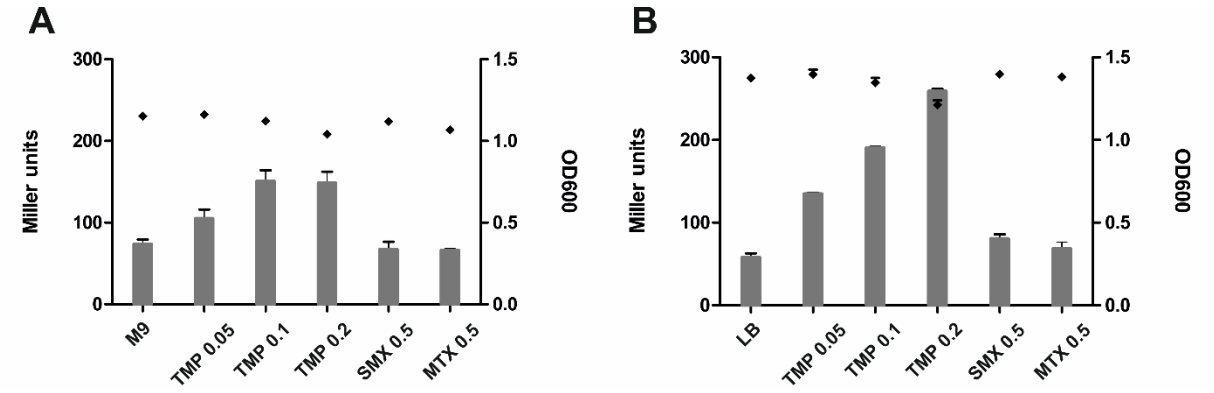


Figure 22: The *folP* motif RNA regulation is effected in response to trimethoprim. (**A**) Reporter activity of *E. coli* cultures carrying *folP*-motif:*lacZ* reporter plasmids incubated in M9 medium with different concentrations of trimethoprim (0.05, 0.1 and 0.2 µg*ml^{-1}) or with 0.5 µg*ml^{-1} of sulfamethoxazole and methotrexate, respectively. Activity is given in Miller units monitored by the ONPG assay. (**B**) Reporter activity of *folP*-motif:*lacZ* cultures grown in LB medium in the absence or presence of different abtifolates. All reporter assays are shown as mean ± standard deviation of triplicate measurements.

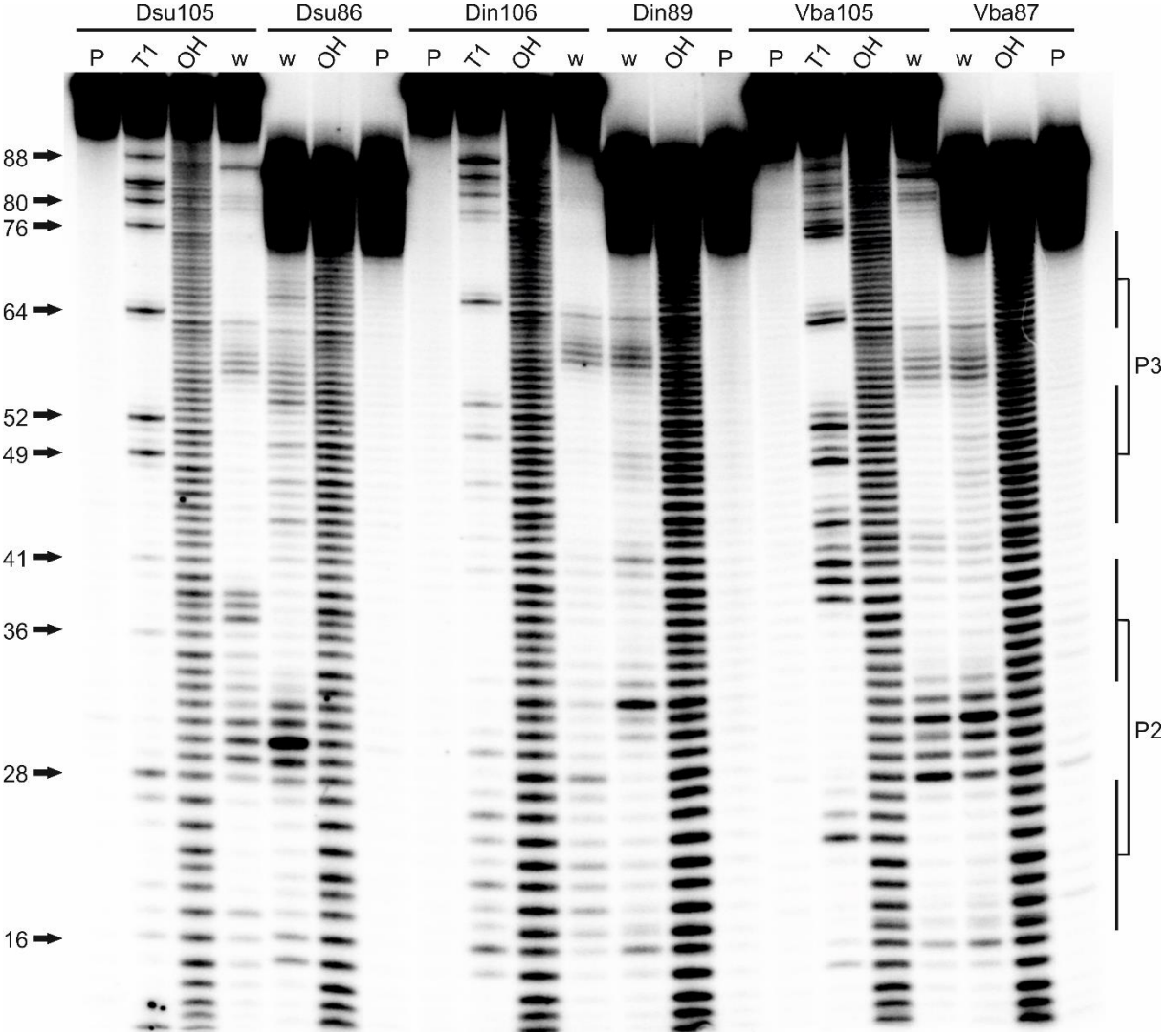


Figure 23: The *folP* RNA motif folds *in vitro* as predicted. PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *folP* RNA of different constructs. Constructs from three organisms have been analysed: *Dialister succinatiphilius* (Dsu), *Dialister invisus* (Din) and *Veillonellaceae bacterium* (Vba). For each organism two different length were tested regarding their folding. P, OH and T1 represent 5´-^{32}P-labeled RNA undergoing no reaction, digest with RNase T1, or digest under alkaline conditions, respectively. Brackets indicate the formation of the stems P2 and P3.

6.5.Conclusion

In this chapter, three riboswitch candidates, identified by our collaboration partner Zasha Weinberg, have been studied. All three motifs exhibit properties that are characteristic for riboswitches.

In the first part of this chapter, we characterized a motif that we call *guanidine-IV variant* (*Gd4v*). The motif was identified using comparative sequence analysis and is almost identical with the *guanidine-IV* riboswitch in its primary and secondary structure. In binding assays, we showed that the *Gd4v* motif undergoes structural modulation in response to guanidine. The *Gd4v* motif however, binds guanidine with a very low affinity, suggesting that guanidine is not its natural ligand, but somehow structurally related to it. Despite the low affinity, we demonstrated that the binding is indeed specific and relays on the presence of some highly conserved nucleotides. In transcription termination assays as well as in an *in vivo* reporter system, no evaluable effect in response to guanidine or a guanidine analog was observed. Our results on the *Gd4v* motif confirm our initial hypothesis that the motif is a riboswitch, action as transcriptional ON-switch that activates gene expression in response to a guanidine derivate.

The second motif characterized, is the *ablB* motif. It is found upstream of various transporters as well as the two enzymes lysine-2,3-aminomutase and β-lysine-*N*-acetyltransferase. These proteins are likely involved in the biosynthesis of N^{ε}-acetyl-β-lysine, a molecule that was shown to be used intracellular for osmoadaption[129]. Based on that, we speculated that the *ablB* motif functions as an ON-switch that activates gene expression in response to osmotic stress. Using transcription termination assay, we could show that the motif indeed likely functions as ON-switch, since in the absence of potential ligands, almost no “full-length” product was formed. Additionally, it was shown that the *ablB* motif RNA undergoes a modulation in the presence of high potassium and sodium concentrations. This modulation was observed using the *ablB* motif sequence of three different organisms and did not occur on unrelated motif RNAs. The band intensity decreased at highly conserved positions in the internal bulge of the third stem. However, mutation of these nucleotides did not diminish modulation. The *ablB* motif is a promising candidate for a riboswitch that is involved in osmotic stress and should further be characterized using an *in vivo* reporter system. Additionally, bacterial extract fractions of cultures that were grown under high osmotic pressure, could be tested binding assays and transcription termination assays. Assuming that the cognate ligand is present when osmotic stress occurs, activity for the responding fraction should be observed.

In the last part of this chapter the *folP* RNA motif was characterized. This motif was identified as riboswitch candidate in 2017 by Weinberg *et al.*[77] via comparative sequence analysis of 5’-UTRs. Almost all representatives of the *folP* motif RNAs are followed by an intrinsic terminator stem, making it likely that its regulation mechanism is based on intrinsic transcription termination. Multiple riboswitches that have been previously identified and found to use terminators for gene regulation, function as

ON-switches[80,99,25]. The *folP* motif is always associated with a gene encoding for a dihydropteroate synthase, that is involved in biosynthesis of THF. Given this information, we hypothesized that the *folP* RNA motif is a riboswitch regulating its downstream genes on the level of transcription termination in response to a metabolite that plays a role in folate synthesis and functioning as ON-switch. A *folP*-motif:*lacZ* reporter strain, cultivated in the presence of the antifolate trimethoprim showed an increased reporter activity, indicating that the *folP* motif upregulated its downstream gene. Trimethoprim inhibits the dihydrofolate synthase and thus leads to a depletion of tetrahydrofolate[144]. Speculating that *folP* acts as an ON-switch, it seems likely that the cognate ligand for the riboswitch candidate *folP* is accumulated in the presence of trimethoprim. This *in vivo* results give evidence that the *folP* motif is indeed involved in the regulation of folate biosynthesis.

Since riboswitches are involved in controlling metabolic pathways and transport, it seems likely that their cognate ligands are not present in the cell under all conditions and growth phases with the same concentration. Thus, to identify ligands, bacterial strains could be cultivated under different conditions and life styles, respectively. Identification of cultivation conditions where the cognate ligand is most abundant also provides an insight into the function of the process controlled by the riboswitch. For all three orphan riboswitch motifs, no activity in bacterial extract screening was observed. It might be helpful to assay a wide variety of cultivation conditions such as different carbon sources or application of stressors (starvation, temperature, oxidative stress, etc.).

7. Materials and Methods

7.1.Materials

7.1.1. Buffers and media

Table 3: Buffers prepared and used. All chemicals were purchased either from Sigma-Aldrich or Roth.

Name	ingredient	final concentration
2x RNA denaturing PAGE loading buffer	Formamide	80 % (v/v)
	Bromphenol blue	0.5 % (w/v)
	Xylene cyanol	0.5 % (w/v)
2x In-line probing reaction buffer (pH 8.3)	Tris	100 mM
	$MgCl_2$ x 6 H_2O	40 mM
	KCl	200 mM
10x Na2CO3 buffer (pH 9)	Na_2CO_3	500 mM
10x Sodium citrate buffer (pH 5)	Tris-sodium citrate	250 mM
2x Urea loading buffer	Urea	9 M
	Sucrose	20 % (w/v)
	SDS	% (w/v)
	Bromphenol blue	0.05 % (w/v)
	Xylene cyanol	0.05 % (w/v)
Permeabilisation Buffer	Na_2HPO_4	100 mM
	KCl	20 mM
	$MgSO_4$	2 mM
	CTAB	0.8 mg/ml
	Sodium deoxycholate	0.4 mg/ml
	β-mercaptoethanol	5.4 % (v/v)
Substrate solution	Na_2HPO_4	60 mM
	Na_2HPO_4 x 6 H_2O	40 mM
	ONPG	1 mg/ml
	β-mercaptoethanol	2.7 % (v/v)
Stop solution	Na_2CO_3	1 M
6x Agarose gel loading buffer	Glycerol	30 % (v/v)

	Bromphenol blue	0.25 % (w/v)
	Xylene cyanol	0.25 % (w/v)
1x TBE buffer	Tris Base	10.8 g/l
	Boric acid	5.5 g/l
	EDTA (pH 8.0)	2 mM
Crush-Soak buffer (pH 7.5)	NaCl	200 mM
	EDTA	1 mM
	HEPES	10 mM
10X TBE, 9 M urea	Tris base	108 g/l
	Boric acid	55 g/l
	EDTA (pH 8.0)	20 mM
	Urea	540 g/l

Table 4: List of bacterial growth media used. All chemicals were purchased from either Sigma-Aldrich or Roth. MiliQ water was used as solvent.

Name	**ingredient**	**final concentration**
Lysogenic broth (LB) medium, liquid	Yeast extract	5 g/l
	Tryptone	10 g/l
	NaCl	10 g/l
Lysogenic broth (LB) medium, solid	Yeast extract	5 g/l
	Tryptone	10 g/l
	NaCl	10 g/l
	Agar-agar	20 g/l
M9 minimal medium	Na_2HPO_4 x 2H2O	8.5 g/l
	KH2PO4	3 g/l
	NaCl	0.5 g/l
	NH4Cl	1 g/l
	MgSO4 x 7H2O	1 mM
	CaCl2	100 µM
	Glucose	0.2 % (v/v)
	Vitamin solution	1 ml/l
Super optimal broth with catabolite repression (SOC medium), liquid	Yeast Extract	5 % (w/v)
	Tryptone	0.2 % (w/v)
	NaCl	10 mM

	KCl	2.5 mM
	$MgCl_2$	10 mM
	$MgSO_4$	10 mM
	Glucose	20 mM
M9, no C-source, liquid	$NaHPO_4$ x $2H_2O$	8.5 g/l
	KH_2PO_4	3 g/l
	NaCl	0.5 g/l
	NH_4Cl	1 g/l
	$MgSO_4$ x 7 H_2O	1 mM
	$CaCl_2$	100 µM
	Vitamin solution	1 ml/l

7.1.2. Enzymes and Kits

Table 5: enzymes and enzyme buffers.

enzyme	**supplier**
Phusion DNA polymerase	New England Biolabs (NEB)
Q5 DNA Polymerase	NEB
T7 RNA Polymerase	Fermentas
T5 *E. coli* RNA Polymerase	NEB
rSAP	NEB
T4 Quick Ligase	NEB
RNase inhibitor	Fermentas
PPase	Thermo Scientific
T4 polynukleotide kinase	Thermo Scientific
NdeI	NEB
XhoI	NEB
enzyme buffer	**supplier**
5x HF-buffer	Thermo Scientific
5x GC-buffer	Thermo Scientific
2x Quick ligase buffer	NEB
10x CutSmart	NEB
PNK A Buffer	Thermo Scientific
T5 RNA Polymerase Buffer	NEB

Table 6: List of Kits.

