Published in final edited form as:

Biochim Biophys Acta. 2009 ; 1789(0): 592–611. doi:10.1016/j.bbagr.2009.07.002.

Amino acid recognition and gene regulation by riboswitches

Alexander Serganov and Dinshaw J. Patel

Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY, 10021, USA

Abstract

Riboswitches specifically control expression of genes predominantly involved in biosynthesis, catabolism and transport of various cellular metabolites in organisms from all three kingdoms of life. Amongst many classes of identified riboswitches, two riboswitches respond to amino acids lysine and glycine to date. Though these riboswitches recognize small compounds, they both belong to the largest riboswitches and have unique structural and functional characteristics. In this review, we attempt to characterize molecular recognition principles employed by amino acid-responsive riboswitches to selectively bind their cognate ligands and to effectively perform a gene regulation function. We summarize up-to-date biochemical and genetic data available for the lysine and glycine riboswitches and correlate these results with recent high-resolution structural information obtained for the lysine riboswitch. We also discuss the contribution of lysine riboswitches to antibiotic resistance and outline potential applications of riboswitches in biotechnology and medicine.

Keywords

lysine riboswitch; glycine riboswitch; RNA structure; X-ray crystallography

Introduction

The evolution of primordial macromolecules has resulted in the domination of protein molecules over nucleic acids in modern organisms. The 20 amino acids that prevail in proteins display a much more diverse set of functional groups as compared to the 4 common nucleotides found in DNA or RNA. Consequently, proteins are utilized for much broader structural and functional applications in virtually all cellular processes. Given their dependence on protein biosynthesis, cells must ensure the constant availability of various amino acids within certain concentration ranges. Many bacteria have the dual ability to either use amino acids necessary for growth from environmental sources or to make these molecules from simpler compounds. To save energy and resources, the biosynthesis of available amino acids needs to be shut down under certain circumstances and the expression of genes encoding biosynthetic enzymes needs also to be switched off. The regulation of gene expression is achieved by tight feedback control at the transcriptional and post-transcriptional levels. Owing to differences in cellular metabolism, the mechanisms of regulation of genes involved in amino acid metabolism and transport vary in evolutionarily

© 2009 Elsevier B.V. All rights reserved.

corresponding author: serganoa@mskcc.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

distant species. Nevertheless, gene expression modulation is typically based on two sophisticated approaches that involve either sensing tRNA molecules or measuring the translation rate for the leader peptide. The former mechanism depends on the selective binding of tRNA by a region of mRNA termed T-box [1]. If the cognate tRNA is not amino acylated, it specifically recognizes the T-box RNA using its anticodon loop for Watson-Crick base pairing with a region called the 'specifier'. This interaction stabilizes the T-box conformation thereby increasing the production of proteins responsible for maintaining an adequate level of amino acylated tRNA. If tRNA is amino acylated, the T-box RNA rejects it and adopts a conformation that inhibits the transcription or translation of the genes under control. The latter mechanism of amino acid homeostasis relies on classical attenuators that change their conformation in response to ribosome stalling [2]. The gene under control typically contains several codons for the 'effector' amino acid in the leader region. If the concentration of the effector amino acid is high, the ribosome translates the leader peptide quickly, thereby allowing for the formation of the transcription terminator in the downstream region of mRNA. When the effector amino acid is in short supply, the ribosome pauses at the corresponding codons and promotes the formation of the alternative anti-terminator mRNA conformation. Though the two systems that control the availability of amino acids are broadly distributed, T-boxes are usually found in Gram-positive bacteria while attenuators are mostly present in Gram-negative bacteria [3, 4]. In addition, RNA binding proteins, such as the well-known TRAP [5], and DNA binding transcriptional factors, such as MtaR [6], contribute to gene expression regulation in a number of bacterial species.

Given the variety of mechanisms involved in the modulation of genes needed for the maintenance of sufficient levels of amino acids [7], the finding of yet another method of regulation, namely riboswitch-based control, was unanticipated and garnered considerable excitement in the field. Most similar to T-boxes and attenuators, riboswitches [8–10], the metabolite-receptive mRNA segments, employ a mechanism of control that depends on the interplay of two alternative conformations in the leading region of mRNA. Nevertheless, the riboswitch mechanism significantly diverges from the strategies implemented by T-boxes and attenuators for the gene expression modulation. In contrast to T-boxes, riboswitches can directly sense their cognate metabolites and, unlike attenuators, the formation of the alternative riboswitch conformations depends on the metabolite binding and not on the kinetics of the leader peptide translation.

Riboswitches are usually found in the 5'-untranslated regions (5'-UTRs) of bacterial mRNAs whose translation products participate in the metabolism and transport of metabolites that control riboswitches. Typical riboswitches consist of two domains: an evolutionarily conserved sensor or aptamer domain that specifically binds the cognate metabolite and a highly variable expression platform that contains gene expression signals (Fig. 1). The metabolite-driven regulation of gene expression is accomplished through the interplay of two alternative riboswitch conformations, which permit gene expression to be turned 'on' or 'off' [11]. Upon metabolite binding, the sensing domain adopts a stable metabolite-bound conformation that aids the formation of a gene expression controlling element, such as a transcription terminator or a hairpin sequestering the ribosome binding site. If the intracellular concentration of metabolite is below the threshold level, the riboswitch forms an alternative structure that engages the complementary regions of both riboswitch domains in base pairing, thereby precluding the formation of gene expression controlling elements. Most amino acid-specific riboswitches use the interplay between alternative RNA conformations to modulate premature transcription termination or translation initiation [12]. As described for some other riboswitches [13, 14], the expression signals of amino acid-dependent riboswitches may overlap with the metabolite-sensing domains, thus simplifying the fold of the regulatory region.

Amongst over a dozen major types of riboswitches identified to date, five riboswitches respond to cofactors used by protein enzymes [8–10, 13, 15–20], one riboswitch binds to a byproduct of cofactor metabolism [21], five riboswitches sense purines and their derivatives [22–26], and two riboswitches interact with amino acids [27–29]. In addition, a mRNA element was found in the 5′-untranslated region (UTR) of the *glmS* gene that specifically binds a phosphorylated amino sugar glucosamine-6-phosphate (GlcN6P) and uses this compound for mRNA cleavage [30]. Recent studies have also identified two classes of mRNA that sense magnesium cations [31, 32]. Although only two riboswitches can directly sense amino acids, the control of amino acid concentrations in bacteria is accomplished by three riboswitches: lysine riboswitches modulate the metabolism of lysine and precursors common to threonine and methionine by binding to lysine [27, 29], glycine riboswitches regulate glycine cleavage via interactions with glycine [28], and *S*-adenosylmethionine (SAM)-responsive riboswitches participate in the regulation of methionine and cysteine genes by sensing the coenzyme SAM. It was also suggested that in *Aspergillus nidulans*, the splicing of the arginase gene *agaA* is modulated by the interaction of arginine with the 5′-UTR [33]. This review will focus on well-studied lysine- and glycine-responsive riboswitches but will not discuss SAM-specific riboswitches, which specifically recognize both methionine and adenosyl moieties of SAM. The candidate riboswitch from *A. nidulans* *agaA* requires further validation and will also not be considered in the review.

Contributing to the great variation in riboswitches and their respective ligands, glycine and lysine-responsive riboswitches possess a number of unique functional and structural characteristics. The lysine-binding riboswitch sensor, centered on a five-way helical junction, is among the longest and most complexly folded metabolite-sensing aptamers. The glycine riboswitch is also long, but is usually composed of two three-way junction-based glycine-binding modules cooperating with each other upon glycine binding. The lysine and glycine riboswitches efficiently discriminate between their cognate amino acids, precursors, related amino acids, and the cognate amino acids within a peptide context. Both riboswitches need also to recognize the carboxylate and ammonium groups of the ‘main chain’. Glycine riboswitches, on the other hand, interact with the smallest and simplest amino acid incapable of providing an extensive interface for specific RNA recognition and adaptive RNA folding. Our review aims to complement recent reviews on the functional [34–41] and structural [42–48] aspects of riboswitches and related gene regulation systems and more comprehensive reviews on purine [49], SAM/SAH [50], and eukaryotic TPP [14] riboswitches by providing a detailed analysis of the structure-function relationships of amino acid-responsive riboswitches. We will discuss the identification of lysine and glycine riboswitches, as well as their distribution and mechanisms of gene control. We will also present an in-depth characterization of the recent three-dimensional structure of a lysine riboswitch solved by X-ray crystallography [51, 52] and comment on the low-resolution structure of a glycine riboswitch determined by small angle X-ray scattering (SAXS) [53]. Lastly, we will provide a comparison of the structural and biochemical data available to date, discuss the contribution of a lysine riboswitch to bacterial antibiotic resistance [54–56] and potential applications of amino acid-responsive riboswitches.

Discovery and distribution of amino acid-responsive riboswitches

Identification of lysine- and glycine-specific riboswitches

The regulation of the aspartate family of biosynthetic enzymes in bacteria attracted the attention of researchers because three amino acids, lysine, methionine, and threonine, share a common biosynthetic pathway that begins with aspartate phosphorylation by the enzyme aspartokinase (Fig. 2a). Named the diaminopimelate (DAP) pathway [57], it is common to most bacteria and higher plants and extensively studied in several species. The biosynthesis of lysine in *Bacillus* species was particularly interesting because dipicolinate, a constituent

of bacterial endospores, is formed from an intermediate of the lysine pathway and the enzymes involved in the dipicolinate synthesis may be co-regulated because their activities increase during sporulation [58]. Though early studies did not identify a consistent pattern of DAP pathway regulation, the repression of biosynthetic enzymes by lysine was established for both Gram-negative and Gram-positive bacteria. In *Escherichia coli*, lysine represses the production of monofunctional aspartokinase LysC [59]. Mutants that constitutively synthesize the enzyme were mapped to the putative operator region [59, 60] located in the 5' UTR of the *lysC* gene [60, 61]. The same *lysC* gene encoding aspartokinase II was found to be repressed by lysine in *Bacillus subtilis* [62, 63]. As in *E. coli*, mutations that abolished lysine-dependent repression and conferred resistance to the toxic lysine analogue *S*-(2-aminoethyl)-L-cysteine (AEC, also known as thialysine) were located in the segment preceding the *lysC* gene [64, 65]. Bacteria bearing these single-nucleotide mutations and grown in the presence of lysine produced the full-length *lysC* mRNA, as well as wild-type bacteria grown under lysine starvation did. On the other hand, mRNA in wild-type bacteria grown in the medium containing lysine was truncated at the site of a putative transcription terminator [66]. Quantitative analysis indicated that the presence of lysine does not change the number of *lysC* mRNA molecules, but causes the stoichiometric replacement of full-length mRNAs with truncated molecules. These experiments suggested that *lysC* expression is controlled by the promotion of premature transcription termination in the presence of lysine, though the exact mechanism of regulation that would account for the large distance between AEC-resistant mutations and the potential terminator/antiterminator region was not resolved. Despite the divergence of these species, sequence similarities were observed in the long 5'-UTRs of *lysC* gene from *E. coli*, *B. subtilis* and thermophilic *Bacillus* sp. [60], supporting the role of the 5'-UTR in gene regulation [66]. In addition, these similarities ruled out a classical attenuation-type mechanism [67] because the *E. coli* gene lacked the *Bacillus* type transcription terminator. Moreover, the sequences of the regions potentially involved in the attenuation mechanism differed significantly in *E. coli* and bacilli. Though the mechanism of *lysC* regulation remained to be resolved, it was clear that it was different from the transcription activation mechanism of *E. coli lysA* gene carried out by the LysR protein [68, 69] and should somehow be related to a large RNA segment upstream of the coding sequence. In many other amino acid synthetic genes, long 5'-UTRs form tRNA recipient modules. However, searches for the T-box consensus sequences upstream of *lysC* were not successful [60]. The existence of a mRNA-binding lysine-sensing protein regulator was hypothesized [60, 66, 70] by analogy to the findings in other operons [5]. Nonetheless, the protein regulator was never found. In addition, the lysine-induced repression mechanism of *B. subtilis lysA* gene also remained a puzzle [71].