Kit	supplier
Zymo DNA Clean & Concentrator TM Kir	Zymo Research
Zymo Gel DNA Recovery Kit	Zymo Research
Zyppy Plasmid MiniPrep Kit	Zymo Research
DNeasy Blood & Tissue Kit	Quiagen

7.1.3. Strains

Escherichia coli XL 10 Gold (Stratagene) strain was used for plasmid expression; Genotype: TetrΔ(*mcrA*)183;Δ(*mcrCB*-hsdSMR-mrr)173; *endA*1;*supE*44; *thi*-1 *recA*1; *gyrA*96; *relA*1; lac; *Hte* [F'; *proAB*; *lacI*$_q$ZΔM15;Tn10;(Tet$_r$); Amy; Cam$_r$].

Escherichia coli BW25113 strain was used for *lacZ* reporter gene studies; Genotype: F-, Δ(*araD-araB*)567, *ΔlacZ*4787(::rrnB-3), λ-, rph-1, *Δ(rhaD-rhaB)568, hsdR514*

Staphylococcus aureus RN4220 (Rosenbach 1884) DSM 26309 was used for controllable gene expression studies.

7.1.4. Plasmids

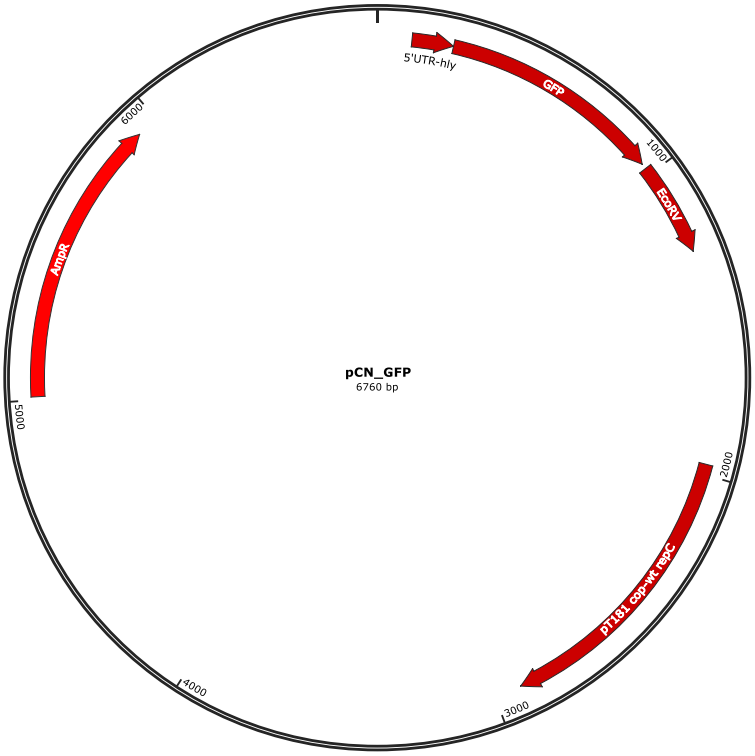


Figure 24: pCN-pBlaZ-GFP vector. Map of plasmid used for *gfp* expression in *S. aureus*. RNA motifs were inserted into the *hly* 5´-UTR. Provided by the Romby lab (University of Strasbourg).

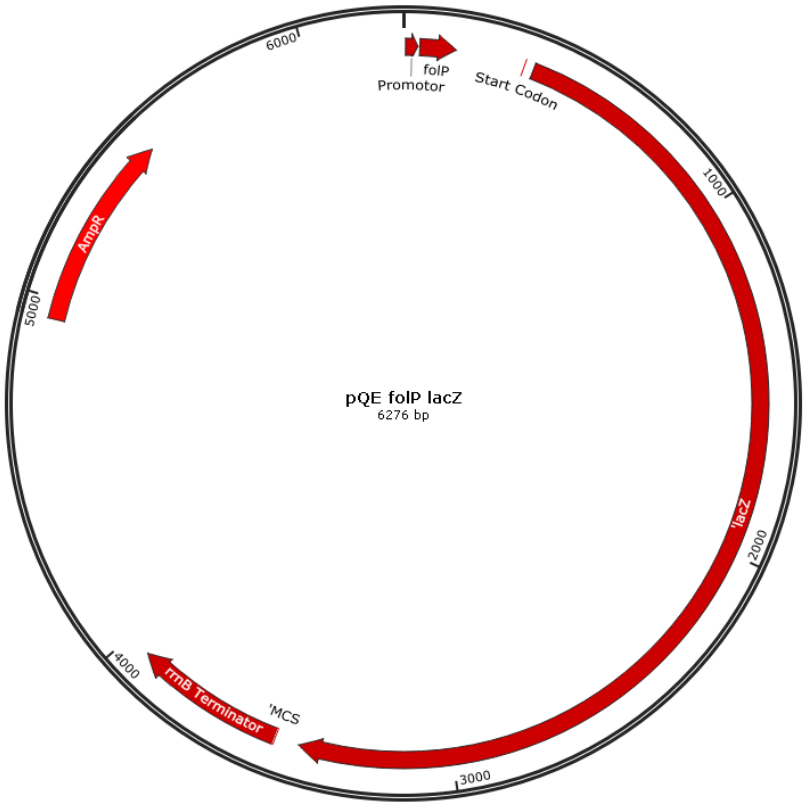


Figure 25: pQE-folP-lacZ vector. Map of pQE31 derived vector containing J06 promoter, *folP* RNA motif and *lacZ* used for reporter experiments in *E. coli* BW25113.

7.2.Methods

7.2.1. General Methods

Phusion or Q5 DNA Polymerase PCR

To amplify DNA from genomic DNA or a plasmid or to insert sequences into a plasmid via whole plasmid PCR, *Phusion* DNA polymerase or Q5 DNA polymerase were used. The reaction mixture as shown in Table x was run according to the protocols shown in Table 9 (Phusion) or Table 10 (Q5).

Table 7: Standard reaction setup for Phusion or Q5 PCR.

Reagent	Volume [µL]	Stock conc.	Final conc.
HF/GC Buffer	10	5x	1x
dNTPs	5	2 mM	200 µM
Primer A	6	5 µM	600 nM
Primer B	6	5 µM	600 nM
Template	1	10 ng/µl	0.2 ng/µl
Phusion polymerase	0.5	2 U/µl	1 U
H_2O	21.5		
total	**50**		

The reaction was performed in a thermocycler using the following PCR protocol:

Table 8: Standard thermocycler protocol for Phusion DNA polymerase PCR.

	Temperature [°C]	time	
Initial denaturation	98	30 s	
Denaturation	98	10 s	25 cycles
Annealing	58-69	30 s	
Extension	72	20 s/kbp	
Final extension	72	7 min	

Table 9: Standard thermocycler protocol for Q5 DNA polymerase PCR.

	Temperature [°C]	time	
Initial denaturation	98	30 s	
Denaturation	98	10 s	35 cycles
Annealing	58-69	30 s	
Extension	72	25 s/kbp	
Final extension	72	2 min	

DpnI digest

To remove methylated template DNA following PCR amplification the product was digested using restriction endonuclease DpnI. 1 µL of DpnI enzyme was added to the PCR reaction mixture in CutSmart buffer and incubated at 37 °C for 20 min.

Restriction enzyme digest and dephosphorylation

Larger DNA sequences were inserted into a plasmid backbone using restriction enzymes. Both the insert and the backbone were digested with the same restriction enzymes to generate complementary 5´-overhangs at the restriction sites. 2 µg of DNA or a whole PCR product was digested in a reaction volume of 50 µL containing 1 µL of the used enzymes in the respective buffer provided by the manufacturer. The reaction was incubated for 2 hours at 37 °C. The digestion reaction of the backbone DNA was followed by a dephosphorylation step to prevent unwanted side products during ligation. For dephosphorylation 1 µL of Shrimp Alkaline Phosphatase (rSAP) in CutSmart buffer was used in a reaction volume of 50 µL and incubated for 1 h at 37 °C. The dephosphorylated backbone DNA and the inserts were then purified using a column (<500 bp) or agarose gel electrophoresis.

Ligation

For a ligation the reaction mixture shown in table x was incubated at 25 °C for 20 min.

Table 10: Quick Ligase protocol for whole plasmid PCR ligation.

Reagent	Volume [µL]	Stock	Final conc.
Quick Ligase Buffer	5	2x	1x
DNA	4.5	25 ng	2.5 ng/µl
Quick Ligase	0.5		

Overlap extension using reverse transcription

To obtain DNA templates for *in vitro* transcription overlap extension with reverse transcription was performed with the following setup:

Table 11: reaction setup for overlap extension reaction.

Reagent	Volume [µL]	Stock conc.	Final conc.
FSB	10	5x	1x
dNTPs	5	10 mM	1 mM
Primer A	3	100 µM	6 µM
Primer B	3	100 µM	6 µM
DTT	5	0.1 M	10 mM
H_2O	22		
total	**50**		

The mixture was incubated at 90 °C for 1 min and then let stand at RT for 3 min. 2 µl of SS II RT enzyme were added and incubated for one hour at 50 °C.

Agarose gel electrophoresis

PCR and overlap extension products were analyzed by electrophoresis using standard agarose gels. Agarose was solved in 0.5% TBE and 1µl Midori Green Advanced was added for DNA staining. DNA samples with loading dye were loaded together with the respective DNA ladders onto the gels. For qualitative agarose gels the gels were run at 120 V for 45 min. For preparative agarose gels they were run at 100 V for 90 min. Detection of DNA bands was conducted under UV-light.

DNA Isolation and Purification Kits

PCR products were directly purified with Zymo DNA Clean & Concentrator™ Kit using the manufacturer's standard protocol, eluting purified DNA with MiliQ ultra-pure water.

PCR products separated by agarose gel electrophoresis were treated with Zymoclean™ Gel DNA Recovery Kit for purification using the manufacturer's standard protocol. Elution was conducted in MiliQ ultra-pure water.

Plasmid DNA was isolated from *E. coli* cultures applying the standard protocol of the Zyppy™ Plasmid MiniPrep Kit and eluted with MiliQ ultra-pure water.

Genomic DNA was isolated from bacteria using the DNeasy® Blood & Tissue Kit first applying the protocol for "Pretreatment for Gram-Negative bacteria" to lyse bacterial cells. Afterwards the standard protocol "Purification of Total DNA from Animal Tissues (Spin-column Protocol)" was further carried out. gDNA was eluted with MiliQ ultra-pure water.

DNA and RNA concentration determination

DNA/RNA concentration was determined by analysing 2 µl of samples on a NanoQuant plate using plate reader Tecan infinite M200.

Sequencing

Sequencing of DNA samples was conducted by GATC Eurofins costume DNA Sanger-Sequencing.

Gibson Assembly

For the insertion of a fragment of 100 bp to 10 kbp, a Gibson Assembly was performed. The backbone and the insert with overhangs complementary to the backbone, were amplified via Phusion PCR. Both were analyzed and purified with gel electrophoresis and Zymo DNA Clean & Concentrator™ Kit, respectively. A Gibson Assembly Master Mix containing T5 Exonuclease, Phusion DNA Polymerase and Taq DNA Ligase was mixed with backbone using the following setup:

Table 12: Gibson Assembly Protocol.

Reagent	Volume [µL]	Stock	Final conc.
backbone	-	50 ng	2.5 ng/µl
insert	-	17.5 ng	0.875 ng/µl
Gibson Assembly Master Mix	10	2 x	1 x

After incubation at 50 °C for 20 min, 2 µl were used to be transformed into electrocompetent cells.

Transformation in *E. coli*

1 µl of purified plasmid was given to 80 µl of electrocompetent *E . coli* cells thawed on ice. The mixture was rapidly given to a pre-cooled electroporation cuvette and electroporation was conducted in an Eppendorf 2510 Electroporator. The mixture was transferred from the cuvette into pre-warmed (37 °C) SOC medium. After incubation for 30 min the SOC culture was plated on LB-Agar plates containing selective antibiotics.

Precipitation of RNA and DNA

DNA or RNA samples were supplemented with 1/10 volumes of 3 M NaAcetate and 3 volumes of 100 % pre-cooled (-80 °C) Ethanol and incubated for 1h at -80 °C. Samples were centrifuged for 20 min at 15,000 rpm at 4 °C. The supernatant was discarded and the pellet was resuspended in MiliQ ultrapure water.

in vitro transcription

Table 13: *in vitro* transcription protocol.

Reagent	Volume [µL]	Stock	Final conc.
NTPs	15	10 mM	1 mM
Transcription buffer	20	5x	1x
RiboLock	0.75	-	-
PPase	1	-	-
T7 RNA polymerase	2.5	-	-
Template DNA	16	From overlap reaction	0.96 µM
Water	39.75		
Total	100		

In vitro transcription mixture was incubated at 37 °C for 3h, stopped by adding denaturing loading dye and purified over a PAGE gel.

Polyacrylamide gel electrophoresis (PAGE)

Polyacrylamide gel electrophoresis (PAGE) was used for separation and purification of RNA. Preparation of a 1.5 mm 10% denaturing PAGE gel was conducting using the following protocol:

Table 14: Protocol of a 10% denaturing PAGE gel.

Name	Ingredients	Final concentration
Denaturing PAGE Gel (10 %)	Urea	6.34 M
	10x TBE, Urea Buffer	1x
	10 % APS	0.08 %
	TEMED	0.04 %

After the Gel was completely polymerized for at least 1.5 h, the RNA samples containing denaturating PAGE loading dye were loaded on the gel. Gel was run at 65 W for app. 2 h. RNA was visualized by UV shadowing or radiograph. Identified RNA band was cut out and crushed manually before suspending with Crush-Soak buffer. The mixture was incubated for 2 h at RT and then filtered through glass wool. The flow through containing RNA was ethanol precipitated.

SDS PAGE

SDS PAGE was used to analyze protein samples. A 1 mm gel was prepared consisting of a stacking and resolving gel. The stacking gel contained of 10-16 % (w/v) Rotiphorese Gel 40 (37.5:1), 250 mM Tris-HCl (pH 8.8), 0.1% SDS (w/v), 0.1% APS (w/v) and 0.1% TEMED (w/v), was overlaid with 100%

isopropanol and allowed to polymerize for at least 30 min. Isopropanol was removed and a resolving gel consisting of 5% Rotiphorese Gel 40 (37.5:1), 125 mM Tris-HCl (pH 6.8), 0.1% SDS (w/v), 0.1% APS (w/v) and 0.1% TEMED (w/v) was prepared and given on top of the stacking gel. After polymerisation of the resolving gel protein samples were loaded and the gel was run for a maximum of 2 h at 120 V in 1x Rotiphorese SDS-PAGE buffer. Gel electrophoresis was conducted using BioRad SDS gel electrophoresis equipment.

7.2.2. In-line probing

DNA templates for in-line probing reactions were generated using T7 promoter containing primers either via PCR amplification from genomic DNA or via overlap extension reaction. DNA templates were purified via Zymo DNA Clean & Concentrator™ Kit and *in vitro* transcription reactions were performed using T7 RNA polymerase. The RNA was purified by PAGE gel.

Dephosphorylation

2 µl of Cut Smart Buffer were added to 70 – 100 pmol RNA and incubated for 1 min at 65 °C. The mixture was cooled down on ice. 2 µl rSAP and water (to to 20 µl) were added and incubated for 1 h at 37 °C. To inactivate the phosphatase, everything was incubated for another 8 min at 75 °C.

Kination

Table 15: [^{32}P]-kination of RNA for in-line probing reaction.

Reagent	Volume [µL]	Stock	Final conc.
RNA	10	From dephosphorylation reaction	-
T4 PNK A Buffer	2	10x	1x
T4 PNK	2.5	200 U/µl	25 U/µl
γ-[^{32}P]-ATP	3	-	-
RiboLock	0.5	-	-
PPase	0.5	-	-
Water	1.5		
Total	20		

The kination mixture was incubate for 1 h at 37 °C. The reaction was stopped with denaturating PAGE loading dye and purified by PAGE gel. After Crush-Soak extraction and RNA precipitation the pellet was solved in water to obtain a concentration of 1 kBq/µl.

In-line probing reaction

In-line probing reactions were set up in 20 µL. The ^{32}P-labeled RNA was incubated for 40 h at 25 °C using the following setup:

Table 16: protocol for in-line probing reactions.

Reagent	Volume [µL]	Stock	Final conc.
RNA	1	1 kBq	-
in-line probing buffer	5	2 x	1 x
water	3		
ligand	1		

For nucleotide assignment partial digestion of the RNA with T1 RNase and alkaline digestion was performed. For RNase T1 digestion 1 kBq RNA was incubated for 8 min in 1x sodium citrate buffer with 1x Urea loading buffer and 1 U T1 RNase at 50°C. For alkaline digestion 1 kBq RNA was incubated for 5 min in 1x Na_2CO_3 buffer at 95°C.

In-line probing reactions and controls were analyzed together by analytical PAGE and imaged using a Typhoon FLA 7000 phosphorimager. Modulated bands were quantified with ImageQuant.

7.2.3. Transcription termination assay

DNA templates containing the T5 promotor followed by the riboswitch motif, natural terminator and about additional 150 nt, were amplified using PCR. PCR products were purified with Zymoclean™ Gel DNA Recovery Kit and diluted to a concentration of 100 ng/µl. 1 µl of the DNA template were mixed on ice with 1 µl of the ligand. The transcription termination reaction was set up using the following protocol:

Table 17: protocol for multiple round transcription assay.

Reagent	Volume [µL]	Stock	Final conc.
DNA template	1	100 ng/µl	10 ng/µl
ligand	1	-	-
Polymerase buffer	2	5 x	1 x
NTP´s (GUC)	0.75	25 mM	1.875 mM
ATP	1	1 mM	0.1 mM
α-[^{32}P]-ATP	0.2	-	-
RiboLock	0.1	-	-

PPase	0.1	-	-
T5 Polymerase	0.2	-	-
Water	3.65		
Total	10		

The reaction was incubated for 10 min at 37 °C and stopped with denaturing loading dye. In order to assign full length product and termination product a DNA template was designed that was identical to the template, but lacking the sequence followed after the terminator. For this short length control the transcription termination assay was performed as described above but using water instead of a ligand. As negative control the transcription termination assay was performed with the full length template but also using water instead of a ligand. The water control, short length control and the other samples were loaded on an analytical PAGE gel. The Gel was imaged using a Typhoon FLA 7000 phosphorimager. Full length product and termination product bands were quantified with ImageQuant.

7.2.4. Preparation of bacterial extracts

Bacterial extract fractions were used to screen riboswitch candidates in the transcription termination assay and in in-line probings. *E. coli* and *B. subtilis* cells grown in M9 medium were centrifuged and washed with -80 °C-cold methanol. This washing step was repeated two times. The methanol extract was lyophilized, dissolved in water and fractionated by HPLC over a reverse-phase C18 column. The obtained fractions were lyophilized and re-suspended in water. 4 fractions were pooled, respectively.

7.2.5. Reporter construction and growth condition

Genetic reporter plasmids were designed by insertion of the *folP*:RNA motif from Dialister invisus including the first eight codons of the downstream ORF into an pQE31 derived vector. The pQE31 vector contains an modified version of the artificial and constitutive active J06 promoter (modified from the Anderson promoter library: http://parts.igem.org/Promoters/Catalog/Anderson) and *lacZ* reporter gene. The reporter plasmid (WT) was transformed into *E. coli* BW25113 via electroporation. Reporter assays were conducted by inoculating the reporter strain in LB medium and in M9 medium containing 0.4% glucose in the presence of different concentrations. Mutated reporter plasmids were constructed via whole plasmid PCR using primer pairs x (Mut1), x (Mut2), x (Mut3) and were transformed into *E. coli* BW25113. WT and Mut1-5 reporter strain were inoculated under the same

conditions. All reporter assays were conducted in a volume of 5 ml in 50 ml falcons incubated overnight at 37 °C in a rotational shaker at 200 rpm.

7.2.6. ONPG Assay

Reporter cultures were diluted 1:10 in LB medium or M9 medium and OD_{600} was measured. 10 µl of diluted cultures were added to 40 µl of Permeabilisation Solution. Pre-incubated at 30 °C for 30 min. In order to start reporter enzyme activity 300 µl of 30 °C-warm Substrate Solution were added. Enzyme reaction was stopped by addition of 350 µl 1 M Na_2CO_3 (Stop Solution) and reaction time was monitored. The solution was analyzed recording absorption at 420 nm in a plate reader (Tecan infinite M200). Reporter activities were calculating in Miller Units using the following equation:

(1) $$\text{Miller Units} = 1000 * \frac{Abs._{420}}{((Abs._{600})*(sampe\ volume\ [0.01ml]*(reaction\ time\ [sec]))}$$

7.2.7. *S. aureus* experiments

Cultivation of *S. aureus*

S. aureus were cultivated in liquid cultures (BHI-medium) at 37 °C and 200 rpm or on BHI-agarplates at 37 °C.

Electrocompetent *S. aureus*

350 mL BHI medium was inoculated with an overnight culture of *S. aureus* to an initial OD_{600} of approximately 0,01. The bacteria were grown for 2 h at 37 °C under agitation until reaching an OD_{600} of 0.4-0.5. 300 mL of the liquid culture were transferred to 50 mL falcons and centrifuged at 3000 g for 15 min at 4 °C. The supernatant was discarded carefully. The pellets were pooled and resuspended in 60 mL of a sterilized 0.5 M sucrose solution. The solution was centrifuged for 10 min at 3000 g at 4 °C. The supernatant was discarded and the pellet was resuspended in 30 mL of a 10 % glycerol solution followed by centrifugation for 5 min at 3000 g at 4 °C. The supernatant was discarded and the pellet was resuspended in 3 mL of a 10 % glycerol solution. The competent cells were stored at -80 °C in 100 µL aliquots.

Transformation in *S. aureus*

Electrocompetent *S. aureus* cells were thawed on ice and incubated with 1 µg of nonmethylated DNA in a volume of no more than 10 µL for 30 min. Cells were transferred to electroporation cuvettes with a 2 mm gap and pulsed with 1.8 kV for 2.5 ms using a Gene Pulser. The electroporated cells were quickly resuspended in 900 µL of pre-warmed BHI medium and incubated for 2 h at 37 °C under agitation. The cells were then transferred to an Eppendorf tube and centrifuged at 3000 g for 2 min.

The supernatant was discarded and the pellet was resuspended in remaining supernatant and spread on BHI plates containing 10 µg/mL erythromycin. The plates were incubated overnight at 37 °C.

eGFP fluorescence measurements

Three single colonies for each transformed plasmid were cultivated in 400 µL BHI-Medium in a 96-deepwell plate. The plates were incubated overnight at 37 °C at 1300 rpm. For technical triplicates, 10 µL of an overnight culture were added to 400 µL of fresh BHI-medium in a 96-deepwell plate. The plates were incubated overnight at 37 °C at 1300 rpm. Of each culture 100 µL were transferred to a UV transparent flat bottomed 96-well plate. eGFP fluorescence intensity and OD_{600} measurements were performed using a tecan plate reader. For eGFP measurements the excitation wavelength was set to 488 nm and emission wavelength to 535 nm. eGFP fluorescence was normalized to OD_{600}.

8. References

(1) Crick, F. H. C., Fenton P D Cotterill (1958) On protein synthesis.

(2) Crick, F. (1970) Central Dogma of Molecular Biology, *Nature 227,* 561–563. DOI: 10.1038/227561a0.

(3) Watson, J. D., and Baker, T. A. (2014) Molecular biology of the gene. 7th ed., Pearson; Cold Spring Harbor Laboratory Press, Boston, Mass., Cold Spring Harbor, NY.

(4) Pan, T., and Sosnick, T. (2006) RNA folding during transcription, *Annual review of biophysics and biomolecular structure 35. DOI:* 10.1146/annurev.biophys.35.040405.102053.

(5) GROS, F., GILBERT, W., HIATT, H. H., Attradi, G., SPAHR, P. F., and WATSON, J. D. (1961) Molecular and biological characterization of messenger RNA, *Cold Spring Harbor symposia on quantitative biology 26. DOI:* 10.1101/sqb.1961.026.01.016.

(6) Pelechano, V., and Steinmetz, L. M. (2013) Gene regulation by antisense transcription, *Nature reviews. Genetics 14. DOI:* 10.1038/nrg3594.

(7) Serganov, A., and Nudler, E. (2013) A decade of riboswitches, *Cell 152,* 17–24. DOI: 10.1016/j.cell.2012.12.024.

(8) Waters, L. S., and Storz, G. (2009) Regulatory RNAs in bacteria, *Cell 136. DOI:* 10.1016/j.cell.2009.01.043.

(9) Banerjee, D., and Slack, F. (2002) Control of developmental timing by small temporal RNAs: a paradigm for RNA-mediated regulation of gene expression, *BioEssays : news and reviews in molecular, cellular and developmental biology 24. DOI:* 10.1002/bies.10046.

(10) Winkler, W., Ali Nahvi, and Ronald R. Breaker (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression, *Nature 419,* 952–956. DOI: 10.1038/nature01145.

(11) Winkler, W. C., Cohen-Chalamish, S., and Breaker, R. R. (2002) An mRNA structure that controls gene expression by binding FMN, *Proceedings of the National Academy of Sciences of the United States of America 99,* 15908–15913. DOI: 10.1073/pnas.212628899.

(12) Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L., and Breaker, R. R. (2002) Genetic control by a metabolite binding mRNA, *Chemistry & biology 9,* 1043. DOI: 10.1016/s1074-5521(02)00224-7.

(13) Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., Perumov, D. A., and Nudler, E. (2002) Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria, *Cell 111,* 747–756. DOI: 10.1016/S0092-8674(02)01134-0.

(14) Mandal, M., and Breaker, R. R. (2004) Gene regulation by riboswitches, *Nature reviews. Molecular cell biology 5,* 451–463. DOI: 10.1038/nrm1403.

(15) Winkler, W. C., and Breaker, R. R. (2005) Regulation of bacterial gene expression by riboswitches, *Annual review of microbiology 59,* 487–517. DOI: 10.1146/annurev.micro.59.030804.121336.

(16) Serganov, A., and Dinshaw J Patel (2007) Serganov, A. & Patel, D.J. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. Nat. Rev. Genet. 8, 776-790, *Nature Reviews Genetics 8,* 776–790. DOI: 10.1038/nrg2172.

(17) Henkin, T. M. (2008) Riboswitch RNAs: using RNA to sense cellular metabolism, *Genes & Development 22,* 3383–3390. DOI: 10.1101/gad.1747308.

(18) Winkler, W. C., Nahvi, A., Sudarsan, N., Barrick, J. E., and Breaker, R. R. (2003) An mRNA structure that controls gene expression by binding S-adenosylmethionine, *Nature structural biology 10,* 701–707. DOI: 10.1038/nsb967.

(19) Corbino, K. A., Barrick, J. E., Lim, J., Welz, R., Tucker, B. J., Puskarz, I., Mandal, M., Rudnick, N. D., and Breaker, R. R. (2005) Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria, *Genome biology 6,* R70. DOI: 10.1186/gb-2005-6-8-r70.

(20) Ames, T. D., Rodionov, D. A., Weinberg, Z., and Breaker, R. R. (2010) A Eubacterial Riboswitch Class That Senses the Coenzyme Tetrahydrofolate, *Chemistry & biology 17,* 681–685. DOI: 10.1016/j.chembiol.2010.05.020.

(21) Mandal, M., Boese, B., Barrick, J. E., Winkler, W. C., and Breaker, R. R. (2003) Riboswitches Control Fundamental Biochemical Pathways in Bacillus subtilis and Other Bacteria, *Cell 113,* 577–586. DOI: 10.1016/S0092-8674(03)00391-X.

(22) Kim, P. B., Nelson, J. W., and Breaker, R. R. (2015) An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism, *Molecular Cell 57,* 317–328. DOI: 10.1016/j.molcel.2015.01.001.

(23) Sherlock, M. E., Sudarsan, N., and Breaker, R. R. (2018) Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems, *PNAS 115,* 6052–6057. DOI: 10.1073/pnas.1720406115.

(24) 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, *Nature chemical biology 9,* 834–839. DOI: 10.1038/nchembio.1363.

(25) 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 (New York, N.Y.) 335,* 233–235. DOI: 10.1126/science.1215063.

(26) Furukawa, K., Ramesh, A., Zhou, Z., Weinberg, Z., Vallery, T., Winkler, W. C., and Breaker, R. R. (2015) Bacterial Riboswitches Cooperatively Bind Ni2+ or Co2+ Ions and Control Expression of Heavy Metal Transporters, *Molecular Cell 57,* 1088–1098. DOI: 10.1016/j.molcel.2015.02.009.

(27) Tucker, B. J., and Breaker, R. R. (2005) Riboswitches as versatile gene control elements, *Current opinion in structural biology 15,* 342–348. DOI: 10.1016/j.sbi.2005.05.003.

(28) Barrick, J. E., and Breaker, R. R. (2007) The distributions, mechanisms, and structures of metabolite-binding riboswitches, *Genome biology 8,* R239. DOI: 10.1186/gb-2007-8-11-r239.

(29) Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., Repoila, F., Buchrieser, C., Cossart, P., and Johansson, J. (2009) A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes, *Cell 139,* 770–779. DOI: 10.1016/j.cell.2009.08.046.