The long-awaited mechanism of gene regulation in several genetic systems has been uncovered by experiments in the Ronald Breaker and Eugene Nudler laboratories, which proved the direct sensing of coenzymes AdoCbl [9], FMN [8, 18] and TPP [8, 10] by mRNA regions termed 'riboswitches'. These findings were supported by extensive bioinformatics searches [72–74], revealing the widespread distribution of riboswitches in bacteria and their presence in eukaryotes [75], and by the experimental identification of SAM-responsive (SAM-I class) [16, 17, 19] and purine-sensing [23] riboswitches. A similar mechanism involving a conserved lysine-sensing riboswitch sequence (LYS or L-box) was predicted in the case of lysine regulation [76] and experimentally confirmed for *B. subtilis lysC* gene regulation [27, 29]. The riboswitch mechanism agreed well with previous experimental data, thereby unraveling the long-standing mystery of gene expression inhibition by lysine.

The availability of more than 90 sequenced bacterial genomes in conjunction with improved bioinformatics strategies rendered possible extensive comparative searches for the identification of novel regulatory motifs. The examination of long intergenic regions for indications of conserved sequences and secondary structure elements yielded several new

candidates for riboswitches [77]. One such wide-spread motif, termed *gcvT*, was found upstream of bacterial genes that express enzymes involved in the metabolism of glycine. In *B. subtilis*, the *gcvT* motif controls the expression of the three-gene operon *gcvT-gcvPA-gcvPB* coding for enzymes that catalyze the initial reactions for the use of glycine as an energy source. The *gcvT* motif is represented by two similar structural types (I and II) that have a common core flanked by distinct termini. The motifs of both types often reside adjacent to each other and function as two cooperative glycine-binding modules accompanied by a single expression platform [28]. This advanced integrated form of a riboswitch most likely evolved to ensure the rapid activation and repression of genes in response to small changes of glycine concentration.

Distribution of lysine and glycine riboswitches

The phylogenetic distribution of lysine and glycine riboswitches is more fragmented than the distribution of the more abundant riboswitches responsive to coenzymes TPP, AdoCbl, FMN and SAM (SAM-I class) [12]. Both glycine and lysine riboswitches are widespread in certain bacterial groups and completely absent in others (Fig. 3). Glycine riboswitches seem to have a much broader distribution than lysine riboswitches do, though both riboswitches have been identified in Gram-positive and Gram-negative bacteria. Glycine riboswitches were found in Firmicutes, Fusobacteria, Actinobacteria and all divisions of Proteobacteria. They are missing in Cyanobacteria, Acidobacteria and deeply branched bacteria, such as Thermotogae and *Deinococcus/Thermus*. Lysine riboswitches are common in Gram-positive Firmicutes and Gram-negative Fusobacteria, and, in contrast to glycine riboswitches, are present in Acidobacteria and absent in Actinobacteria. From all divisions of Proteobacteria, only species of γ -proteobacteria, including enterobacteria, Pasteurellaceae and Vibrionaceae, have lysine riboswitches. Together with TPP, AdoCbl and FMN riboswitches, lysine riboswitches were identified in *Thermotoga*. Glycine riboswitches are highly abundant in the sequences obtained from bacterial environmental communities [78, 79] and presumably belong to the species from Proteobacteria [80]. The majority of glycine riboswitches were observed in the Sargasso Sea metagenome. Moreover, 83% of riboswitches apparently originate from one species, *Candidatus Pelagibacter ubique*, from the SAR11 clade, which prevails in marine surface waters [81]. Lysine riboswitches are rarer than glycine riboswitches. They are present in the Sargasso Sea and Whale Fall metagenomes and are absent in the Minnesota Soil sample [12].

Genes under the control of amino acid-dependent riboswitches

Most of the identified lysine riboswitches reside upstream of lysine biosynthetic genes (Fig. 2a), which in different bacteria could be found as single genes scattered along the chromosome (e.g., *E. coli*) or as multiple gene clusters organized in operons (e.g., *Staphylococcus aureus*). In accordance with the organization of lysine biosynthetic genes, the distribution of lysine riboswitches varies significantly, though most often L-boxes are detected upstream of the *lysC*, *lysA* and *dapA* (dihydropicolinate synthase) genes of Gram-positive bacteria, including *B. subtilis*, and upstream of the *lysC* genes of γ -proteobacteria, including *E. coli* [27, 29, 76]. Occasionally, lysine riboswitches control other genes, such as the *dapX* gene from *Leuconostoc mesenteroides* that encodes a putative DAP epimerase that possibly substitutes for the product of the *dapF* gene found in other species [76]. In the case of operon organization, riboswitch candidates also precede other genes, for instance the *asd* (aspartate semialdehyde dehydrogenase) gene positioned in the beginning of the large cluster *asd-dapF-dapA-dapB-lysC-lysA-ukuR* in *Thermotoga maritima* genome.

In addition to the genes involved in lysine biosynthesis, L-boxes are also located upstream of genes encoding several lysine transporters. In Gram-positive bacteria, lysine riboswitches regulate lysine transport carried out by the lysine-specific permease encoded by *lysP* and a

putative transporter *yvsH* (the *B. subtilis* gene name). In several species, L-boxes were found upstream of the candidate components of an ATP-dependent lysine transport system, designated *lysX-lysY* [76]. Another gene under the control of the lysine riboswitch, *nhaC* (named *lysW* in [76]), belongs to the NhaC Na⁺/H⁺ antiporter superfamily. Though this gene might code for a candidate lysine transporter in species that lack the *lysP*, *yvsH* and *lysX-lysY* transporter genes [76], another intriguing possibility is that the gene could encode a functional Na⁺/H⁺ antiporter. It was suggested that if NhaC participate in H⁺ transport, the lysine riboswitch might be involved in the regulation of cellular pH [29]. In certain organisms, the activity of lysine decarboxylase, which consumes a proton upon conversion of L-lysine to cadaverine, is used to control pH. Instead of the lysine decarboxylase system, the riboswitch-controlled NhaC may assume the regulatory role, becoming deactivated in the presence of high levels of cellular lysine [29]. The role of NhaC in the pH homeostasis of the alkaliphilic *Bacillus firmus*, however, appears to be subtle at alkaline pH and more pronounced when cation concentrations are suboptimal [82]. Moreover, the *nhaC* gene was unable to complement the alkali-sensitive phenotypes of neutrophilic *B. subtilis* and alkaliphilic *B. firmus* species [82] arguing against its dominant role in pH homeostasis [83].

The search for other genes under the control of lysine riboswitches has suggested a possible role of riboswitches in lysine utilization. In *Thermoanaerobacter tengcongensis*, one of two L-boxes was identified upstream of the *pspF3* gene cluster [29, 76] containing the genes homologous to the *kamA*, *kamD* and *kamE* lysine aminomutase genes, which are involved in lysine catabolism in *Clostridium* [84]. Candidates for lysine riboswitches were also located upstream of the *kam* cluster in *Fusobacterium nucleatum* [76]. Unlike the biosynthetic genes that are obviously repressed by lysine, the catabolic genes should be activated by lysine.

In contrast to lysine riboswitches, glycine riboswitches most often occur upstream of the *gcvTHP* operon encoding the glycine cleavage system, which breaks down glycine to produce 5–10-methylene-tetrahydrofolate, ammonia and carbon dioxide [77]. 5–10-methylene-tetrahydrofolate can be used directly in the synthesis of serine and 2'-deoxythymidyl-5'-monophosphate and can be converted to the compounds involved in variety of pathways, including methionine and purine biosynthesis. The second frequently observed glycine riboswitch, accounting for ~20% glycine riboswitches in metagenomes, originated from the most abundant marine bacterium *Cand. P. ubique*, where it controls the malate synthetase gene *glcB* [80] (Fig. 2b). In *Cand. P. ubique*, many genes involved in carbon metabolism and regulation are missing but the bacterium retains the ability to utilize an exogenous supply of glycine derivatives to make malate from glyoxylate using malate synthetase [85]. Therefore, the *glcB* and *gcvT* glycine riboswitches modulate the production of critical enzymes and play a central role in the carbon metabolism of *Cand. P. ubique* by regulating the flow of carbon through the glycine cleavage system and tri-carboxylic acid (TCA) cycle in response to the presence of glycine. The selective pressures that led to the control of TCA cycle by an indirectly involved compound are not clear, though they are likely connected with the adaptation of the bacterium to highly oligotrophic marine ecosystems through the effective use of dissolved organic carbon [85]. Glycine-sensing modules were located upstream of several other genes involved in the conversion and synthesis of glycine [28, 77, 80]: *glyA* (glycine/serine hydroxymethyltransferase), *sdaA* (L-serine deaminase), *dadA* (glycine/D-amino acid oxidase), and *serC* (transaminase). The *gcvT* element can often be traced nearby to the *yaaJ* genes for putative Na⁺/alanine symporters, and it is sometimes located upstream of the *ansP* gene for γ -aminobutyrate permease and several other uncharacterized genes and genes not directly related to glycine biosynthesis/catabolism.