(30) Nudler, E., and Mironov, A. S. (2004) The riboswitch control of bacterial metabolism, *Trends in biochemical sciences 29,* 11–17. DOI: 10.1016/j.tibs.2003.11.004.

(31) Irnov, I., and Winkler, W. C. (2010) A regulatory RNA required for antitermination of biofilm and capsular polysaccharide operons in Bacillales, *Molecular microbiology 76,* 559–575. DOI: 10.1111/j.1365-2958.2010.07131.x.

(32) McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E., and Breaker, R. R. (2017) Riboswitch diversity and distribution, *RNA (New York, N.Y.) 23,* 995–1011. DOI: 10.1261/rna.061234.117.

(33) Sudarsan, N., Barrick, J. E., and Breaker, R. R. (2003) Metabolite-binding RNA domains are present in the genes of eukaryotes, *RNA (New York, N.Y.) 9,* 644–647. DOI: 10.1261/rna.5090103.

(34) Croft, M. T., Moulin, M., Michael E Webb, and Alison G Smith (2007) Thiamine biosynthesis in algae is regulated by riboswitches, *undefined*.

(35) Bocobza, S. E., and Aharoni, A. (2014) Small molecules that interact with RNA: riboswitch-based gene control and its involvement in metabolic regulation in plants and algae, *The Plant journal : for cell and molecular biology 79,* 693–703. DOI: 10.1111/tpj.12540.

(36) Kubodera, T., Mutsumi Watanabe, Kumi Yoshiuchi, Nobuo Yamashita, and Hideo Hanamoto (2004) Thiamine-regulated gene expression of Aspergillus oryzae thiA requires splicing of the intron containing a riboswitch-like domain in the 5'-UTR, *FEBS Letters 555,* 516–520. DOI: 10.1016/S0014-5793(03)01335-8.

(37) Serganov, A., and Patel, D. J. (2012) Molecular recognition and function of riboswitches, *Current opinion in structural biology 22,* 279–286. DOI: 10.1016/j.sbi.2012.04.005.

(38) Serganov, A., Yuan, Y.-R., Pikovskaya, O., Polonskaia, A., Malinina, L., Phan, A. T., Hobartner, C., Micura, R., Breaker, R. R., and Patel, D. J. (2004) Structural Basis for Discriminative Regulation of Gene Expression by Adenine- and Guanine-Sensing mRNAs, *Chemistry & biology 11,* 1729–1741. DOI: 10.1016/j.chembiol.2004.11.018.

(39) Savinov, A., Perez, C. F., and Block, S. M. (2014) Single-molecule studies of riboswitch folding, *Biochimica et biophysica acta 1839. DOI:* 10.1016/j.bbagrm.2014.04.005.

(40) Hammes, G. G., Chang, Y. C., and Oas, T. G. (2009) Conformational selection or induced fit: a flux description of reaction mechanism, *Proceedings of the National Academy of Sciences of the United States of America 106. DOI:* 10.1073/pnas.0907195106.

(41) Suddala, K. C., Wang, J., Hou, Q., and Walter, N. G. (2015) Mg2+ Shifts Ligand-Mediated Folding of a Riboswitch from Induced-Fit to Conformational Selection, *Journal of the American Chemical Society 137,* 14075–14083. DOI: 10.1021/jacs.5b09740.

(42) Suddala, K. C., and Walter, N. G. (2014) Riboswitch structure and dynamics by smFRET microscopy, *Methods in enzymology 549,* 343–373. DOI: 10.1016/B978-0-12-801122-5.00015-5.

(43) Haller, A., Soulière, M. F., and Micura, R. (2011) The dynamic nature of RNA as key to understanding riboswitch mechanisms, *Accounts of Chemical Research 44. DOI:* 10.1021/ar200035g.

(44) Wickiser, J. K., Wade C. Winkler, Ronald R. Breaker, and Donald M. Crothers (2005) The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch, *Molecular Cell 18,* 49–60. DOI: 10.1016/j.molcel.2005.02.032.

(45) Wickiser, J. K., Cheah, M. T., Breaker, R. R., and Crothers, D. M. (2005) The Kinetics of Ligand Binding by an Adenine-Sensing Riboswitch, *Biochemistry 44,* 13404–13414. DOI: 10.1021/bi051008u.

(46) Lemay, J. F., Penedo, J. C., Tremblay, R., Lilley, D. M., and Lafontaine, D. A. (2006) Folding of the adenine riboswitch, *Chemistry & biology 13. DOI:* 10.1016/j.chembiol.2006.06.010.

(47) Rieder, R., Lang, K., Graber, D., and Micura, R. (2007) Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control, *Chembiochem : a European journal of chemical biology 8. DOI:* 10.1002/cbic.200700057.

(48) Salvail, H., Balaji, A., Yu, D., Roth, A., and Breaker, R. R. (2020) Biochemical Validation of a Fourth Guanidine Riboswitch Class in Bacteria, *Biochemistry 59,* 4654–4662. DOI: 10.1021/acs.biochem.0c00793.

(49) Zhang, J., Lau, M. W., and Ferré-D'Amaré, A. R. (2010) Ribozymes and riboswitches: modulation of RNA function by small molecules, *Biochemistry 49. DOI:* 10.1021/bi1012645.

(50) Xayaphoummine, A., Viasnoff, V., Harlepp, S., and Isambert, H. (2007) Encoding folding paths of RNA switches, *Nucleic Acids Res 35,* 614–622. DOI: 10.1093/nar/gkl1036.

(51) Garst, A. D., Edwards, A. L., and Batey, R. T. (2011) Riboswitches: structures and mechanisms, *Cold Spring Harbor perspectives in biology 3 published online* Jun 1, 2011. DOI: 10.1101/cshperspect.a003533.

(52) Al-Hashimi, H. M., and Walter, N. G. (2008) RNA dynamics: It is about time, *Current opinion in structural biology 18,* 321–329. DOI: 10.1016/j.sbi.2008.04.004.

(53) Chauvier, A., Picard-Jean, F., Berger-Dancause, J.-C., Bastet, L., Naghdi, M. R., Dubé, A., Turcotte, P., Perreault, J., and Lafontaine, D. A. (2017) Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation, *Nature Communications 8. DOI:* 10.1038/ncomms13892.

(54) Wickiser, J. K., Wade C. Winkler, Ronald R. Breaker, and Donald M. Crothers (2005) The Speed of RNA Transcription and Metabolite Binding Kinetics Operate an FMN Riboswitch, *Molecular Cell 18,* 49–60. DOI: 10.1016/j.molcel.2005.02.032.

(55) Strobel, E. J., Cheng, Katherine E. Berman, Paul D. Carlson, and Julius B. Lucks (2019) A ligand-gated strand displacement mechanism for ZTP riboswitch transcription control, *Nat Chem Biol 15,* 1067–1076. DOI: 10.1038/s41589-019-0382-7.

(56) Bastet, L., Turcotte, P., Wade, J. T., and Lafontaine, D. A. (2018) Maestro of regulation: Riboswitches orchestrate gene expression at the levels of translation, transcription and mRNA decay, *RNA biology 15,* 679–682. DOI: 10.1080/15476286.2018.1451721.

(57) Roth, A., and Breaker, R. R. (2009) The structural and functional diversity of metabolite-binding riboswitches, *Annual review of biochemistry 78,* 305–334. DOI: 10.1146/annurev.biochem.78.070507.135656.

(58) Breaker, R. R. (2018) Riboswitches and Translation Control, *Cold Spring Harbor perspectives in biology 10. DOI:* 10.1101/cshperspect.a032797.

(59) Roberts, J. W. (2019) Mechanisms of Bacterial Transcription Termination, *Journal of Molecular Biology 431,* 4030–4039. DOI: 10.1016/j.jmb.2019.04.003.

(60) Ray-Soni, A., Bellecourt, M. J., and Landick, R. (2016) Mechanisms of Bacterial Transcription Termination: All Good Things Must End, *Annual review of biochemistry 85,* 319–347. DOI: 10.1146/annurev-biochem-060815-014844.

(61) Hollands, K., Proshkin, S., Sklyarova, S., Epshtein, V., Mironov, A., Nudler, E., and Groisman, E. A. (2012) Riboswitch control of Rho-dependent transcription termination, *Proceedings of the National Academy of Sciences of the United States of America 109,* 5376–5381. DOI: 10.1073/pnas.1112211109.

(62) Gusarov, I., and Nudler, E. (2001) Control of intrinsic transcription termination by N and NusA: the basic mechanisms, *Cell 107,* 437–449. DOI: 10.1016/s0092-8674(01)00582-7.

(63) Yarnell, W. S., and Roberts, J. W. (1999) Mechanism of Intrinsic Transcription Termination and Antitermination, *Science 284,* 611–615. DOI: 10.1126/science.284.5414.611.

(64) Gusarov, I., and Nudler, E. (1999) The Mechanism of Intrinsic Transcription Termination, *Molecular Cell 3,* 495–504. DOI: 10.1016/S1097-2765(00)80477-3.

(65) Sinn, M. (2019) New assays for riboswitch ligand discovery and a novel lysine degradation pathway in bacteria.

(66) Mehdizadeh, A. E., Hejazi, M. S., and Barzegar, A. (2016) Riboswitches: From living biosensors to novel targets of antibiotics, *Gene 592. DOI:* 10.1016/j.gene.2016.07.035.

(67) Parashar, A. (2016) Aptamers in Therapeutics, *Journal of clinical and diagnostic research : JCDR 10. DOI:* 10.7860/JCDR/2016/18712.7922.

(68) Kourlas, H., and Schiller, D. S. (2006) Pegaptanib sodium for the treatment of neovascular age-related macular degeneration: a review, *Clinical therapeutics 28. DOI:* 10.1016/j.clinthera.2006.01.009.

(69) Eugene W. M. Ng, David T. Shima, Perry Calias, Emmett T. Cunningham, David R. Guyer, and Anthony P. Adamis (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease, *Nat Rev Drug Discov 5,* 123–132. DOI: 10.1038/nrd1955.

(70) Deigan, K. E., and FerrÉ-D'AmarÉ, A. R. (2011) Riboswitches: Discovery of Drugs That Target Bacterial Gene-Regulatory RNAs, *Accounts of Chemical Research 44,* 1329–1338. DOI: 10.1021/ar200039b.

(71) Lünse, C. E., Schüller, A., and Mayer, G. (2014) The promise of riboswitches as potential antibacterial drug targets, *International Journal of Medical Microbiology 304,* 79–92. DOI: 10.1016/j.ijmm.2013.09.002.

(72) Navani, N. K., and Li, Y. (2006) Nucleic acid aptamers and enzymes as sensors, *Current opinion in chemical biology 10,* 272–281 published online May 4, 2006. DOI: 10.1016/j.cbpa.2006.04.003.

(73) Sefah, K., Phillips, J. A., Xiong, X., Meng, L., van Simaeys, D., Chen, H., Martin, J., and Tan, W. (2009) Nucleic acid aptamers for biosensors and bio-analytical applications, *The Analyst 134,* 1765–1775 published online Jun 23, 2009. DOI: 10.1039/b905609m.

(74) Breaker, R. R. (2012) Riboswitches and the RNA World, *Cold Spring Harbor perspectives in biology 4. DOI:* 10.1101/cshperspect.a003566.

(75) 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, *Proceedings of the National Academy of Sciences of the United States of America 101,* 6421–6426. DOI: 10.1073/pnas.0308014101.

(76) Weinberg, Z., Barrick, J. E., Yao, Z., Roth, A., Kim, J. N., Gore, J., Wang, J. X., Lee, E. R., Block, K. F., Sudarsan, N., Neph, S., Tompa, M., Ruzzo, W. L., and Breaker, R. R. (2007) Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline, *Nucleic acids research 35,* 4809–4819. DOI: 10.1093/nar/gkm487.

(77) 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 research 45,* 10811–10823. DOI: 10.1093/nar/gkx699.

(78) 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 biology 11,* R31. DOI: 10.1186/gb-2010-11-3-r31.

(79) Weinberg Z., James W. Nelson, Christina E. Lünse, Madeline E. Sherlock, and Ronald R. Breaker (2017) Bioinformatic analysis of riboswitch structures uncovers variant classes with altered ligand specificity, *Proceedings of the National Academy of Sciences 114,* 201619581. DOI: 10.1073/pnas.1619581114.

(80) Meyer, M. M., Hammond, M. C., Salinas, Y., Roth, A., Sudarsan, N., and Breaker, R. R. (2011) Challenges of ligand identification for riboswitch candidates, *RNA biology 8,* 5–10.

(81) Soukup, G. A., and Breaker, R. R. (1999) Relationship between internucleotide linkage geometry and the stability of RNA, *RNA 5,* 1308–1325.

(82) Regulski, E. E., and Breaker, R. R. (2008) In-line probing analysis of riboswitches, *Methods in molecular biology (Clifton, N.J.) 419,* 53–67. DOI: 10.1007/978-1-59745-033-1_4.

(83) Westheimer, F. H. (1968) Pseudo-Rotation in the Hydrolysis of Phosphate Esters, *Accounts of Chemical Research 1,* 70–78.

(84) Ciesiollka, J., Lorenz, S., and Erdmann, V. A. (1992) Different conformational forms of Escherichia coli and rat liver 5S rRNA revealed by Pb(II)-induced hydrolysis, *European journal of biochemistry 204,* 583–589. DOI: 10.1111/j.1432-1033.1992.tb16671.x.

(85) 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, *Molecular Cell 65,* 220–230. DOI: 10.1016/j.molcel.2016.11.019.

(86) Sherlock, M. E., Malkowski, S. N., and Breaker, R. R. (2017) Biochemical Validation of a Second Guanidine Riboswitch Class in Bacteria, *Biochemistry 56,* 352–358. DOI: 10.1021/acs.biochem.6b01270.

(87) Sherlock, M. E., and Breaker, R. R. (2017) Biochemical Validation of a Third Guanidine Riboswitch Class in Bacteria, *Biochemistry 56,* 359–363. DOI: 10.1021/acs.biochem.6b01271.

(88) Greene, R. F., and Pace, C. N. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin, *The Journal of biological chemistry 249*.

(89) Jack, D. L., Storms, M. L., Tchieu, J. H., Paulsen, I. T., and Saier, M. H. (2000) A broad-specificity multidrug efflux pump requiring a pair of homologous SMR-type proteins, *Journal of bacteriology 182. DOI:* 10.1128/JB.182.8.2311-2313.2000.

(90) Schuldiner, S. (2009) EmrE, a model for studying evolution and mechanism of ion-coupled transporters, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 1794,* 748–762. DOI: 10.1016/j.bbapap.2008.12.018.

(91) Schuldiner, S. (2012) Undecided membrane proteins insert in random topologies Up, down and sideways: it doesn't really matter, *Trends in biochemical sciences 37,* 215–219. DOI: 10.1016/j.tibs.2012.02.006.