Organization and mechanism of riboswitch control

Mechanisms of gene expression control by lysine and glycine riboswitches

The metabolite-sensing domains of lysine riboswitches are centered on a five-way helical junction, and, with the exception of aptamers in some bacteria such as *T. tengcongensis* and *F. nucleatum*, are followed by expression platforms that form regulatory hairpins. In Gram-positive bacteria, e.g. the *B. subtilis lysC* gene, these hairpins are flanked at the 3' side by stretches of uridines, characteristic of ρ -independent transcriptional terminators (Fig. 4a). The 3' region of the sensing domain and the 5' region of the terminator are typically complementary to each other and, in the absence of lysine, can form an alternative 'anti-terminator' stem-loop structure. Therefore, riboswitch-dependent regulation proceeds through a termination/antitermination mechanism that results in the elevated expression of genes in the absence of lysine and in the inhibition of expression at excess concentrations of lysine. This mechanism was experimentally confirmed both by a single-round transcription *in vitro* and by a *lysC-lacZ* transcriptional fusion assay *in vivo*, using riboswitches from the *B. subtilis lysC* leader [27, 29]. In the presence of lysine, the sensing domain was stabilized and a terminator hairpin was formed, thereby causing RNA polymerase to stall at the terminator and release the nascent RNA to prematurely terminate gene transcription. In the absence of lysine, the lysine-dependant structural transition did not occur and mRNA adapted an anti-terminator conformation, thereby allowing RNA polymerase to pass through the entire gene. In Gram-negative bacteria, in the presence of lysine the regulatory hairpin overlaps with the ribosome-binding site (RBS or Shine-Dalgarno sequence) of the downstream gene and prevents ribosome loading onto mRNA thus functioning as a translational sequester (Fig. 4b). In the absence of lysine, as in transcriptional regulation, complementary regions of the sensing domain and the RBS-sequestering hairpin form an alternative hairpin (anti-sequester) that does not engage the ribosome-binding site in base pairing, rendering it accessible to ribosomes. Both transcriptional termination and translational occlusion mechanisms repress the expression of lysine-related genes in the presence of lysine. On the other hand, the lysine utilization genes in *T. tengcongensis* and *F. nucleatum* should be activated in the presence of lysine. The sensing domains and transcriptional terminators in these genes overlap and, therefore, a mode of regulation most likely involves the direct interplay between two alternative conformations. In the presence of lysine, the lysine-bound sensing domain prevents the formation of the terminator and acts as an anti-terminator. In the absence of lysine, the terminator conformation prevails leading to the pre-mature termination of transcription.

Unlike other riboswitches, glycine riboswitches typically contain tandem receptor domains followed by a single expression platform that may contain either transcription terminator, as in the *B. subtilis gcvT* riboswitch [28], or translation initiation signals, as in the *Cand. P. ubique gcvT* riboswitch [85]. Since most glycine riboswitches modulate genes related to the utilization of glycine, glycine riboswitches serve as activators of gene expression in the presence of glycine. The mechanism of gene expression activation is similar to the mechanism employed by lysine riboswitches that control lysine utilization genes. In the absence of glycine, the glycine-sensing domains are not stabilized and the expression platform forms a transcription terminator that prevents gene transcription (Fig. 4c). In the presence of glycine, the glycine-sensing domains adopt a stable conformation. The 3'-terminal sensing module, adjacent to the expression platform, precludes the formation of the transcription terminator, thereby turning gene expression on. Despite the obvious similarity in the gene expression mechanisms used by glycine and other rare riboswitches that activate gene expression, tandem glycine riboswitches possess a unique feature that distinguishes them from all riboswitches including riboswitches with other tandem arrangements [86, 87]. Though each individual sensing domain of the glycine riboswitch is capable of specific binding to a separate target molecule, the tandem domains display a cooperative mechanism

of ligand binding. The concerted response to glycine was first observed following in-line probing of the tandem riboswitch from *Vibrio cholerae* [28]. The experiment showed that each section of the riboswitch, namely type I domain, interdomain linker, and type II domain, undergo equal structural changes at various glycine concentrations, suggesting that either both aptamers have identical affinities for glycine or that they bind glycine in a highly cooperative manner. The cooperativity of glycine binding was supported by other experiments that documented different ranges of response to glycine in a single domain and tandem domains. While the change from ~10 to ~90% of glycine-bound single domain occurred over a ~100-fold increase in glycine concentration, the same adjustment in the double-domain riboswitch was detected over only a ~10-fold increase in glycine concentration. Consequently, the glycine binding curve for the single domain is normal hyperbolic, as opposed to the sigmoidal or *S*-shaped curve for the double domain. The reduction in the dynamic range of the response and the sigmoidal shape of the binding curve can be explained by an improvement in glycine binding affinity at one site when the second site is occupied by the ligand. Indeed, a mutation that disrupted binding in one domain caused a significant decrease in the binding affinity of the adjacent unaltered site. The degree of cooperativity for the *V. cholerae* riboswitch, expressed by the Hill coefficient, gave a value of 1.64, whereas the maximum value for two cooperative binding sites is 2 and the noncooperative value is 1. The *yaaJ* glycine riboswitch from *F. nucleatum* [85], and the *gcvT* riboswitches from *B. subtilis* [28] and *Cand. P. ubique* exhibited only slightly smaller Hill coefficient of 1.4 [88]. The degree of cooperativity in glycine riboswitches is comparable with the value of 2.8 (maximum for four binding sites is 4) determined for the oxygen-transport protein hemoglobin [89], one of the best characterized biological systems that use binding cooperativity [90]. Interestingly, the *glcB* glycine riboswitch from *Cand. P. ubique* was practically devoid of cooperativity (Hill coefficient of 1.1), leading to a more gradual activation of gene expression in response to increasing concentrations of glycine [85]. Therefore, two glycine riboswitches from *Cand. P. ubique* are tuned to different ranges of glycine concentration and seem to act collectively but not necessarily synergistically. This suggestion agrees with the fact that the riboswitches control two indirectly connected systems, the glycine cleavage and TCA cycle [85]. Since no intrinsic terminators are apparent, both riboswitches appear to utilize the mechanism of ribosomal binding site occlusion for gene regulation (Fig. 4d).

Sequences and secondary structures of lysine and glycine riboswitches

Lysine and glycine riboswitches were characterized using several biochemical, genetic and biophysical techniques aimed to identify the riboswitch sequence, its effector ligand, and the mode of genetic control. Aside from the demonstration of a genetic control *in vivo*, each riboswitch candidate was subjected to the ultimate test - direct binding with the candidate ligand *in vitro*, as well as to a careful phylogenetic analysis. The consensus sequences and secondary structures of riboswitches (Fig. 5) [12, 27–29, 76] showed that both lysine and tandem glycine riboswitches are among the longest and most complexly organized riboswitches. The metabolite-sensing domain of the lysine riboswitch adopts a composite fold closed by a single helix P1 and organized around a highly conserved purine-rich five-way helical junction between base-paired elements P1, P2/P2a/P2b, P3, P4 and P5 (Fig. 5a). Each paired element, except P5, conforms to specific length restrictions, suggesting that they all directly participate in structure formation. In certain sequences, an additional stem-loop structure P2c is predicted [12], while the majority of riboswitches contain a loop E [91, 92] or *S*-turn motif between P2 and P2a helices [27, 29] and a K-turn motif [93] in the P2b element [27]. The apical loops of stem-loops P2b and P3 exhibit complementary residues with an extensive level of covariation indicative of pairing with each other [27, 29, 76]. A mutational analysis showed that both K-turn and ‘kissing’ loop-loop interactions are crucial for lysine-dependent transcription attenuation [94].

Unlike the lysine riboswitch, the glycine-sensing module is composed of two separate aptamer domains (type I and type II), each capable of independent binding with a ligand [28] and cooperative functioning in a tandem arrangement (Fig. 5b). The domains are typically connected by a ~8 nt linker of limited sequence conservation [12, 28]. The length of the linker might be important for the cooperative response to glycine. *Cand. P. ubique glcB* riboswitch with a 14 nt linker does not support positive interdomain cooperativity [85] while riboswitches with shorter 7–10 nt linkers exhibit a normal cooperative response upon glycine binding [28, 85, 88]. Both glycine-sensing domains demonstrate a high degree of similarity in the core region centered on the evolutionarily conserved purine-rich three-way junction between the helix P1, slightly variant in type I and II domains, the variable helix P2 and the conserved helix P3/P3a. An additional hairpin may extend from the P3a element and radiate from the junctional region in certain species [12, 28]. The P3 and P3a elements are separated by a conserved purine-rich internal loop that contains consecutive adenines often found in “A patches” involved in tertiary A-minor motifs [95]. The Rfam database [96] lists several single domain glycine riboswitches that appear to belong to both aptamer types; further studies are needed to evaluate the response to glycine and the biological role of these riboswitches.

Ligand recognition by amino acid-responsive riboswitches

Ligand recognition: binding specificity

A key characteristic of any riboswitch is the ability to discriminate between similar natural compounds in order to modulate the gene expression of a particular gene/set of genes in the presence of a correct chemical cue. If the wrong compound is recognized by a riboswitch, bacteria may reduce or shut down the expression of genes essential for their viability. Bacteria may also waste resources and energy to activate the expression of genes not needed at the moment, thereby decreasing their chances to outcompete for survival other species inhabiting the same ecological niche. Riboswitches often regulate biosynthetic genes via feedback inhibition using their end product. In this case, riboswitches have to be tuned to differentiate between the synthesized effector molecule and its precursors or derivatives. For example, SAM and SAH riboswitches distinguish coenzyme SAM from its reaction byproduct SAH, which lacks only a single methyl group and a positive charge of a sulfur atom [50]. Glycine and lysine riboswitches share the ‘main chain’ amine and carboxyl functionalities with all amino acids and, therefore, must be adapted to discriminate against 19 other common amino acids, their precursors and conversion products. In addition, the amino acid-responsive riboswitches should interact with a free amino acid outside the peptide context.

In-line probing, equilibrium dialysis, and *in vitro* transcription experiments showed that both lysine and glycine riboswitches exhibit high specificity towards lysine and glycine, respectively. The lysine riboswitch binds L-lysine with an apparent K_D of 1 μ M, while biosynthetic precursor DAP and many closely related analogs, including D-lysine, possess much lower binding affinities [29, 56] and reduced abilities to terminate transcription of the gene [27, 29, 56]. The riboswitch, however, can tolerate chemical modifications to certain groups of lysine [56]. A large functional group can be accommodated on N6, as long as this amine carries a hydrogen bond donor and is adjacent to a positive charge, as in L-homoarginine (HArg) and N6-1-iminoethyl-L-lysine (IEL) (Fig. 6a). On the contrary, the main chain amine, essential for lysine recognition, cannot be modified without the loss of binding. Modifications of various sizes can also be accommodated at position 4 of the ‘side chain’, including the double bond between C4 and C5 in DL-*trans*-2,6-diamino-4-hexenoic acid (DHA), and substitutions of the carbon atom by oxygen in L-4-oxalysine, sulfur in antibiotic AEC, and even the bulkier group in L-3-[(2-aminoethyl)-sulfonyl]-alanine (ASA). Most changes, however, decrease the binding affinity by ~2.5–13 times in equilibrium

dialysis experiments [56] and reduce the stability of the RNA fold by ~8–100 times in primer extension experiments [52]. The removal of the carboxylate group of lysine ablates riboswitch binding, though the group can be modified with only modest reduction in L-lysine hydroxamate [29]. The large collection of lysine analogs tested for riboswitch binding has allowed the definition of molecular recognition determinants of the lysine riboswitch receptor [56] (Fig. 6b). This summary indicates that all lysine functionalities, the carboxylate and ammonium groups of the main chain, and the ammonium group of the side chain, are required for efficient binding to the riboswitch. In addition, lysine-like ligands for the lysine riboswitch should obey certain length restrictions and their side chains must fulfill several additional criteria.

Despite a large number of candidate effector molecules for the glycine riboswitch, glycine was found to be the only potent ligand with an apparent K_D of 10 μ M for the individual type II riboswitch domain [28]. Other small amino acids, such as alanine and serine, did not activate the riboswitch [28, 85, 88]. Among the glycine derivatives tested, only four analogs, glycine methyl ester, glycine ethyl ester, glycine tertiary butyl ester, and glycine hydroxamate showed measurable binding affinities that were, however, reduced 10–100 times in comparison to glycine [28, 88] (Fig. 6c). All of these analogs carry modifications at the carboxylate moiety and demonstrate a decrease in binding affinity with an increase in the size of the substituent group. Modification of the amino group abolished binding to riboswitch, suggesting that the riboswitch envelopes the amino terminus of glycine while the ligand binding pocket around the carboxy terminus is either semi-opened or flexible enough to tolerate bulky substitutions (Fig. 6d). Unlike the glycine riboswitch, the ligand binding pocket in the lysine riboswitch has two openings (see further) that tolerate certain extensions of the carboxy and N6 amino groups (Fig. 6b). Interestingly, the carboxylate shared by all amino acids does not seem to be occluded by glycine and lysine riboswitches. It is not clear whether this observation reflects a lack of discriminatory functionality in the carboxy group or difficulty in recognizing a negatively charged group by RNA.

Ligand recognition: location of the binding sites and ligand-induced riboswitch folding

Methods to study ligand binding and riboswitch folding—One of the most intriguing questions in the riboswitch field is how RNA molecules, composed of only four simple and similar building blocks, can form a structure that recognizes small organic molecules with exquisite specificity. To address this question, numerous biochemical and biophysical techniques, ranging from RNA footprinting to three-dimensional structure determination, have been employed to various riboswitches, including glycine and lysine riboswitches. The easiest approach to elucidating the secondary RNA structure, mapping a ligand binding site, and detecting ligand-induced conformational changes, involves RNA structure probing and footprinting with chemical and enzymatic probes [97, 98]. One of the simplest probing/footprinting techniques, termed in-line probing, takes advantage of the inherent chemical instability of RNA [99]. Internucleotide linkages of conformationally unrestricted RNA segments, such as bulged nucleotides and loops, are cleaved faster and can be easily detected by subjecting the radioactively-end-labelled RNA incubated in Mg^{2+} -containing buffer at elevated pH to denaturing polyacrylamide gel electrophoresis (PAGE). Since ligand binding typically restricts RNA dynamics, the ligand-binding sites can be mapped on RNA if the RNA-ligand complex was assayed in parallel with the RNA sample. Unfortunately, in-line probing assesses the ligand-induced cleavage restriction of only nucleotides that are reactive in the free RNA. As some other footprinting techniques do, in-line probing cannot distinguish between nucleotides directly bound to the ligand and nucleotides involved in conformational adjustments induced by ligand binding. Alternatively, ligand-induced modulations can be visualized by the selective 2' hydroxyl acylation analyzed by primer extension (SHAPE) method, which also probes the relative

flexibility of RNA linkages. SHAPE is based on the reactivity of the 2' ribose hydroxyls of RNA to an electrophile, e.g. N-methylisatoic anhydride [100, 101]. The nucleophilic character of the hydroxyl is influenced by its proximity to the adjacent 3'-phosphodiester anion, which is, in turn, governed by local structural context. Another chemical footprinting technique utilizes hydroxyl radicals generated in the Fenton reaction [102]. Due to their small size and high reactivity, hydroxyl radicals cleave the RNA backbone that is accessible to solvent, irrespective of the sequence and almost independently of secondary structure. Hydroxyl radical footprinting can probe all nucleotides in a RNA molecule and can be used for mapping both ligand binding sites and tertiary interactions. However, hydroxyl radical footprinting generates high background and is not easily reproducible for weak complexes. Another powerful method, called nucleotide analog interference mapping (NAIM), can assess individual functional groups in various nucleotides [103]. NAIM is based on the random incorporation of modified nucleotides into RNA, as well as the physical separation of the modified RNAs that retain binding capacity from those that lost this ability [104]. To assign the modified positions, all nucleotide analogs contain a phosphorothioate linkage (5' nonbridging oxygen atom replaced with a sulfur atom) which can be cleaved with iodine. Using various nucleotide analogs, every occurrence of a particular functional group in an RNA molecule can be simultaneously tested for its importance in the formation of the ligand-recognizing structure and its direct contribution to ligand binding. Chemical probing of RNA is often complemented by enzymatic footprinting which utilizes various structure-sensing ribonucleases, such as nuclease V1, specific to helical RNA regions, and nuclease T2, specific to non-paired RNA segments [97]. In addition to secondary structure probing and footprinting, ribonucleases, due to their large size, are particularly useful for monitoring large ligand-induced conformational re-arrangements. Conformational changes can be also tracked by gel-retardation and fluorescent assays, successfully utilized to study folding of lysine riboswitch [94] and other riboswitches [105–107], and by more sophisticated NMR [108, 109] and single-molecule force [110] approaches.

Recognition of glycine by the glycine riboswitch—In-line probing experiments performed on the tandem *V. cholerae gcvT* riboswitch are in overall agreement with the secondary structure proposed on the basis of phylogenetic analysis [28, 88] (Fig. 7a). Virtually all predicted bulges and unpaired loops in riboswitch domains, interdomain linker, and four-way junctional region in the 5' receptor (type I domain) are cleaved spontaneously in the absence of glycine. On the other hand, the presumably unpaired nucleotides of the three-way junction in the 3' receptor (type II domain), with the exception of two purines, are not cleaved, thereby suggesting their participation in the stable RNA conformation. The discrepancies in the cleavage patterns between the two domains might be associated with the tandem organization of the riboswitch and not with the differences in the junctional organization of the glycine receptors. Indeed, the individual *V. cholerae* type II domain has additional cleavage sites in the junctional region (nts 212–214) that corresponds to nts 83–86 of domain I. In addition, the in-line probing performed on the tandem riboswitch from *F. nucleatum* demonstrates that the three-way junction-based *F. nucleatum* type I domain produces a similar cleavage pattern to the *V. cholerae* four-way junctional domain, while the three-way junction-based *F. nucleatum* type II domain is more reactive in the junction (the region corresponds to nts 212–216 and nt 218 in *V. cholerae*, Fig. 7a) as compared to the corresponding domain from *V. cholerae*, similar in sequence and organization [88].

In the *V. cholerae gcvT* riboswitch, glycine-induced cleavage modulations coincide with the sites of conserved nucleotides and cluster in the internal loop between P3 and P3a, the apical/internal loop adjacent to P3a, the linker region, and the junctional region of domain I [28]. The cleavage reduction/enhancement pattern is largely preserved in *F. nucleatum* [88], with an additional region of reduced cleavage on the 3' side of the junction that overlaps with the glycine-induced 'protections' in the individual *V. cholerae* domain II [28]. The

glycine-dependant modulations are similar in both domains of the riboswitch, suggesting that each domain utilizes similar means for individual glycine binding. The sites with equivalent modulation may also be involved in cooperative interactions that are common to both aptamer domains. On the other hand, the absence of sites with asymmetrical modulations, aside from the species-variable modulations in the three-way junction of domain II, impairs the prediction of RNA regions that participate in cooperative interactions. The asymmetric sites are good candidates for nucleotides responsible for domain cooperativity because these sites should not participate in glycine binding expected to be similar in both riboswitch domains.

The in-line probing data are consistent overall with the hydroxyl radical footprinting results from the *V. cholerae gcvT* riboswitch (Fig. 7a) [53]. The strongest glycine-induced hydroxyl radical footprints were observed in the junctional regions and the P3/P3a stems of both domains and are overlapped or expanded around the positions showing in-line cleavage reductions. The hydroxyl radical footprints also extend to the P1 and P3a helices and the junctional nts 147–150, which were not reactive in the in-line probing experiments.

To gain further insights on glycine recognition by the riboswitch and the molecular basis of cooperativity, NAIM was employed on *V. cholerae* and *F. nucleatum* tandem riboswitches [88] using backbone modifications (parental phosphorothioate nucleotides and 2'-deoxynucleotides) and purine analogs specific for N7 (7-deazapurine nucleotides) and Watson-Crick (N-methylpurine nucleotides) positions. The experiments revealed multiple sites of interference that are identical in both species and mostly overlap with regions that show glycine-induced cleavage reduction in in-line probing and hydroxyl radical footprinting (Fig. 7a). Predominantly base-specific interferences were observed in the bottom halves of the junctions, while mixed base- and backbone-specific interferences were found in the top part of the P1 helix and the P3/P3a region. In contrast to cleavage reductions in in-line probing and similar to those in hydroxyl radical footprints, symmetrical interferences were detected in the bottom part of the junctions in both domains; no interferences were observed in the upper junctional regions. Therefore, the bottom junctional regions and the P3/P3a stems likely participate in the formation of glycine binding pockets. The interdomain linker does not appear to be directly involved in glycine recognition and domain cooperativity, given that this region does not demonstrate any interferences and shows only weak hydroxyl radical footprints. The glycine-induced modulations in this region, detected by in-line probing [28], probably reflect global conformational adjustments upon glycine binding. Given the large number of evolutionarily conserved nucleotides modulated by glycine in footprinting/NAIM experiments, the up-to-date data do not permit the precise identification of the residues that interact with glycine. By analogy with the purine and lysine riboswitches that recognize small ligands [51, 52, 111, 112], one may consider the junctional region, which shows multiple base-specific interferences, to be the primary glycine recognition determinant (Fig. 7b). The P3/P3a stem might then be folded back towards the junction to sandwich the ligand. Mutations in the junction (G17C and G15C) significantly reduce glycine binding, supporting their role in the ligand recognition [28]. The same mutations negatively affect domain cooperativity; therefore the junctional regions might also participate in tertiary interdomain interactions.