(92) Kermani, A. A., Macdonald, C. B., Gundepudi, R., and Stockbridge, R. B. (2018) Guanidinium export is the primal function of SMR family transporters, *Proceedings of the National Academy of Sciences of the United States of America 115,* 3060–3065. DOI: 10.1073/pnas.1719187115.

(93) Sherlock, M. E., and Breaker, R. R. (2020) Former orphan riboswitches reveal unexplored areas of bacterial metabolism, signaling, and gene control processes, *RNA (New York, N.Y.) 26,* 675–693. DOI: 10.1261/rna.074997.120.

(94) McCown, P. J., Liang, J. J., Weinberg, Z., and Breaker, R. R. (2014) Structural, Functional, and Taxonomic Diversity of Three PreQ1 Riboswitch Classes, *Chemistry & biology 21,* 880–889. DOI: 10.1016/j.chembiol.2014.05.015.

(95) Sudarsan, N., Lee, E. R., Weinberg, Z., Moy, R. H., Kim, J. N., Link, K. H., and Breaker, R. R. (2008) Riboswitches in eubacteria sense the second messenger cyclic di-GMP, *Science (New York, N.Y.) 321,* 411–413. DOI: 10.1126/science.1159519.

(96) Fuchs, R. T., Grundy, F. J., and Henkin, T. M. (2006) The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase, *Nature structural & molecular biology 13. DOI:* 10.1038/nsmb1059.

(97) Lenkeit, F., Eckert, I., Hartig, J. S., and Weinberg, Z. (2020) Discovery and characterization of a fourth class of guanidine riboswitches, *Nucleic Acids Res 48,* 12889–12899. DOI: 10.1093/nar/gkaa1102.

(98) 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. DOI:* 10.1021/acs.biochem.6b01269.

(99) Battaglia, R. A., and Ke, A. (2018) Guanidine-sensing riboswitches: How do they work and what do they regulate?, *Wiley interdisciplinary reviews. RNA. DOI:* 10.1002/wrna.1482.

(100) Huang, L., Wang, J., Wilson, T. J., and Lilley, D. M. J. (2017) Structure of the Guanidine III Riboswitch, *Cell chemical biology 24,* 1407-1415.e2 published online Oct 5, 2017. DOI: 10.1016/j.chembiol.2017.08.021.

(101) Reiss, C. W., and Strobel, S. A. (2017) Structural basis for ligand binding to the guanidine-II riboswitch, *RNA (New York, N.Y.) 23,* 1338–1343 published online Jun 9, 2017. DOI: 10.1261/rna.061804.117.

(102) Weinberg, Z., Regulski, E. E., Hammond, M. C., Barrick, J. E., Yao, Z., Ruzzo, W. L., and Breaker, R. R. (2008) The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches, *RNA 14,* 822–828. DOI: 10.1261/rna.988608.

(103) Poiata, E., Meyer, M. M., Ames, T. D., and Breaker, R. R. (2009) A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria, *RNA 15,* 2046–2056. DOI: 10.1261/rna.1824209.

(104) Arachchilage, G. M., Sherlock, M. E., Weinberg, Z., and Breaker, R. R. (2018) SAM-VI RNAs selectively bind S-adenosylmethionine and exhibit similarities to SAM-III riboswitches, *RNA biology 15,* 371–378. DOI: 10.1080/15476286.2017.1399232.

(105) Roth, A., Winkler, W. C., Regulski, E. E., Bobby W K Lee, Jinsoo Lim, Inbal Jona, Jeffrey E Barrick, Ankita Ritwik, Jane N Kim, Rüdiger Welz, Dirk Iwata-Reuyl, and Breaker, R. R. (2007) A riboswitch selective for the queuosine precursor preQ 1 contains an unusually small aptamer domain, *Nat Struct Mol Biol 14,* 308–317. DOI: 10.1038/nsmb1224.

(106) Meyer, M. M., Roth, A., Chervin, S. M., George A. Garcia, and Breaker, R. R. (2008) Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria, *RNA 14,* 685–695. DOI: 10.1261/rna.937308.

(107) Lee, E. R., Baker, J. L., Weinberg, Z., Sudarsan, N., and Breaker, R. R. (2010) An Allosteric Self-Splicing Ribozyme Triggered by a Bacterial Second Messenger, *Science 329,* 845–848. DOI: 10.1126/science.1190713.

(108) Sinn, M., Hauth, F., Lenkeit, F., Weinberg, Z., and Hartig, J. S. (2021) Widespread bacterial utilization of guanidine as nitrogen source, *Molecular microbiology. DOI:* 10.1111/mmi.14702.

(109) Ogston, A. (1882) Micrococcus Poisoning, *Journal of Anatomy and Physiology 17,* 24–58.

(110) Lowy, F. D. (1998) Staphylococcus aureus infections, *The New England journal of medicine 339,* 520–532. DOI: 10.1056/NEJM199808203390806.

(111) van Belkum, A., Melles, D. C., Nouwen, J., van Leeuwen, W. B., van Wamel, W., Vos, M. C., Wertheim, H. F.L., and Verbrugh, H. A. (2009) Co-evolutionary aspects of human colonisation and infection by Staphylococcus aureus, *Infection, Genetics and Evolution 9,* 32–47. DOI: 10.1016/j.meegid.2008.09.012.

(112) Kloos, W. E. (1980) Natural populations of the genus Staphylococcus, *Annual review of microbiology 34. DOI:* 10.1146/annurev.mi.34.100180.003015.

(113) Bradley, A. (2002) Bovine mastitis: an evolving disease, *Veterinary journal (London, England : 1997) 164. DOI:* 10.1053/tvjl.2002.0724.

(114) Menzies, P. I., and Ramanoon, S. Z. (2001) Mastitis of sheep and goats, *The Veterinary clinics of North America. Food animal practice 17. DOI:* 10.1016/s0749-0720(15)30032-3.

(115) Holmes, M., and Zadoks, R. (2011) Methicillin Resistant S. aureus in Human and Bovine Mastitis, *J Mammary Gland Biol Neoplasia 16,* 373–382. DOI: 10.1007/s10911-011-9237-x.

(116) Fitzgerald, J. R. (2012) Livestock-associated Staphylococcus aureus: origin, evolution and public health threat, *Trends in Microbiology 20,* 192–198. DOI: 10.1016/j.tim.2012.01.006.

(117) Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C., and Ha, T. (2008) Advances in single-molecule fluorescence methods for molecular biology, *Annual review of biochemistry 77. DOI:* 10.1146/annurev.biochem.77.070606.101543.

(118) Nienhaus, G. U. (2008) The Green Fluorescent Protein: A Key Tool to Study Chemical Processes in Living Cells, *Angewandte Chemie International Edition 47,* 8992–8994. DOI: 10.1002/anie.200804998.

(119) Finke, M. Engineering and characterization of RNA-based switches for control of gene expression in model organisms. Engineering and characterization of RNA-based switches for control of gene expression in model organisms.

(120) Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res 31,* 3406–3415. DOI: 10.1093/nar/gkg595.

(121) Hebisch, E., Knebel, J., Landsberg, J., Frey, E., and Leisner, M. (2013) High Variation of Fluorescence Protein Maturation Times in Closely Related Escherichia coli Strains, *PLOS ONE 8,* e75991. DOI: 10.1371/journal.pone.0075991.

(122) Malone, C. L., Boles, B. R., Lauderdale, K. J., Thoendel, M., Kavanaugh, J. S., and Horswill, A. R. (2009) Fluorescent Reporters for Staphylococcus aureus, *Journal of microbiological methods 77,* 251–260. DOI: 10.1016/j.mimet.2009.02.011.

(123) Blouin, S., Mulhbacher, J., Penedo, J. C., and Lafontaine, D. A. (2009) Riboswitches: Ancient and Promising Genetic Regulators, *ChemBioChem 10,* 400–416. DOI: 10.1002/cbic.200800593.

(124) Weigand, J. E., and Suess, B. (2009) Aptamers and riboswitches: perspectives in biotechnology, *Applied microbiology and biotechnology 85. DOI:* 10.1007/s00253-009-2194-2.

(125) Huang, C.-C., and Chang, H.-T. (2008) Aptamer-based fluorescence sensor for rapid detection of potassium ions in urine, *Chem. Commun.,* 1461–1463. DOI: 10.1039/B718752A.

(126) Ehrentreich-Förster, E., Orgel, D., Krause-Griep, A., Cech, B., Erdmann, V. A., Bier, F., Scheller, F. W., and Rimmele, M. (2008) Biosensor-based on-site explosives detection using aptamers as recognition elements, *Analytical and bioanalytical chemistry 391. DOI:* 10.1007/s00216-008-2150-5.

(127) de-los-Santos-Álvarez, N., Lobo-Castañón, M. J., Miranda-Ordieres, A. J., and Tuñón-Blanco, P. (2007) Modified-RNA Aptamer-Based Sensor for Competitive Impedimetric Assay of Neomycin B, *Journal of the American Chemical Society 129,* 3808–3809. DOI: 10.1021/ja0689482.

(128) Greenlee, E. B., Stav, S., Atilho, R. M., Brewer, K. I., Harris, K. A., Malkowski, S. N., Mirihana Arachchilage, G., Perkins, K. R., Sherlock, M. E., and Breaker, R. R. (2018) Challenges of ligand identification for the second wave of orphan riboswitch candidates, *RNA biology 15,* 377–390. DOI: 10.1080/15476286.2017.1403002.

(129) Triadó-Margarit, X., Vila, X., and Galinski, E. A. (2011) Osmoadaptative accumulation of Nε-acetyl-β-lysine in green sulfur bacteria and Bacillus cereus CECT 148T, *FEMS Microbiol Lett 318,* 159–167. DOI: 10.1111/j.1574-6968.2011.02254.x.

(130) Sowers, K. R., Robertson, D. E., Noll, D., Gunsalus, R. P., and Roberts, M. F. (1990) N epsilon-acetyl-beta-lysine: an osmolyte synthesized by methanogenic archaebacteria, *PNAS 87,* 9083–9087. DOI: 10.1073/pnas.87.23.9083.

(131) Galinski, E. A., and Trüper, H. G. (1994) Microbial behaviour in salt-stressed ecosystems, *FEMS Microbiol Rev 15,* 95–108. DOI: 10.1111/j.1574-6976.1994.tb00128.x.

(132) Eisenberg, H., and Wachtel, E. J. (1987) Structural studies of halophilic proteins, ribosomes, and organelles of bacteria adapted to extreme salt concentrations, *Annual review of biophysics and biophysical chemistry 16. DOI:* 10.1146/annurev.bb.16.060187.000441.

(133) Roberts, M. F. (2005) Organic compatible solutes of halotolerant and halophilic microorganisms, *Saline Systems 1,* 5. DOI: 10.1186/1746-1448-1-5.

(134) Empadinhas, N., and da, C. M. S. (2008) Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes, *International microbiology : the official journal of the Spanish Society for Microbiology 11*.

(135) Brown, A. D. (1976) Microbial water stress, *Bacteriological Reviews 40,* 803–846.

(136) Ruzicka, F. J., Lieder, K. W., and Frey, P. A. (2000) Lysine 2,3-aminomutase from Clostridium subterminale SB4: mass spectral characterization of cyanogen bromide-treated peptides and cloning, sequencing, and expression of the gene kamA in Escherichia coli, *Journal of bacteriology 182. DOI:* 10.1128/JB.182.2.469-476.2000.

(137) Csonka, L. N. (1989) Physiological and genetic responses of bacteria to osmotic stress, *Microbiological reviews 53. DOI:* 10.1128/mr.53.1.121-147.1989.

(138) McLaggan, D., Naprstek, J., Buurman, E. T., and Epstein, W. (1994) Interdependence of K+ and glutamate accumulation during osmotic adaptation of Escherichia coli, *The Journal of biological chemistry 269*.

(139) McLaggan, D., and Epstein, W. (1991) Escherichia coli accumulates the eukaryotic osmolyte taurine at high osmolarity, *FEMS microbiology letters 65. DOI:* 10.1016/0378-1097(91)90304-s.

(140) Measures, J. C. (1975) Role of amino acids in osmoregulation of non-halophilic bacteria, *Nature 257. DOI:* 10.1038/257398a0.

(141) Bermingham, A., and Derrick, J. P. (2002) The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery, *BioEssays : news and reviews in molecular, cellular and developmental biology 24,* 637–648. DOI: 10.1002/bies.10114.

(142) Suh, J. R., Herbig, A. K., and Stover, P. J. (2001) New perspectives on folate catabolism, *Annual review of nutrition 21,* 255–282. DOI: 10.1146/annurev.nutr.21.1.255.

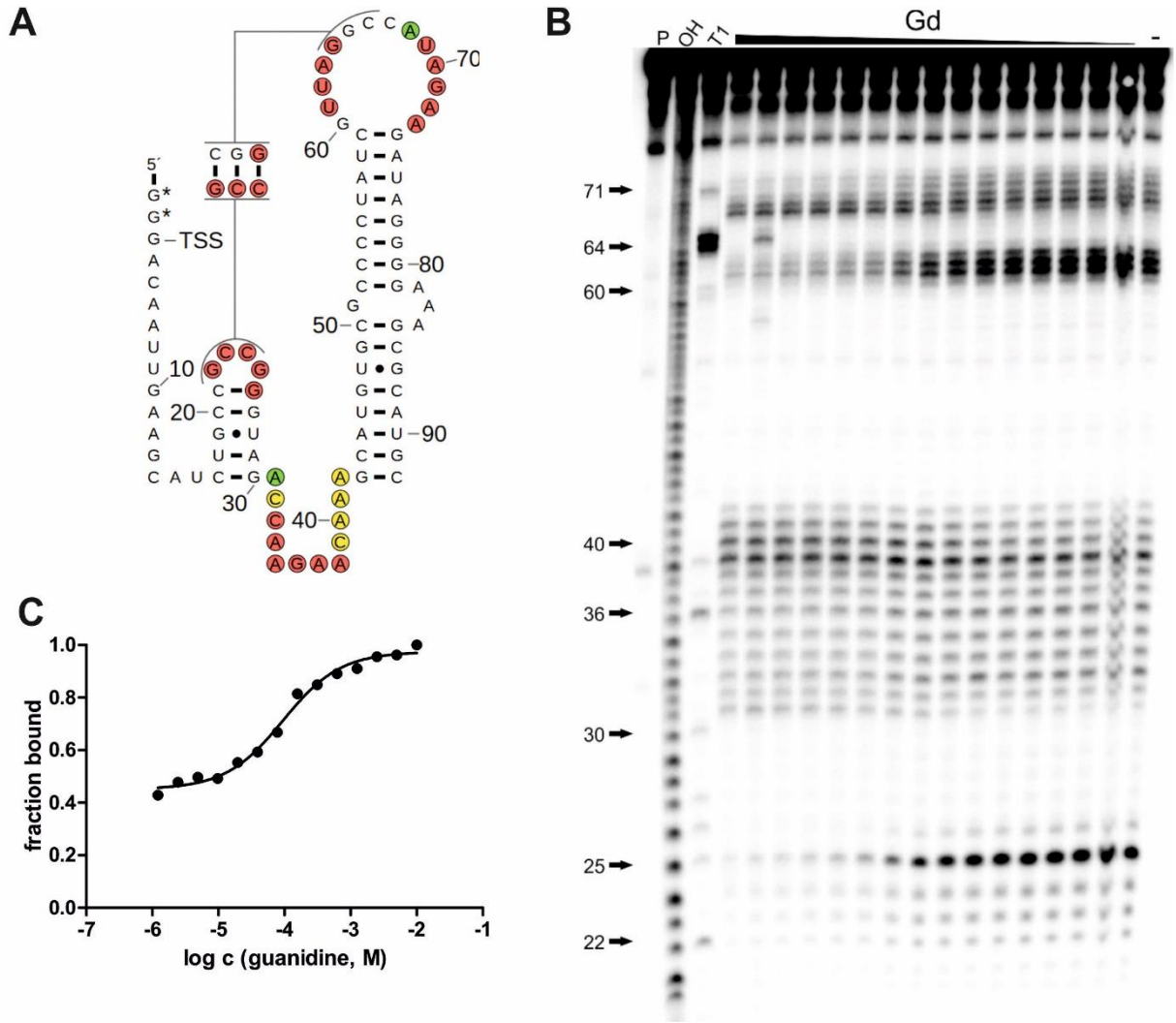
(143) Stover, P. J. (2009) One-Carbon Metabolism–Genome Interactions in Folate-Associated Pathologies12, *The Journal of Nutrition 139,* 2402–2405. DOI: 10.3945/jn.109.113670.

(144) Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M., and Hagler, A. T. (1988) Structure and energetics of ligand binding to proteins: Escherichia coli dihydrofolate reductase-trimethoprim, a drug-receptor system, *Proteins 4. DOI:* 10.1002/prot.340040106.

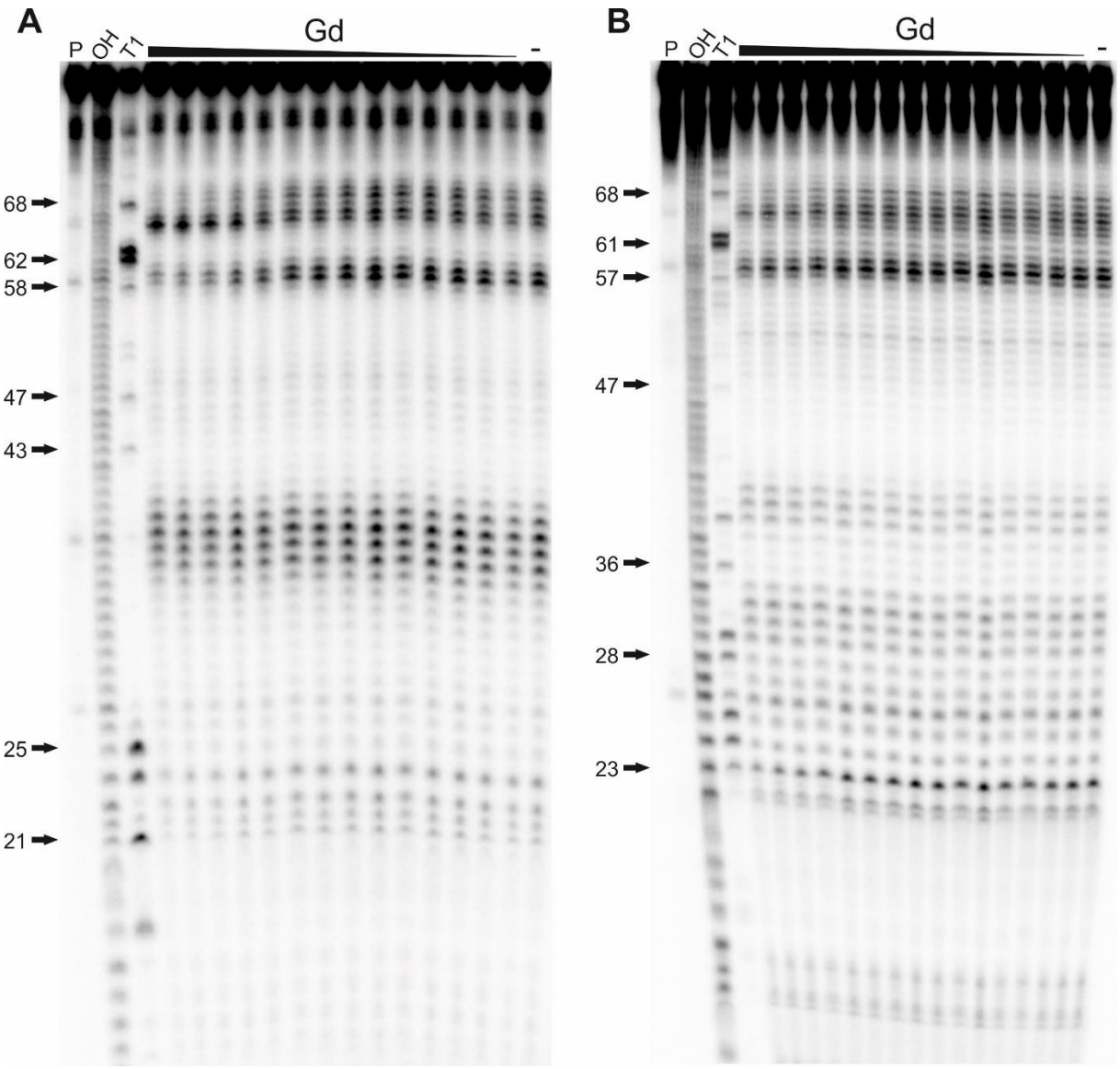
(145) Sköld, O. (2009) Sulfonamides and Trimethoprim. in *Mechanisms of drug resistance.* (Mayers, D. L., Ed.), Humana Press, New York, NY.

(146) Rajagopalan, P. T. R., Zhang, Z., McCourt, L., Dwyer, M., Benkovic, S. J., and Hammes, G. G. (2002) Interaction of dihydrofolate reductase with methotrexate: Ensemble and single-molecule kinetics, *Proceedings of the National Academy of Sciences of the United States of America 99,* 13481–13486. DOI: 10.1073/pnas.172501499.

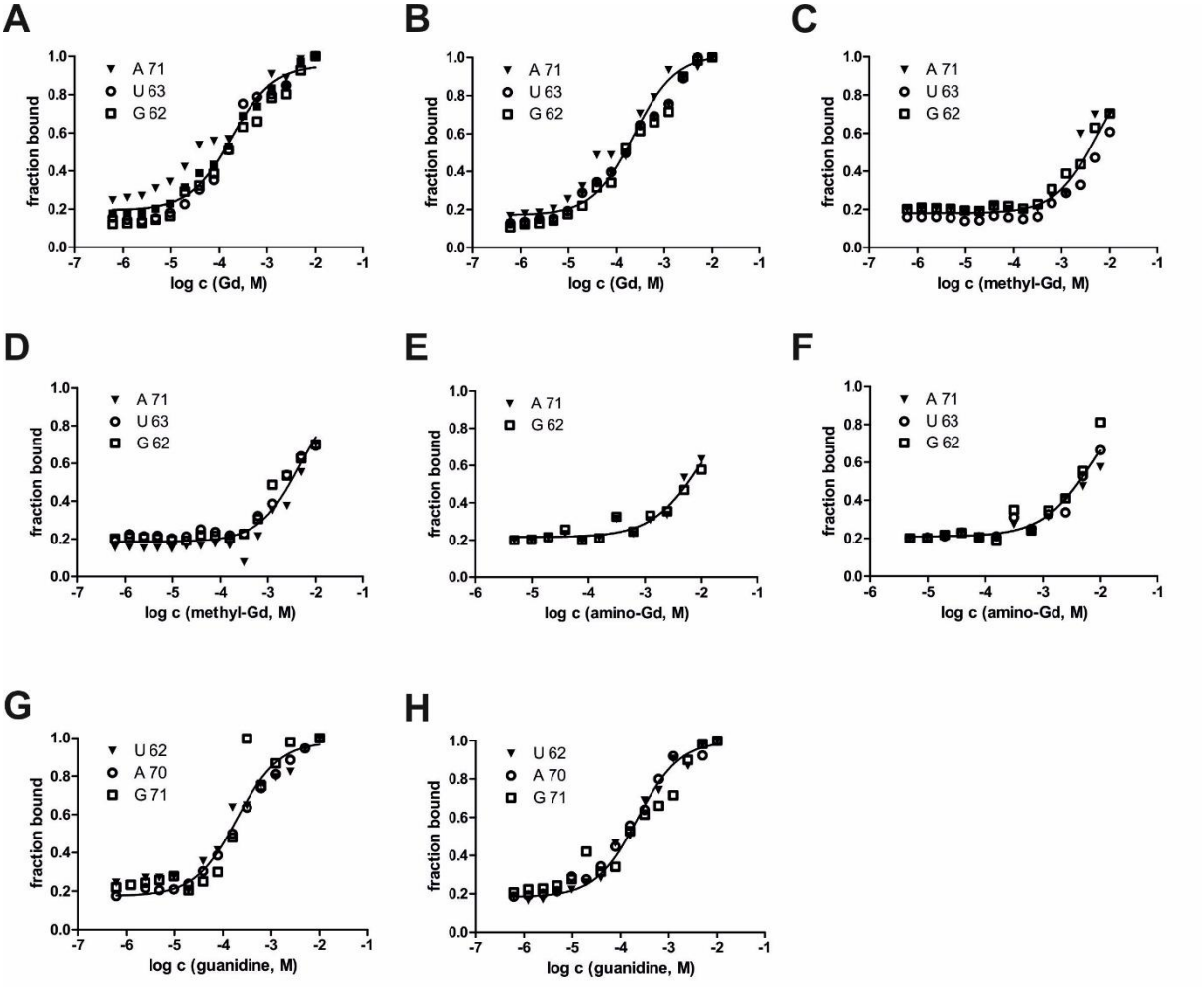
(147) David S. Goodsell (1999) The Molecular Perspective: Methotrexate, *The Oncologist 4,* 340–341. DOI: 10.1634/theoncologist.4-4-340.



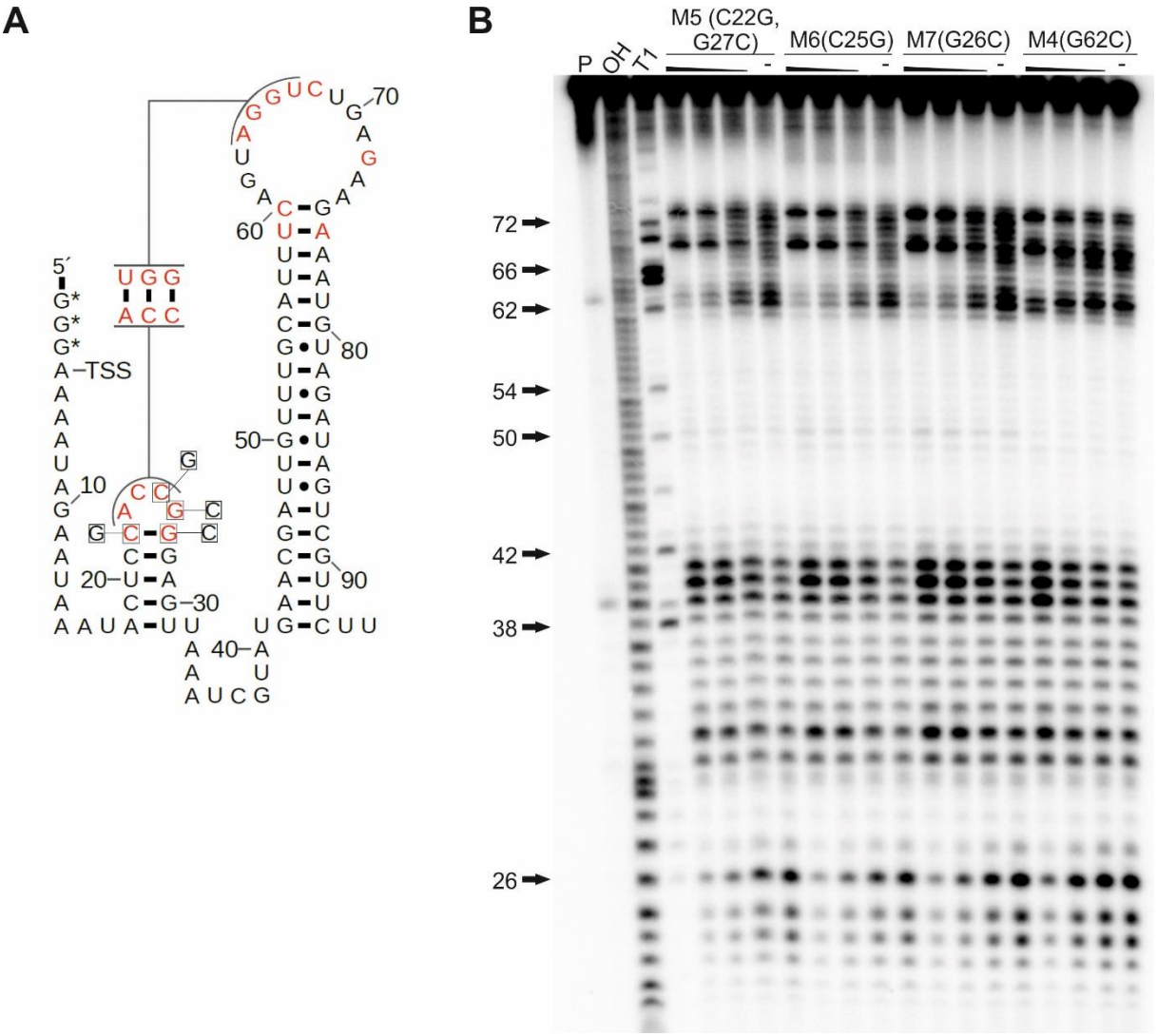
Supplementary Figure 1: Guanidine binding by *GGAM-1* motif from *Raoultibacter timonensis*. (**A**) Sequence and secondary structure of *92 Rti* RNA construct from the 5´-UTR of the *emrE* gene of *Raoultibacter timonensis*. The 5´ terminus of the construct includes two additional guanosine nucleotides to improve *in vitro* transcription efficiency. The annotations are the same as in Figure 2A. (**B**) PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *92 Rti* RNA without (-) or with guanidine hydrohydrochloride in a range of 0.61 µM – 10 mM. P, OH and T1 represent 5´-^{32}P-labeled RNA undergoing no reaction, digest with RNase T1, or digest under alkaline conditions, respectively. (**C**) Plot of the fraction of RNA bound to ligand as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined based on quantification of band intensity changes at U 62, A 70 and G 72, normalized by the intensity of the constant band A 40. A trendline was generated using a sigmoidal dose-response curve fit (maximum value equal to 1) to determine an apparent KD value. A mean apparent KD of 190 µM (standard deviation: 30 µM) was determined in three independent experiments (Supplementary Figure 3 G,H).