Three-dimensional folding of the glycine riboswitch and the molecular basis for cooperativity—Though the majority of interferences in *V. cholerae* and *F. nucleatum* tandem riboswitches were identical in both domains, several sites were found in one domain and not in another (Fig. 7b) [88]. These asymmetric positions might be involved in the tertiary interactions responsible for riboswitch cooperativity because, as stated earlier, they do not likely associate with the chemical features important for ligand recognition by a single aptamer. The P1 helix of domain 1 displays asymmetric interferences in the ribose of

A13 and the N2 of G127, suggesting the importance of functional groups in the helical minor groove. The removal of the minor groove exocyclic amine by replacing the C12-G127 pair with U-A decreased the glycine binding affinity and abolished cooperativity, while the alteration of the G127 conformation by the introduction of the wobble U12-G127 pair retained glycine binding but eliminated cooperativity. A mutation in the analogous pair in domain II did not affect glycine affinity or cooperativity supporting the hypothesis that the minor groove of P1 in domain I participates in the tertiary interactions involved in riboswitch cooperativity. The second region of asymmetric interference was observed within the P3a hairpin of domain I. Two consecutive N-methyl A interferences (A67 and A68) suggest a tertiary contact in the major groove of helix P3a. Extensive mutational analysis supports the possible role of the exocyclic amine of A67 in the tertiary structure of the riboswitch and suggests the involvement of the P3a helices from both domains in cooperative glycine binding. Despite the recognition of the P1 helix from domain I as a key contact for glycine riboswitch cooperativity, the binding partner for the helix is not convincingly identified. The logical assumption that the interdomain interaction occurs between the P1 helix from domain I and the P3/P3a hairpin from domain II is not fully supported by experimental data [88]. It should be noted that the alternative possibility is not completely ruled out. If the P3/P3a stem folds towards the junction, the apical P3a loop of the domain I may reach the P1 helix of the same aptamer and make contacts with the minor groove of the helix. The asymmetric interferences then can be explained by variations in the loop regions in both domains. Therefore, the interaction partner for the P1 helix and the molecular mechanism of domain cooperativity are yet to be determined.

Glycine binding and formation of tertiary interactions significantly change the shape of the tandem riboswitch, as revealed by a small-angle X-ray scattering (SAXS) experiments [53, 113]. The low resolution SAXS solution structure of the *V. cholerae* riboswitch in the free form represents an elongated “T”-shaped conformation, which becomes more compact upon the addition of millimolar concentration of Mg^{2+} . Though the assignment of the secondary structure elements to the density is not reliable, possible scenarios are described in ref. [53]. The addition of glycine to the Mg^{2+} -folded riboswitch caused a significant shift in electron density from the periphery towards the middle of the structure. The change is likely related to rearrangements upon the formation of glycine-binding pockets and molecular contacts involved in allosteric “communication” between riboswitch domains. The transition from the Mg^{2+} -folded riboswitch to the glycine-bound form seems to require the additional uptake of magnesium, thereby implying the Mg^{2+} dependence of glycine binding affinity. It will be interesting to understand whether metal cations directly mediate glycine-riboswitch interactions, as in the TPP [114–117], FMN [118] and lysine [52] riboswitches, or they mostly play an indirect structural role, as in purine riboswitches [106, 119].

The three-dimensional structure of the lysine riboswitch—The in-line probing experiments on the lysine-sensing receptor of the *B. subtilis* *lysC* riboswitch showed limited cleavage restricted predominantly to the five-way junction [29]. Other cleavage sites were mapped to the loop E motif, the K-turn, and the apical loops, except for loops L2 and L3, which are apparently engaged in the predicted kissing loop pairing. Lysine-induced cleavage reductions were mapped exclusively to the five-way junction and the regions of the P1 and P5 helices adjacent to the junction. The same experiments performed on the *T. maritima* *asd* riboswitch revealed a similar cleavage reduction pattern with more reductions in the bottom of P4, fewer reductions in P5, and weak reductions in the loop E motif and the P2a-L2 turn detected at large excess of lysine (Fig. 8a,b) [52]. The results from SHAPE experiments on riboswitches from both species generally agreed with those from in-line probing experiments [51].

The lysine-sensing domain from *T. maritima* *asd* riboswitch was crystallized in the lysine-bound form and its three-dimensional X-ray structure was determined independently by two groups. The first structure, determined at a resolution of 1.9 Å, includes a full-length 174-nt lysine sensor (Fig. 9a,b) [52], whereas the second structure encompasses the riboswitch with a shortened helix P5 and has a diffraction limit of 2.8 Å [51]. Both structures are in good agreement except for a slight variation in the conformation of loop L4 and differences in bound cations. According to the phylogenetic analysis and in-line probing, the tertiary structure of the L-box consists of three-helical and two-helical bundles radiating from a compact lysine-bound five-way junction. The junction is organized through the collinear stacking of helices P1 and P2, and helices P4 and P5. The P2/P2a stem reverses its orientation by making two turns around the canonical loop E motif and the novel P2a/L2 turn, which replaces a canonical K-turn found in other lysine riboswitches. Despite different nucleotide composition and pairing, the conformation of the P2a/L2 turn is reminiscent of the canonical K-turns [93] and a turn from the P4-P6 domain of group I introns [120]. A kissing loop interaction between loops L2 and L3 aligns stems P2 and P3; stems P2 and P4 are anchored parallel to each other by novel interactions between the terminal pentaloop of L4 and the irregular helix P2. Unlike previously described kissing loop interactions [121], the L2-L3 interactions are stabilized by a molecular 'staple' composed of a pair of stacked residues G43 and U94 which are positioned perpendicular to the bases of the kissing loop helix and which form hydrogen bonds with three consecutive base pairs (Fig. 9c). Such an element might constitute an adaptation to the optimal physiological temperature of 80° C for *T. maritima*, which inhabits geothermal-heated marine sediment. Another possible thermostable adaptation involves the lysine-induced long-range stacking interaction between G110, not present in riboswitches from mesophilic bacteria, and A164 (Fig. 9d). This interaction restricts the movements of helices P1 and P3 and should stabilize the overall riboswitch structure. Similar to the previous observation about RNA-protein complexes, the thermophilic riboswitch demonstrates ~30-fold higher affinity for lysine than does its counterpart from mesophilic *B. subtilis*.

Molecular basis of lysine recognition by a riboswitch—Lysine binds the riboswitch in an extended conformation and is positioned centrally within the pocket formed by the evolutionarily conserved core of the five-way junction. The amino acid sits on the top of purine tetrad formed by nucleotides of the P1 and P5 helices, is flanked by the first G-C and G•U base pairs of the P2 and P4 helices, and is covered by the G-C pair of the helix P2 and two purines of J2-3, located in the beginning of the helix P5 (Fig. 10a) [51, 52]. The carboxylate and ammonium groups of the lysine main chain form direct hydrogen bonds with the minor groove edges of guanosine bases and 2'-hydroxyls of sugars (Fig. 10b). The additional interaction of carboxylate with RNA is mediated by a K⁺ cation [52], which binds nucleotides from different layers of the junction. The side chain ammonium group provides hydrogen bond donors to a sugar ring oxygen, a non-bridging phosphate and a tightly bound water molecule. The methylene linker of the side chain threads through the tight pocket, which cannot fit bulky and branched amino acid side chains. Nevertheless, the pocket is not tightly packed against lysine between positions C4 and N6; therefore lysine analogs with modifications at C4 should be accommodated within the pocket (Fig. 10c). The crystal structures of the riboswitch bound to L-4-oxalysine and to the antibiotic AEC confirm the biochemical experiments and demonstrate that these compounds can be placed inside the pocket without causing changes in RNA conformation [52]. The binding experiments also suggested that certain analogs with extensions at the side chain ammonium group can be accommodated by the riboswitch [56]. Indeed, L-homoarginine and IEL take advantage of the opening next to the ε-ammonium group and, as revealed by the crystal structures of the complexes, bind the riboswitch with only a slight shift towards G80 in the terminal segment of the side chain and a small adjustment in the adjacent ribose (Fig. 10d) [52]. Though these

analogs lack the positive charge at the N6 atom and lose two hydrogen bonds from this position, the nitrogen atoms of the guanidinium group of L-homoarginine substitute water molecules w1 and w2 and restore their hydrogen bond patterns. Despite emergence of better stacking interactions and additional hydrogen bonds between the guanidinium group and RNA, L-homoarginine binds the riboswitch weaker than lysine does, emphasizing the potential contributions of electrostatic interactions and the fine structural complementarity between the ϵ -ammonium group and the binding pocket. Interestingly, not all contacts made by the guanidinium group of L-homoarginine seem important since IEL, which has a nitrogen atom substituted by a carbon atom, retains its binding ability (Fig. 6a). The removal of the remaining positive charge by the oxygen substitution in N⁶-acetyl-L-lysine abolishes binding [56]. The contacts formed by the ϵ -ammonium group of lysine play a direct role in the discrimination against other amino acids, e.g. glycine and alanine, which lack this functional group but can fit the pocket and form the same interactions using the main chain functionalities. Among the amino acids that contain the terminal ammonium group, the shorter L-ornithine should not make productive interactions, whereas the longer L- α -homolysine should have its side chain protruding out of the pocket. As expected, L-ornithine does not bind the riboswitch while L- α -homolysine bind the riboswitch, but weakly [29]. The vast majority of other amino acids and their derivatives, e.g. D-lysine and DAP, are rejected through indirect readout by the tight binding pocket. Interactions with an amino acid within a peptide context are also prevented by a steric block achieved through the formation of a rigid pocket around main chain functionalities. The RNA, however, does not completely envelope the carboxylate moiety and leaves an opening occupied by a K⁺ cation (Fig. 10c). The existence of this opening draws a striking parallel with the glycine riboswitch, which can tolerate substitutions in the carboxylate functionality. Consistent with the weak binding affinity observed for L-lysine hydroxamate [29], the opening might accommodate methyl, ethyl, or even longer esters. Such modifications will most likely replace the K⁺ cation, and given the important contribution of this cation to lysine binding affinity, the effect of the carboxylate alterations could adversely impact on the biological function of the riboswitch.