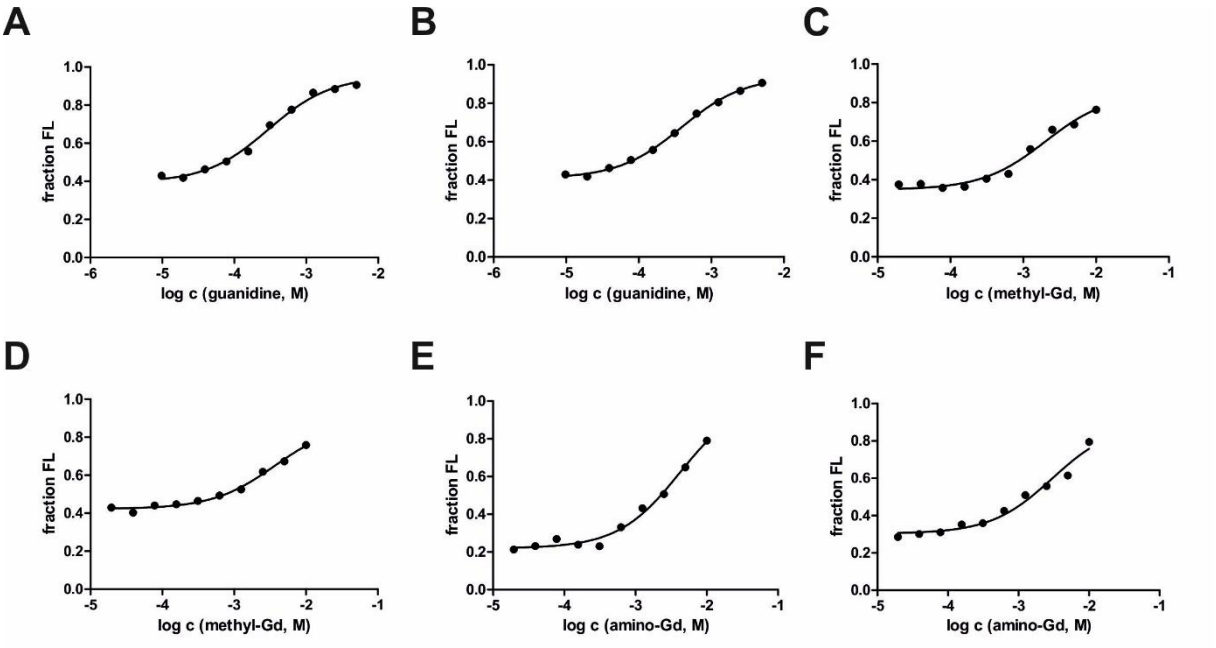
Supplementary Figure 2: Guanidine binding by *GGAM-1* motif from *Brachyspira alvinipulli and Ruminococcus albus*. (**A**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *89 Bal* RNA (sequence from *Brachyspira alvinipulli*) without (-) or with guanidine hydrohydrochloride in a range of 1.22 µM – 10 mM. (**B**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *89 Ral* RNA (sequence from *Ruminococcus albus*) without (-) or with guanidine hydrohydrochloride in a range of 0.61 µM – 10 mM. P, OH and T1 represent 5′-^{32}P-labeled RNA undergoing no reaction, digest with RNase T1, or digest under alkaline conditions, respectively.



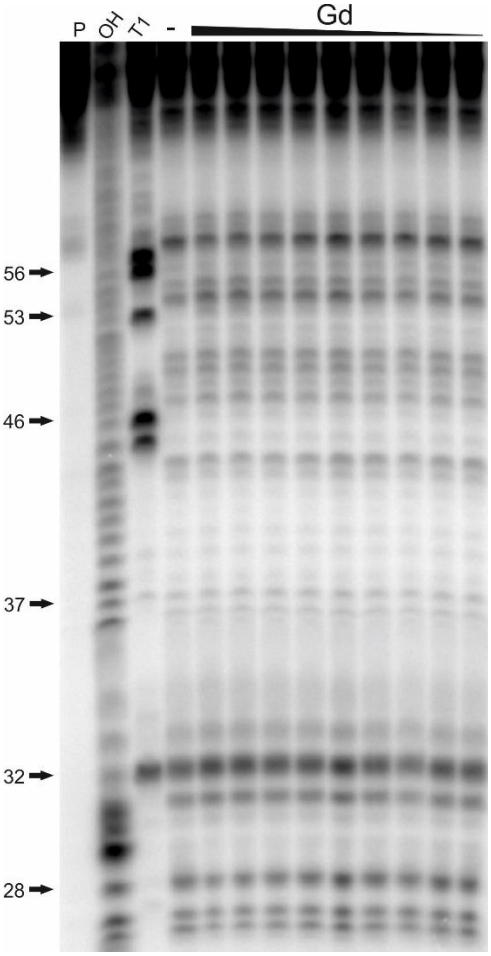
Supplementary Figure 3: Independent experiments for guanidine, methyl-guanidine and amino-guanidine binding by the *GGAM-1* motif. Data shown in this figure reflects experiments that are independent from each other and from figures in the main text. (**A**), (**B**): Plot of the fraction of *95 Lla* RNA bound to guanidine as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined as described in Figure 2. (**C**)**,** (**D**): Plot of the fraction of *95 Lla* RNA bound to methyl-guanidine as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined as described in Figure 2. (**E**), (**F**): Plot of the fraction of *95 Lla* RNA bound to amino-guanidine as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined as described in Figure 2. (**G**)**,** (**H**): Plot of the fraction of *92 Rti* RNA bound to guanidine as a function of the logarithm (base 10) of the molar guanidine hydrochloride concentration. Fraction of RNA bound was determined as described in Supplementary Figure S1.



Supplementary Figure 4: Point mutations of conserved nucleotides in P1 do not prevent guanidine binding. (**A**) Sequence and secondary structure of *95 Lla* RNA with the location of mutations in constructs M5(C22G, G27C), M6(C25G) and M7(G26C). Highly conserved positions (97% nucleotide identity) are shown in red. (**B**) PAGE analysis of an in-line probing reaction of 5′-^{32}P-labeled *95 Lla* wt RNA and mutants M5(C22G, G27C), M6(C25G), M7(G26C) and M4(G62C) without (−) or with guanidine hydrochloride with concentrations of 10 mM, 1 mM and 100 µM. P, OH and T1 represent 5′-^{32}P-labeled RNA undergoing no reaction, digest under alkaline conditions, or digest with RNase T1, respectively.



Supplementary Figure 5: Transcription termination control by guanidine and guanidine derivatives, independent experiments. (**A**), (**B**): Plot of the fraction of full length *147 Lla* product relative to the total number of transcripts as a function of the guanidine hydrochloride concentration. (**C**), (**D**): Plot of the fraction of full length *147 Lla* product relative to the total number of transcripts as a function of the methylguanidine hydrochloride concentration. (**E**), (**F**): Plot of the fraction of full length *147 Lla* product relative to the total number of transcripts as a function of the amino-guanidine hydrochloride concentration.



Supplementary Figure 6: shorter construct of *Gd4v* motif RNA binds guanidine with lower affinity. PAGE analysis of an in-line probing reaction of 5´-^{32}P-labeled *81 Csp* RNA (sequence from *Cloabacillus sp.*) without (-) or with guanidine hydrohydrochloride in a range of 39 µM – 10 mM.

Supplementary Table 1: Sequence of oligonucleotides used for the analysis of the *GGAM-1* motif. Sequences were taken from *Lactococcus lactis (Lla)*, *Raoultibacter timonensis (Rti)*, *Ruminococcus albus (Ral)* and *Brachyspira alvinipulli (Bal)*, respectively. T7 RNA Polymerase Promotor is shown in green, T5 RNA Polymerase Promotor in blue. Lowercase letters identify guanosine nucleotides, added to enhance transcription. Nucleotides that were mutated relative to the wt sequence are shaded yellow, start codon is shaded grey.

Construct name	Sequence (5´-3´)
95 Lla	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACCGGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGGTCTGAGAAGAAATGTAGATAGTCGTTCTT
95 Lla M1 (G66C)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACCGGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGCTCTGAGAAGAAATGTAGATAGTCGTTCTT
95 Lla M2 (U67A)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACCGGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGGACTGAGAAGAAATGTAGATAGTCGTTCTT
95 Lla M3 (G72C)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACCGGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGGTCTGACAAGAAATGTAGATAGTCGTTCTT
95 Lla M4 (G62C)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACCGGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCACTAGGTCTGAGAAGAAATGTAGATAGTCGTTCTT
95 Lla M5 (C22G G27C)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCGACCGCGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGGTCTGAGAAGAAATGTAGATAGTCGTTCTT
95 Lla M6 (C25G)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACGGGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGGTCTGAGAAGAAATGTAGATAGTCGTTCTT
95 Lla M7 (G26C)	TAATACGACTCACTATAgggAAAATAGAATAAATACTCCACCCGGAGTTAAATCGT ATGAACGATTGTTTGCATTTCAGTAGGTCTGAGAAGAAATGTAGATAGTCGTTCTT
147 Lla	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGTCTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAGGAGGCGAAAATGACTTGG CTATATCTACTAATAGCAGGAATTT
92 Rti	TAATACGACTCACTATAggGACAATTGAAGCATCTGCCGCCGGGTAGACCAAGAA CAAAGCATGTGCGCCCCTATCGTTAGGCCATAGAAGATAGGGGAAAGCGCATGC
89 Ral	TAATACGACTCACTATAggGCAATATGAAAACTGCCACCGGGTAGGAATGCTGAA AGCATTTGTTCACACATTGTTAGGTCTTAGAAGATGTGTGCGCAGATGCTT
89 Bal	TAATACGACTCACTATAggGAAATGATAAAATTTGTCGCCGGATAAATTTTTTTAAA ACGTTTGTATTCATACCGTAGGCCTTAGAAAGTATGAACGCAAACGTTT

Supplementary Table 2: Sequence of oligonucleotides used for the analysis of the *GGAM-2* to *6* motifs. Sequences were taken from *Cellulomonas fimi*, *Burkholderia ubonensis*, *Methyloceanibacter superfactus*, *Novosphingobium sp.*, *Selenomonas ruminantium*, respectively. T7 RNA Polymerase Promotor is shown in green. Lowercase letters identify guanosine nucleotides, added to enhance transcription.

Construct name	Sequence (5´-3´)
63 Cfi (GGAM-2)	TAATACGACTCACTATAGGAATGGTACGGTCGTACCAGAACCGTGCGCGGGACG GCGGACCTGACACCCGGCGCGCAGCG
80 Cfi (GGAM-2)	TAATACGACTCACTATAgGTCGCAGGAATGGTACGGTCGTACCAGAACCGTGCGC GGGACGGCGGACCTGACACCCGGCGCGCAGCGGTCGCCCACC
73 Bub (GGAM-3)	TAATACGACTCACTATAggGAUCCAUCCGGGGUACCGCUCCCGCAAGUUGAUCG CGCUCGGUACAAGCGCGAAAUCAGGAGCGGUACGCC
59 Bub (GGAM-3)	TAATACGACTCACTATAggGAUCCAUCCGGGGUACCGCUCCCGCAAGUUGAUCG CGCUCGGUACAAGCGCGAAAUC
80 Msu (GGAM-4)	TAATACGACTCACTATAgggAAACTAAGGTCCGCCCGCTAGGCCGGCCGCGCCTCGA CGCGCCGGCTCTCGTCAATCGAACACGCGCGGTACCCATG
98 Msu (GGAM-4)	TAATACGACTCACTATAgggCGACAGCCTCTCCTCGCGAAACTAAGGTCCGCCCGCTA GGCCGGCCGCGCCTCGACGCGCCGGCTCTCGTCAATCGAACACGCGCGGTACCCAT G
77 Nov (GGAM-5)	TAATACGACTCACTATAggGTGGATGCGCCGAGAATAGAGTAGCAGGGTGGTGTCTC CTGCGCCCGTAAGGGTTCTTCCCAGCCACGGCGTCAA
111 Sru (GGAM-6)	TAATACGACTCACTATAgggTCAATACTCGTATCGGACGGCAAGCAATCGCCTGCTTG CCAGATAAGGTGTCCATAGACACTGGAACGGGTTATTCAGGCCAGATAAGATCGCAT TCTTATCTGGCCT

Supplementary Table 3: Sequences of *hly* 5´-UTR and constructs embedded into the *hly* 5´-UTR for conditional *egfp* expression in *S. aureus*. *GGAM-1* motif sequences from *L. lactis* are underlined. Shine Dalgarno sequence is shown in orange. Nucleotides that were mutated relative to the wt sequence are shaded yellow, start codon is shaded grey.

Construct name	Sequence (5´-3´)
hly 5´-UTR	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATTAGGTTAAAAA ATGTAGAAGGAGAGTGAAACCC
Gd4 (33 (1))	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGTCTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAAGGAGAGTGAAACCCATG
Gd4 (33 (1)) M1	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGCTCTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAAGGAGAGTGAAACCCATG

Gd4 (33 (1)) M2	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGACTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAAGGAGAGTGAAACCCATG
Gd4 (33 (1)) M3	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGTCTG ACAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAAGGAGAGTGAAACCCATG
Gd4 (33 (1)) M4	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCACTAGGTCTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAAGGAGAGTGAAACCCATG
Gd4 (33 (2))	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGTCTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTTAGGTTAAAAAATGTAGAAGGAGAG TGAAACCCATG
Gd4 (33 (3))	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATAAAATAGAATA AATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGTCTG AGAAGAAATGTAGATAGTCGTTCTTTTTTTAGCTGAGGAGGCGAAAATG
Gd4 (47)	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGAAAATAGAATAAATACTCCACCGG GAGTTAAATCGTATGAACGATTGTTTGCATTTCAGTAGGTCTGAGAAGAAATGTA GATAGTCGTTCTTTTTTTGTATTTGGCATTATTAGGTTAAAAAATGTAGAAGGAGA GTGAAACCCATG
Gd4 (18)	ATAAAGCAAGCATATAATATTGCGTTTCATCTTTAGAAGCGAATTTCGCCAATATT ATAATTATCAAAAGAGAGGGGTGGCAAACGGTATTTGGCATTATTAGGTTAAAAA ATGTAAAATAGAATAAATACTCCACCGGGAGTTAAATCGTATGAACGATTGTTTGC ATTTCAGTAGGTCTGAGAAGAAATGTAGATAGTCGTTCTTTTTTTAGAAGGAGAGT GAAACCCATG

Supplementary Table 4: Sequence of oligonucleotides used for the analysis of the *Gd4 variant* motif. Sequences were taken from *Cloacibacillus sp*., *Dialister succinatiphilus* and *Mitsuokella jalaludinii*. T7 RNA Polymerase Promotor is shown in green, T5 RNA Polymerase Promotor in blue. Lowercase letters identify guanosine nucleotides, added to enhance transcription. Nucleotides that were mutated relative to the wt sequence are shaded yellow, start codon is shaded grey.

Construct name	Sequence (5´-3´)
95 Csp	TAATACGACTCACTATAggAATGCTTCCTCCGCCACCGGGCGGAGGGTGAAAAGCGT CCACGGTCTCCCGTAGGTACGGCGCGCTCTGCGCGTTACAAGGGAGCCGCGGGCG
95 Csp M1 (G75C)	TAATACGACTCACTATAggAATGCTTCCTCCGCCACCGGGCGGAGGGTGAAAAGCGT CCACGGTCTCCCGTAGCTACGGCGCGCTCTGCGCGTTACAAGGGAGCCGCGGGCG
95 Csp M2 (U58A)	TAATACGACTCACTATAggAATGCTTCCTCCGCCACCGGGCGGAGGGTGAAAAGCGT CCACGGTCTCCCGTAGGAACGGCGCGCTCTGCGCGTTACAAGGGAGCCGCGGGCG
81 Csp	TAATACGACTCACTATAggAATGCTTCCTCCGCCACCGGGCGGAGGGTGAAAAGCGT CCACGGTCTCCCGTAGGTACGGCGCGCTCTGCGCGTTACAA
95 Dsu	TAATACGACTCACTATAggAAACTCTTCCCTGTCACCGGATAGGGAGCCAAGAGGGT TTATTCCTGCTGCAGGTAATATGGCGAAAACCATTGAAAGGCAGGAATAAATCCT
170 Csp	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAAATGCTTCCTCC GCCACCGGGCGGAGGGTGAAAAGCGTCCACGGTCTCCCGTAGGTACGGCGCGCT CTGCGCGTTACAAGGGAGCCGCGGGCGCTTTATTTTTTATAAGGAGGGTTCTGCA TGTATATCAAGCCGTTTGAAGTGGAGGAATGGATGAACGCATGGGAAAC
162 Dsu	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAAAACTCTTCCCT GTCACCGGATAGGGAGCCAAGAGGGTTTATTCCTGCTGCAGGTAATATGGCGAA AACCATTGAAAGGCAGGAATAAATCCTCTTTTTTTTAAGGAGGTTATATCATGAAC ATTAAACCCTTTGCTGTCGAAGAATGGATGAATGCCTATG
105 Mja	TAATACGACTCACTATAggAACACCGCTTCTGTCACCGGACAGGAGAAATGGGGT GTCTGGGAGTTGTCGTAGGTACGGTATGGTTGCATGCTATAAGACAATTCCTGGG CACCCCTGTTTT

Supplementary Table 5: Sequence of oligonucleotides used for the analysis of the *ablB* motif. Sequences were taken from *Acetobacterium woodii*, *Psychrilyobacter atlanticus* and *Dethiobecter alkaliphilus*. T7 RNA Polymerase Promotor is shown in green, T5 RNA Polymerase Promotor in blue. Lowercase letters identify guanosine nucleotides, added to enhance transcription. Nucleotides that were mutated relative to the wt sequence are shaded yellow, start codon is shaded grey.

Construct name	Sequence (5´-3´)
130 Awo	TAATACGACTCACTATAggTAATAGAACTGGGGCTTGTATAGGATCATAATGGGTC ATATACTATACTATTTGGTAAGGTGGAATTCTTATCGTTAATGGTGTAATGTGAGT TATTTTTAGCGAAACTTAGGCACTCTAACTCTGTG
109 Pat	TAATACGACTCACTATAggTTATAGAACTGGGGTTTATTCAGGGTCTCTTTTGAGGC AATGTATTATGCTACTTATGTGGCATATTGTGAGTCTATTAGACGAAACTTAAACA CTCTAACTCTATT
92 Pat	TAATACGACTCACTATAggCTGGGGTTTATTCAGGGTCTCTTTTGAGGCAATGTATT ATGCTACTTATGTGGCATATTGTGAGTCTATTAGACGAAACTTAAACACTCTA
98 Dal	TAATACGACTCACTATAggAAAGGGGTTCCTGGCTGGTTCCCTGTTGGGCCAGCAG TAGGCTACATCTGAAGATGTGGTTTAAGGTGAGTCGAAAGGCGAAACCTGGGCA CTCTA
111 Dal	TAATACGACTCACTATAggGGAGGAAAGGGGTTCCTGGCTGGTTCCCTGTTGGGC CAGCAGTAGGCTACATCTGAAGATGTGGTTTAAGGTGAGTCGAAAGGCGAAACC TGGGCACTCTAGTCTCTTC
331 Awo	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAAAATATTTATAG AACTGGGGTTTATTCAGGGTCTCTTTTGAGGCAATGTATTATGCTACTTATGTGGC ATATTGTGAGTCTATTAGACGAAACTTAAACACTCTAACTCTATTACCTTTATGGTA GTAGAATTAGAGTGTTTTTTTGTTATTAGGAGGAAAAAATGGAGAAAAAAAAGAA AACAGTACAAAAATCAAAAACAAGGTTAAAAAAACCAAAGTTTGAAACTGATAAA C
236 Pat	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAAAACTCTTCCCT GTCACCGGATAGGGAGCCAAGAGGGTTTATTCCTGCTGCAGGTAATATGGCGAA AACCATTGAAAGGCAGGAATAAATCCTCTTTTTTTTAAGGAGGTTATATCATGAAC ATTAAACCCTTTGCTGTCGAAGAATGGATGAATGCCTATG
130 Awo M1 (A93U)	TAATACGACTCACTATAggTAATAGAACTGGGGCTTGTATAGGATCATAATGGGTC ATATACTATACTATTTGGTAAGGTGGAATTCTTATCGTTAATGGTGTAATGTGTGT TATTTTTAGCGAAACTTAGGCACTCTAACTCTGTG
130 Awo M2 (G92C, A93U)	TAATACGACTCACTATAggTAATAGAACTGGGGCTTGTATAGGATCATAATGGGTC ATATACTATACTATTTGGTAAGGTGGAATTCTTATCGTTAATGGTGTAATGTCTGTT ATTTTTAGCGAAACTTAGGCACTCTAACTCTGTG
130 Awo M3 (A107U, A108U)	TAATACGACTCACTATAggTAATAGAACTGGGGCTTGTATAGGATCATAATGGGTC ATATACTATACTATTTGGTAAGGTGGAATTCTTATCGTTAATGGTGTAATGTGAGT TATTTTTAGCGTTACTTAGGCACTCTAACTCTGTG
109 Pat M1 (A73U)	TAATACGACTCACTATAggTTATAGAACTGGGGTTTATTCAGGGTCTCTTTTGAGGC AATGTATTATGCTACTTATGTGGCATATTGTGTGTCTATTAGACGAAACTTAAACA CTCTAACTCTATT
109 Pat M2 (G72C, A73U)	TAATACGACTCACTATAggTTATAGAACTGGGGTTTATTCAGGGTCTCTTTTGAGGC AATGTATTATGCTACTTATGTGGCATATTGTCTGTCTATTAGACGAAACTTAAACAC TCTAACTCTATT

Supplementary Table 6: Sequence of oligonucleotides used for the analysis of the *folP* motif. T7 RNA Polymerase Promotor is shown in green, T5 RNA Polymerase Promotor in blue. Lowercase letters identify guanosine nucleotides, added to enhance transcription. The start codon is shaded grey.

Construct name	Sequence (5´-3´)
105 Dsu	TAATACGACTCACTATAgggTTACATGGAGGAAATGCCATGGGACGGGTTTATCTG TCCGGTCTCATTGATGATTATTTCCCTGATTATCAATCTGACTGCGGATTCGTCATA CTCCTCAAT
86 Dsu	TAATACGACTCACTATAgggTACATGGAGGAAATGCCATGGGACGGGTTTATCTGT CCGGTCTCATTGATGATTATTTCCCTGATTATCAATCTGACTGCGGA
106 Din	TAATACGACTCACTATAgggTACATGGAGGAAATGCCATGGCGGTGGATGTACGTC CATCGAGCCGGTTTGATGATTATTTCCCTGATTATCAATTCGGCTGCGGATGATCA TACTCCTCAAT
89 Din	TAATACGACTCACTATAgggTACATGGAGGAAATGCCATGGCGGTGGATGTACGTC CATCGAGCCGGTTTGATGATTATTTCCCTGATTATCAATTCGGCTGCGGA
105 Vba	TAATACGACTCACTATAgggCACATGGAGGAAATGCCATGTCTATGAATAATTTCAT AGGGTCGAATTGATGATTATTTCCCTGATTATCAATTTGGTTGCAGGTAGTTCATA CTCCTCAAT
87 Vba	TAATACGACTCACTATAgggCACATGGAGGAAATGCCATGTCTATGAATAATTTCAT AGGGTCGAATTGATGATTATTTCCCTGATTATCAATTTGGTTGCAGG
369 Din	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATATACATGGAGGA AATGCCATGGCGGTGGATGTACGTCCATCGAGCCGGTTTGATGATTATTTCCCTGA TTATCAATTCGGCTGCGGATGATCATACTCCTCAATGAGCATCTTGCGCAAAGGAG GCGGTGTTACACCGCTTTTTTTGTTGTCCGATTTCATTGCTACTGTCTTCTGGTGGT GATAAAATAACAATATTACAGAGTATGGAGTTATGATTTCTGCAAGGAGGAAAAT TGGAACGGGCTTATACGTGGTGTGACGGAAAGAGATTGGAAGTTACAGCAAAAA CGCTGATTATGGGGATATTGAATGTAACCCCCGATTCCTTTTCTGATGGCGGAAG GTGGAATACGGCGGAAAAGGCACAA

10. List of Abbreviations

DNA	deoxyribonucleic acid
RNA	Ribonucleic acid
mRNA	messenger RNA
(d)AT(M/D)P	deoxy)adenosine tri-(mono-/di-)phosphate
TPP	thiamine triphosphate
FMN	Flavin mononucleotide
SAM	S-adenosyl methionine
THF	tetrahydrofolate
AdoCbl	Adenosylcobalamine
Moco	Molybdenum cofactor
ZTP	5-amino-4-imidazole carboxamide riboside 5'-triphosphate
ppGpp	Guanosine-3',5'-bispyrophosphate
preQ$_1$	Pre-queuosine$_1$
TEC	transcription elongation complex
SD	Shine-Dalgarno
GlcN6P	Glucoseamine-6-phosphate
mRNA	messenger RNA
GGAM	Guanidine-Gene-Associated Motifs
L. lactis	*Lactococcus lactis*
R. timonensis	*Raoultibacter timonensis*
UTR	untranslated region
Gd	Guanidine
eGFP	enhanced green fluorescent protein
OD	optical density
S. aureus	*Staphylococcus aureus*
B. subtilis	*Bacillus subtilis*
MRSA	methicillin-resistant *S. aureus*
E. coli	*Escherichia coli*
P. canavaninivorans	*Pseudomonas canavaninivorans*
EC50	half maximum effective concentration
ncRNA	noncoding RNA
TMAO	Trimethylaminoxid
AAT	aspartate aminotransferase

PLP	pyridoxal phosphate
10f-THF	10-formyl-tetrahydrofolate
ONPG	ortho-Nitrophenyl-β-galactoside
°C	degree Celsius
A	adenosine, deoxyadenosine
C	cytidine, deoxycytidine
G	guanosine, deoxyguanosine
U	uridine
T	deoxythymidine
a.u.	arbitrary units
BHI	Brain Heart Infusion
LB	lysogeny broth
min	minute
mL	milliliter
M	molar
mM	millimolar
ng	nanogram
µg	microgram
µL	microliter
PCR	Polymerase Chain Reaction

11. Acknowledgements

Zuerst möchte ich mich ganz besonders bei Prof. Dr. Jörg Hartig bedanken, da er mir die Möglichkeit gegeben hat die letzten vier Jahre in seiner Arbeitsgruppe zu arbeiten und zu forschen. Ich bedanke mich für wissenschaftlichen Austausch, für die Freiheit eigenen Ideen nachzugehen und für die vielseitige Unterstützung, sowohl wenn es gut lief, aber auch wenn es mal nicht so recht klappen wollte.

Darüber hinaus möchte ich mich bedanken bei Prof. Dr. David Schleheck für die Übernahme des Zweitgutachtens sowie für das Mitwirken in meinem Promotions-Komitee. Bei Prof. Dr. Valentin Wittmann bedanke ich mich für die Übernahme des Prüfungsvorsitzes.

Bei Prof. Dr. Ronald Breaker bedanke ich mich für die Unterstützung als Mentor innerhalb des MEiN Mentoring Programmes, sowie für die Möglichkeit zwei Monate in seinem Labor zu arbeiten und viele Geheimnisse des Breaker Labs zu erlernen. Ich möchte mich auch bedanken bei der gesamten Breaker Arbeitsgruppe. Danke dafür, dass ich so nett aufgenommen wurde und für bis heute andauernden wissenschaftlichen Austausch.

Ein großer Dank geht auch an Dr. Zasha Weinberg. Die gemeinsame Kooperation war mir immer eine große Freude und ich bedanke mich für viele Ideen, wissenschaftliche Diskussionen und eine wirklich gute Zusammenarbeit.

Ich möchte mich bei der ganzen Arbeitsgruppe bedanken ohne die die letzten Jahre wesentlich weniger spannend und witzig gewesen wären. Hier findet man immer bei jedem ein offenes Ohr für wissenschaftlichen Austausch, für das Diskutieren neuer Ideen und Methoden und auch für Privates. Ich schätze unseren Zusammenhalt und Teamwork, die angenehme Arbeitsatmosphäre und unsere Kaffeepausen sehr. Außerdem geht ein großer Dank an Astrid, die für jedes Klonierungs-Dilemma eine Lösung parat hat und ohne die wir alle im Chaos versinken würden.

Ein riesiges Dankeschön geht an meine Freunde und an meine Familie. Insbesondere danke ich meiner Mutter, meinem Vater, Hannah und Romy dafür, dass sie bis zum Ende immer versucht haben zu verstehen was eigentlich ein Riboswitch ist. Ein riesiger Dank gilt Doreen, die mich im Bachelor, Master und in der Promotion begleitet hat. Danke für den sportlichen Ausgleich, für Kaffee und jede Menge positive Energie. Florian danke ich für all die emotionale Unterstützung, Rückendeckung, Abenteuer, für Stockbrot und für seine Zuversicht.