Lysine-induced folding of the lysine riboswitch—The lysine-bound riboswitch structure demonstrates that lysine, encapsulated within the junctional core, cannot be bound inside the pocket without certain changes in the riboswitch structure. In contrast to other riboswitches, in-line probing and chemical footprinting showed only limited lysine-induced conformational adjustments located around the five-way junction. In support of the probing and footprinting experiments [29, 51, 52, 56], SAXS experiments did not reveal significant alterations in the global riboswitch conformation upon the addition of lysine to RNA pre-incubated in the presence of Mg²⁺ cations [51]. Sensitive to global conformational rearrangements, nuclease footprinting also detected the compaction of the junction and suggested the stabilization of the parallel alignment of the helices P2, P3 and P4 in the presence of lysine [52]. The key kissing loop interactions were proven to be preformed and only slightly strengthened by lysine binding, in agreement with previous biochemical data [94]. Taken together, these data suggest that the lysine riboswitch may adopt a near native global conformation in the absence of lysine, and that lysine binding promotes the formation and/or stabilization of the junction. To test this hypothesis, we [52] and others [51] determined the crystal structure of the lysine riboswitch in the absence of lysine. Surprisingly, the bound and free riboswitch structures show nearly identical RNA conformations with only several minor differences, e.g. ~0.5 Å shift (more pronounced in [51]) in the G12 position (Fig. 10e) and 0.2–1.3 Å perturbations of the stacked G110/A164 pair. Although the lack of the expression platform, the formation of long helix P1, and the presence of crystal packing contacts facilitate formation of the riboswitch conformation without bound ligand, the lysine riboswitch, in contrast to most other riboswitches, appears

to be largely pre-formed and requires predominantly local adaptive adjustments for efficient ligand binding.

Metal cations in the lysine binding and riboswitch folding— Mg^{2+} cations play a critical role in the function of various cellular RNAs mostly because of their ability to shield negative charges and facilitate the proper folding of RNA molecules [122]. Several natural and artificial riboswitches, such as TPP [115, 116] and FMN [118] riboswitches, the *glmS* riboswitch/ribozyme [123, 124], and the tetracycline sensor [125], employ Mg^{2+} to overcome the negative electrostatic character of their ligands. Divalent cations also form water-mediated contacts with adenine [112] and preQ_1 [126] in the corresponding riboswitch structures. Since bound adenine [112] and preQ_1 [126–128] form multiple hydrogen bonds and extensive stacking interactions with RNA, the divalent cations appear to be much less crucial for the ligand recognition than for riboswitch folding in these systems [107, 119, 127, 129].

It was rather surprising to find out that in contrast to all known riboswitches, lysine riboswitch utilizes the K^+ cation to partially mediate interactions with the carboxylate group of the bound lysine. Moreover, primer extension and equilibrium dialysis experiments proved the essential contribution of K^+ cations at concentrations over 10 mM to high affinity lysine binding, even in the presence of 20 mM concentration of Mg^{2+} [52]. The replacement of K^+ by Na^+ or the omission of K^+ from the reaction buffer dramatically decreased affinity for lysine (at least 30 folds) in both thermophilic *T. maritima* and mesophilic *B. subtilis*. Despite the fact that K^+ cations were previously found in key positions of RNA structures [130, 131] and interfaces of RNA-protein complexes [132], they have not been implicated as crucial factors for riboswitch function. Anomalous scattering data showed that the lysine-bound K^+ cation has a much higher occupancy than the other two K^+ cations found in the riboswitch structure [52]. Though transiently bound K^+ cations could play a role in riboswitch folding, it is reasonable to assume that the strategically placed K^+ cation of the highest occupancy that bridges RNA and bound ligand should account for the maximum contribution of K^+ to lysine binding.

If K^+ cations are so important for tight lysine binding, does the lysine riboswitch critically depend on Mg^{2+} cations, as many other riboswitches do? Earlier studies on the *B. subtilis* *lysC* riboswitch demonstrated that kissing loop interactions crucial for riboswitch function cannot be formed in the absence of Mg^{2+} , even in the presence of 100 mM K^+ [94]. Recent SAXS measurements established the significant compaction of riboswitch RNA after the addition of Mg^{2+} [51]. Consistent with SAXS data, SHAPE mapped multiple Mg^{2+} -induced conformational alterations to all helices of the riboswitch [51]. Nevertheless, Mg^{2+} alone does not provide a measurable lysine binding affinity, while the combination of Mg^{2+} and K^+ gives a high apparent K_D of ~ 3.0 μM for lysine binding [52]. Since Na^+ cations only partially compensate for the missing K^+ cations, it is conceivable that both Mg^{2+} and K^+ cations play essential roles in the function of the lysine riboswitch. The cations probably do not share responsibilities: Mg^{2+} cations likely maintain the overall riboswitch fold whereas K^+ cations stabilize the binding pocket and mediate lysine-RNA interactions.

The dependence of riboswitch folding on Mg^{2+} cations may be less pronounced in riboswitches from thermophilic bacteria, which have to utilize more stable secondary and tertiary RNA elements for maintaining the correct structure at elevated temperature. Preliminary equilibrium dialysis experiments demonstrated that the *T. maritima* *asd* riboswitch does not absolutely require Mg^{2+} and can bind lysine in the presence of K^+ with $K_D \sim 3$ μM (L. Huang & A. Serganov, unpublished), though the mixture of Mg^{2+} and K^+ still gave the best $K_D = 0.1$ μM . These experiments, however, were conducted at room temperature, which is much lower than the optimal growth temperature of *T. maritima*.

Since mesophilic temperature can negatively affect thermophilic systems [133], the riboswitch might show higher Mg^{2+} dependence at elevated temperature.

The most surprising result came from the analysis of the high resolution crystal structure of the lysine riboswitch determined from crystals grown in a mixture of Mg^{2+} , K^+ and Na^+ [52]. Practically all of almost 30 cations with octahedral coordination geometry had the coordination bond distances 2.25–2.85 Å characteristic for Na^+ and longer than 2.0–2.3 Å distances typical for Mg^{2+} . If Na^+ cations, indeed, occupy the majority of the cation-binding sites in the structure, can they replace Mg^{2+} in the functional assay? An equilibrium binding assay showed that Na^+ can partially substitute Mg^{2+} . The addition of Na^+ to K^+ improved the binding affinity at least ~5-fold (L. Huang & A. Serganov, unpublished). The preferential finding of Na^+ over Mg^{2+} in the crystal structure may occur due to the entrapment of Mg^{2+} cations by citrate ions used as a buffer for crystallization. Though citric acid has a weak chelating capacity and can potentially compete with some Mg^{2+} binding sites on RNA, citrate ions are important for growing large crystals and cannot be replaced by other anions, such as acetate ions. The chelating ability of citric acid (the stability constant $\log K_1$ is 2.8) [134], however, is in the same range as the values of individual nucleosides and nucleotides (e.g., guanosine: 3.0; ADP: 3.11; ATP: 4.0; UDP: 3.17; UTP: 4.02). Clearly, the folded RNA can form stronger binding sites for Mg^{2+} cations than its individual components, and these sites may be filled with Mg^{2+} cations even in the presence of citrate. Such strong and specific Mg^{2+} binding sites are apparently not present in the riboswitch. This conclusion is supported by the crystallographic identification of a strong binding site for Mn^{2+} (mimic of Mg^{2+} [135]) in the kissing loop region, aside from the previously established cation binding sites [52]. The second weak Mn^{2+} cation was found in the expected position close to the 5'-terminal tri-phosphate. Consistent with the hypothesis that the thermophilic riboswitch needs lesser contributions from Mg^{2+} , SHAPE located 4 regions of Mg^{2+} -induced adjustments in the *T. maritima* riboswitch, much fewer than ~10 clusters in the *B. subtilis* riboswitch (Fig. 8a,b) [51]. Single nucleotide changes were located in the middle of P1, in the P4 loop and in the loop E motif; and a cluster of modulations was identified in the P2a/L2 turn. These sites, however, do not match the cation-binding sites in the high-resolution structure [52], though in the lower resolution structure [51], 3 sites neighbor water molecules, which might be reinterpreted as Mg^{2+} cations with low occupancy. Lastly, the *T. maritima* riboswitch can be folded and crystallized in the presence of K^+ and Na^+ and in the absence of Mg^{2+} [52]. The resulting Mg^{2+} -free structure does not deviate from the structure obtained in the presence of Mg^{2+} ; however, the moderate 2.85 Å resolution precludes the reliable identification of all cation-binding sites and their comparison with sites from other structures. Taken together, the biochemical and structural data demonstrate that the *T. maritima* riboswitch requires Mg^{2+} cations for efficient ligand binding, but depends less on Mg^{2+} , at least at mesophilic temperature, than the *B. subtilis* riboswitch does. Mg^{2+} cations most likely impact riboswitch folding, yet the precise contribution of Mg^{2+} to the folding trajectory remains to be elucidated.

The exact role of the junctional K^+ cation is emerging from structural data but still needs further clarification. The cation forms many long (2.7–3.3 Å) coordination bonds and cannot be replaced by the larger hydrated Mg^{2+} or Na^+ cations, suggesting that it plays a highly specific structural role in stabilizing the junctional region. However, the lysine-binding pocket can be formed without K^+ , as evidenced by the lack of the cation in the lysine-free structure and in the lysine-bound structure determined from crystals grown in the absence of K^+ [52]. The simultaneous presence of K^+ and lysine strongly suggests that the cation serves as a lysine counter-ion to shield, at least partially, the negative electrostatic charge of the lysine carboxylate. Protein molecules have more means to compensate for the negative charges and do not use counter-ions for lysine binding. The evolutionary pressure for the preference of K^+ in the lysine riboswitch system over Mg^{2+} , successfully utilized as a

counter-ion in several other riboswitches, is not understood. Possibly K^+ is simply better suited to interact with highly charged carboxylate because of the larger size of the cation and its ability to lose associated waters easier than Mg^{2+} .

It is also tempting to speculate that lysine riboswitches might play a role, perhaps indirect, in K^+ homeostasis. Some lysine riboswitches precede the *nhaC* genes, which encode putative Na^+/H^+ antiporters [29, 76]. The NhaC antiporter together with other transporters may be involved in the maintenance of monovalent cations and pH regulation. In *B. subtilis*, a dominant antiporter involved in Na^+ resistance and pH homeostasis is encoded by *tetA(L)*. This tetracycline-metal/ H^+ antiporter exhibits monovalent cation/ H^+ antiporter activity and can exchange both Na^+ and K^+ cations [136]. Therefore, the uptake of K^+ cations by TetA(L) protein might be influenced by the concentration of Na^+ and H^+ , which, in turn, is governed by the activities of all cellular antiporters, including NhaC. In *B. subtilis*, however, the deletion of *nhaC* does not influence bacterial growth at various concentrations of K^+ [136], though the activity of NhaC may depend on certain environmental conditions.

Lastly, the preference of potassium over magnesium may be explained by higher concentration of potassium in bacterial cells. *E. coli* cells contain up to 450 mM potassium but are unable to grow if the internal potassium concentration drops below 150 mM [137]. In *B. subtilis*, 300 mM potassium was found to be the cytoplasmic concentration [138]. The physiological concentration of Mg^{2+} cations in bacterial cells is generally accepted to be at low millimolar range, though some studies report higher Mg^{2+} concentration in growing *E. coli* cells [139, 140]. The high cation concentration might be essential for the binding of lysine and the folding decision within the short window provided by the transcribing RNA polymerase.

Molecular basis of the riboswitch-based control

The key feature that defines riboswitch activity is the ability to form a stable metabolite-bound conformation, which can include the entire regulatory element (e.g., a RBS) or part of it (e.g., a 5' segment of the anti-terminator). The molecular mechanisms involved in the "entrapment" of regulatory elements are delineated by the overall riboswitch conformation and by the principles used for ligand recognition. Generally, the mechanisms can be classified based on the position of a ligand with respect to a regulatory element. According to this criterion, riboswitches can be divided into two categories, where the ligand binding site and the regulatory element are either separated or close to each other. The TPP riboswitch represents an example of the former mechanism [115, 116]. Two irregular helices of the riboswitch recognize extremities of TPP by their middle regions and become stabilized by the ligand 'bridge'. The stabilization propagates to the bottom regions of the helices and contributes to the formation of two stacked tetrads within the three-way junction, which, in turn, facilitate the formation of the top part of the regulatory helix P1. Riboswitches from the second category either make direct hydrogen bonds with the regulatory element, as in the SAM-III riboswitch [141], or participate in ligand binding less directly and specifically, as in purine riboswitches [111, 112], through stacking and the formation of the local RNA conformation that traps the regulatory elements. The lysine riboswitch clearly exploits the second molecular mechanism. In the riboswitch structure, the bound lysine and its partner base pairs G12-C79 and G114•U140 are situated on top of the purine quartet that constitutes the top part of the regulatory P1 helix (Fig. 10f). Therefore, lysine participates in the formation of the P1 helix through stacking and tertiary contacts involved in the stabilization of the junctional region adjacent to the P1 helix. The projection of the lysine-induced in-line cleavage reductions on the structure prompts the speculation that the main structural changes associated with the regulatory mechanism involve the stabilization of G80 and the formation of the G12-C79 and G114•G163 junctional base

pairs, followed by the stabilization of the surrounding regions and the P1 helix (Fig. 10a). Stabilization of the P1 helix by the bound lysine prevents alternative base pairing of the 3' segment of the helix (24 nts downstream of G160) with the 191–215 segment and precludes formation of the anti-terminator (Fig. 4a, left panel). Instead, stabilization of the P1 helix facilitates the formation of the stem-loop structure that causes premature termination of transcription in the riboswitch upstream of the *T. maritima asd* gene (Fig. 4a, right panel). In some other lysine riboswitches, the stem-loop structure prevents initiation of translation (Fig. 4b). Interestingly, the anti-terminator of the *T. maritima asd* riboswitch contains a very long helix composed of large number of base pairs (24 base pairs with two bulged out nucleotides) so that it can efficiently compete with the formation of the P1 helix at high temperature environment.

Based on the glycine-specific in-line cleavage reductions [29, 88] and by analogy with lysine and purine riboswitches, we anticipate that the stabilization/formation of the regulatory helix P1 in domain II of glycine riboswitches could be induced by glycine binding to the bottom region of the three-way junction. Conformational adjustments might involve cooperative tertiary contacts with domain I; however, in the absence of the riboswitch structure, the molecular details of these interactions cannot be adequately predicted at this time.

The data available to date do not allow a reliable prediction of whether lysine and glycine riboswitches are kinetic or thermodynamic systems. The lysine riboswitch binds lysine with an apparent K_D in the micromolar range but requires millimolar lysine concentration to undergo transcription termination *in vitro* [27, 29]. The fact that a much higher lysine concentration is needed to fold the RNA during transcription than to reach equilibrium between the bound and unbound states indicates that *in vivo* the lysine riboswitch may operate under certain kinetic constraints, such as timing of RNA folding, ligand association, and speed of transcription [105, 110, 142–144]. It is interesting to consider whether the kinetic mechanism tuned to the rather high millimolar concentration of lysine also demands a high concentration of the counter-ion for the ligand binding. The requirement for a highly abundant cation for the riboswitch mechanism could explain the involvement of K^+ cations, which are typically considered present in cells at a higher concentration than Mg^{2+} cations.

Unlike the lysine riboswitch, the glycine riboswitch appears to require similar glycine concentrations for both binding and the modulation of transcription activation, suggesting that the thermodynamic component might play a role in riboswitch function [28]. Nevertheless, like the lysine riboswitch, the glycine riboswitch forms a complex three-dimensional structure, which may not have sufficient time to equilibrate with its surroundings at low ligand concentrations.

The schematic in Fig. 4 shows the commonly accepted general scenario of the lysine and glycine riboswitch-based regulation supported by available biochemical, genetic and phylogenetic data. However, the detail mechanism of the ligand binding and the molecular features that lead to the ultimate decision on the lysine and glycine riboswitch folding pathways are not well understood and await further clarification. The formation and/or stabilization of the key helix P1 in the presence of the cognate ligand appears to be the most widespread riboswitch mechanism that contributes to the folding of the downstream gene regulatory elements. This mechanism has been corroborated by the recent single-molecule force experiments performed on the adenine riboswitch, which convincingly and accurately demonstrated formation of the P1 helix in response to the ligand [110]. The direct involvement of lysine in the stabilization of the P1 helix is apparent from the three-dimensional structure that places bound ligand on the top of the P1 helix. However, the direct contribution of glycine to the formation of the P1 helix may be problematic given the

small size of the amino acid. The tolerance of the riboswitch to some glycine analogs also suggests that glycine might not be completely encapsulated in RNA, so that the binding interface between the ligand and the riboswitch is further reduced. How small glycine manages to effectively participate in the riboswitch-based gene expression control remains one of the most interesting unsolved problems in the riboswitch field.

Lysine and glycine riboswitch-based applications

Many studies demonstrated that RNA molecules are able to fold into intricate three-dimensional structures and form binding pockets for small compounds. These properties of RNA have been explored in the search for RNA-binding drugs that target the ribosome, ribonuclease P and viral RNAs [145]. The association of identified drugs with these large RNAs is a fortuitous rather than a natural interaction and, despite the discovery of a number of important ribosome-bound antibacterials, the design of principally new compounds for these targets can be problematic due to hardly predictable binding affinities and limited selectivity [145–147]. The finding of ribosome-binding antibiotics that are effective over the long term can also suffer from the acquisition of antibiotic resistance mutations in both the RNA and protein components of this ribonucleoprotein machinery [148].

Given the direct involvement of riboswitches in the regulation of important biosynthetic, catabolic, and transport genes, it is reasonable to assume that riboswitches may serve as excellent targets for manipulations with metabolic networks in a wide range of microorganisms. Since riboswitches are naturally adept for binding to small drug-like compounds that are often easily taken from the environment, riboswitches have the exceptional potential to exercise effective control in response to external chemical cues. Riboswitches are fundamentally different from other RNA drug targets since they evolved for the purpose of selectively binding to low-molecular weight ligands. Riboswitch structures show that the complexity of metabolite recognition by a riboswitch is comparable to the features observed in metabolite-binding proteins. Recent studies demonstrated that metabolite-like analogs with anti-microbial properties can interact with TPP [149] and FMN riboswitches [118, 150, 151] and suppress the expression of downstream genes. Though it is unclear whether riboswitch-binding drugs are potent enough to inhibit bacterial propagation exclusively by riboswitch-based gene repression or if they necessitate additional targets to complete the antimicrobial effect, the finding of antibiotic-resistant mutations within TPP, FMN and lysine riboswitch sequences undoubtedly indicates the involvement of riboswitches in cellular processes related to antibiotic action [55]. Antibiotic-resistant mutations typically deregulate the expression of the riboswitch-controlled genes. In the case of the lysine riboswitch, such a loss of gene repression may boost the biosynthesis of lysine. In fact, it would be interesting to see if the discovery of the lysine riboswitch and the determination of its three-dimensional structure will impact on the design of microbial strains used for the fermentation of this commercially important amino acid with an annual production about 800,000 tons/year [152]. Unlike most lysine and other riboswitches, glycine riboswitches predominately induce the expression of the catabolic genes and there have been no reports of bacterial toxicity associated with glycine analogs. On the other hand, the glycine-induced stimulation of gene expression by the riboswitch may be exploited in the development of an inducible and possibly tunable gene expression system. Several ongoing studies are attempting to explore the biotechnological and pharmaceutical potential of amino acid-responsive riboswitches.

Riboswitches as antibiotic targets?

The emergence of antibiotic resistance to the most powerful antibiotics highlights the need for a systematic search and subsequent development of new antimicrobial compounds, preferably those that are specific for new molecular targets and those that exhibit novel

mechanisms of action. The biosynthesis of the amino acid lysine and its immediate precursor *meso*-DAP represents one of the promising cellular pathways to be exploited pharmaceutically [153, 154]. Lysine is synthesized *de novo* in bacteria and plants, but not in mammals, and it is therefore one of the essential amino acids that must be provided to mammals through a dietary source. The absence of lysine biosynthesis in mammals suggests that the specific inhibitors of the pathway may display high antibacterial activity with low toxicity to mammals. In addition to being a crucial component of cellular proteins, lysine and *meso*-DAP are vital constituents of the bacterial peptidoglycan cell wall in Gram-positive and Gram-negative bacteria, respectively. Several enzymes of the lysine biosynthetic pathway were shown or implicated to be essential for growth under normal conditions in both Gram-negative and Gram-positive bacteria [155–157]. For instance, the *asd* gene was found to be essential and broadly preserved in a diverse range of bacteria [156]. This gene, however, is controlled by the lysine riboswitch in only some bacterial species [76]. Members of the lysine riboswitch family were identified in several clinically relevant human pathogens, such as *B. anthracis*, *S. aureus*, and *V. cholerae* [55, 56], where they most often downregulate the expression of the *lysC* gene. Given the importance of lysine and intermediates of lysine biosynthesis, it was proposed that the repression of important lysine-related enzymes by riboswitches may be sufficient to inhibit microbial growth [56].

The mapping of mutations that provide resistance towards a lysine analog AEC in the riboswitch sequence can be considered the first hint that points to the lysine riboswitch as a potential target for lysine-like antibiotics [64, 65]. Following these initial observations, it was shown that the AEC-resistant mutants do not bind lysine and can disrupt riboswitch-based regulation, and that AEC binds to the riboswitch with only a small decrease in affinity [27, 29]. Efficient binding to the *B. subtilis* *lysC* riboswitch, the repression of gene expression, and the inhibition of *B. subtilis* growth in the chemically defined minimal medium was later demonstrated for three more lysine analogs: L-3-[(2-aminoethyl)-sulfonyl]-alanine, L-4-oxalysine and DL-*trans*-2,6-diamino-4-hexenoic acid [56] (Fig. 6a). One of the best riboswitch binders, homoarginine, supported the growth of a *B. subtilis* lysine auxotroph strain and was probably consumed by bacteria to serve as a precursor for lysine production, thereby allowing the potential toxic effect on the bacterium to be circumvented [56]. The projection of the mutations that confer AEC resistance and derepress *lysC* expression onto the three-dimensional structure of the lysine riboswitch showed that the mutations either affect tertiary RNA contacts or disrupt the lysine-binding pocket [51, 52]. Several mutations of the first class prevent the base pairing within the regulatory P1 helix, thereby facilitating the formation of an alternative structure and obliterating riboswitch regulation. Other mutations probably impact on stability and the parallel alignment of the P2 and P4 stems and thereby indirectly affect lysine binding. Mutations of the second class disrupt the lysine-binding pocket and eliminate lysine-RNA contacts by substituting conserved guanosines with other nucleotides. Collectively, the biochemical and structural data argue for the concept that the antibacterial effect of the lysine analogs may, at least in part, be explained by the lysine riboswitch-mediated repression of the *lysC* gene [29, 56]. In rich medium, the growth of neither *B. subtilis* nor *B. anthracis* was inhibited, thereby indicating that the extracellular supply of lysine overcomes the full repression of riboswitch-controlled genes within the host environment [56]. Similar predictions may also be put forward for some other pathogenic bacteria since single genes placed under the control of the lysine riboswitch do not appear to be essential due to the metabolic plasticity of bacteria, namely the presence of several enzymes with a similar function and the ability to utilize critical nutrients from the host environment [55]. These considerations lead to the conclusion that the compound that targets the lysine riboswitch exclusively may not be potent enough to stop the growth of certain pathogenic microbial species [56].

Earlier studies demonstrated the toxic effect of AEC on yeast [158] and *Corinebacterium glutamicum* [159], organisms which, as recently demonstrated, do not possess lysine riboswitches [160]. AEC tolerance was correlated with the activities of lysyl-tRNA synthetase and aspartokinase, and alternative mechanisms of AEC toxicity, such as the incorporation of AEC into proteins, were suggested (e.g., [161–165]). A recent report has attributed the mechanism of AEC resistance to point mutations in the lysyl-tRNA synthetase encoded by the *lysS* gene [54]. The mutant synthetases demonstrated ~50-fold decrease in their binding affinities for AEC and conferred antibiotic resistance even in the presence of the wild type lysine riboswitch. These data support the view that the lysyl-tRNA synthetase can efficiently incorporate both lysine and AEC into a protein chain and that AEC-containing proteins cause toxicity to microbial cells. Antibiotic resistant mutations in the lysine riboswitch elevate the synthesis of the enzymes involved in lysine biosynthesis and increase the production of lysine. Lysine at higher concentrations outcompetes AEC for binding with lysyl-tRNA synthetase and reduces the toxic effect of AEC. Therefore, the future design of effective lysine-like antibiotics should possibly aim at targeting both riboswitches and lysyl-tRNA synthetases. The utility of the lysine riboswitches as drug targets is, however, not assured, given that some bacteria have pathways for the biosynthesis and acquisition of lysine from the environment which are not controlled by riboswitches. In addition, the designed drugs must retain high selectivity toward their microbial targets but avoid interference with the functions of mammalian enzymes in order to reduce the off-target toxic effects.

Amino acid-inducible expression system

Gene expression approaches based on strong and tightly regulated promoters play an important role in many applications that require elevated protein production. Among various induction systems constructed for *B. subtilis*, the IPTG- and xylose-inducible systems were found to be the most convenient and are broadly used in academic research [151]. The cost of these inducers, however, restricts their usage in industry; therefore, the search for an inexpensive alternative turned to the glycine riboswitch, capable of co-transcriptional gene activation in response to a low glycine concentration. The experiments conducted with the *B. subtilis gcv* glycine riboswitch demonstrated the feasibility to use the riboswitch as an inducible element of the gene expression system [166]. Under the control of the glycine riboswitch, the production of the recombinant β -galactosidase had a 6-fold induction factor and reached a level comparable to the yields in xylose- and IPTG-driven expression systems. Gene overexpression can be further manipulated by combining the riboswitch with natural and artificial promoters of different strengths. However, the increase in glycine-induced protein production driven by a stronger promoter correlated with a higher basal level of the recombinant protein, necessitating additional adaptations in the system to reduce the basal level.

Regulation of amino acid metabolism at the crossroads of cellular processes: conclusions and perspectives

The availability of amino acids is vital to all living entities. It is not surprising that the anabolism, catabolism, transport, homeostasis and other processes that involve amino acids play a major role in the well-being of organisms with different complexities, ranging from microbes to humans [167, 168]. In bacteria, in addition to their fundamental role as building blocks for protein molecules, glycine and lysine participate in a number of other physiologically significant cellular processes. For instance, as mentioned earlier, glycine can be used as a source of carbon and energy [28, 85], while lysine and its derivatives were implicated in different stress responses, resistance to antibiotics, and pathogenic virulence (reviewed in [169]). The discovery of the glycine and lysine riboswitches emphasizes that

deviations from the common mechanisms of regulation described for the biosynthesis of related compounds can and should be expected. Though it is still puzzling why only glycine and lysine metabolism retained riboswitch control in addition to protein-driven, T-box, and attenuation mechanisms of gene expression regulation, both lysine and glycine riboswitches make important contributions to the sustenance of adequate levels of the cognate amino acids in response to internal and external cues. The surprising discovery of the narrowly distributed 2'-deoxyguanosine-specific riboswitch [22] that only slightly differs from the adenine and guanine riboswitches [23, 24] implies that the existence of other perhaps less abundant amino acid-responsive riboswitches cannot be completely ruled out. A recent report pointed to the leader region of a fungal gene as a possible candidate for riboswitch activity in response to arginine [33].

The molecular characterization of lysine and glycine riboswitches by biochemical techniques revealed no similarities in the sequences, secondary structures, and ligand-induced formation of tertiary interactions between riboswitches that recognize compounds from the same class of macromolecules. Not surprisingly, riboswitches specific to unrelated metabolites also do not resemble the amino acid-responsive riboswitches. The three-dimensional structure of the lysine riboswitch further entertains dissimilarities with other riboswitch folds and demonstrates a unique five-way junction architecture stabilized through conserved peripheral interactions [51, 52]. Based on secondary structure and footprinting data, we foresee that the three-dimensional structure of the glycine riboswitch will be entirely different from the lysine riboswitch, although some structural features, such as the emplacement of the bound ligand between two helical regions and the opening of the binding pocket from the carboxylate end might be preserved in both riboswitches. The structure determination of the tandem glycine riboswitch is challenging, given the anticipated flexibility of the linker region and possible cooperative interdomain contacts. Hopefully, the more feasible determination of the structure of an individual glycine-binding domain combined with the NAIM and mutagenesis data could point to the nucleotides involved in riboswitch cooperativity and dissect the mechanism of the tandem riboswitch action.

The practical applications of riboswitches entirely depend on the ability of riboswitches to bind small-molecule ligands. Natural riboswitches, often superior in sensitivity and specificity to *in vitro*-selected RNA aptamers, greatly enlarged the collection of available RNA domains that sense small molecules. The assortment of riboswitches may be further expanded by complementary changes in the RNA sequences and their respective ligands. In the event that the altered ligand sufficiently diverges from the natural metabolite, the modified riboswitches may, in theory, be excluded from the circuits controlled by the 'parental' cellular metabolite, thereby creating a precise post-transcriptional tool for the induction or repression of the gene of interest. Though the modification of riboswitches for the design of novel specificities is a subject of future research, a recent study has already reported a successful blend of the natural TPP-sensing sensor and the hammerhead ribozyme to build TPP-responsive molecular scissors [170]. Amino acids are cheaper, better soluble, and superior in stability as compared to other riboswitch ligands and, therefore, are well-suited for practical applications. Moreover, the glycine riboswitch is a natural gene expression activator and, given the cooperative mechanism of glycine riboswitch functioning, may be tuned to respond to various dynamic ranges of the inductor. Nevertheless, since free amino acids are highly abundant in all organisms, the potential biotechnological employment of the natural lysine and glycine riboswitches as parts of 'artificial' expression systems might only be of limited use. However, it should not be entirely surprising to discover that some of the bacterial strains used for the production of lysine for fifty years did contain deregulatory mutations in lysine riboswitches [152], though the current producer of choice, bacterium *C. glutamicum*, appears to be devoid of lysine

riboswitches [160]. Despite the vital importance of lysine biosynthesis for bacteria and the direct interaction of lysine riboswitches with antibacterial compounds, the pharmacological targeting of the lysine riboswitch and the design of lysine-like antibiotics is not trivial and cannot be assured due to the control of a limited number of biosynthetic genes in the medically relevant microbial strains and the metabolic plasticity of many pathogenic species. We believe that the utilization of the structural information complemented by biochemical data will allow for a more complete view of regulation mechanisms employed by amino acid-responsive riboswitches and will facilitate their modification for research, clinical and industrial applications.

Acknowledgments

We thank Dr. J. Barrick for providing riboswitch distribution data. This work was supported by National Institutes of Health grant GM073618. Lily Huang contributed to the structure determination of the lysine riboswitch in our laboratory.

References

1. Grundy FJ, Henkin TM. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell*. 1993; 74:475–482. [PubMed: 8348614]
2. Yanofsky C. Attenuation in the control of expression of bacterial operons. *Nature*. 1981; 289:751–758. [PubMed: 7007895]
3. Gutierrez-Preciado A, Henkin TM, Grundy FJ, Yanofsky C, Merino E. Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol Mol Biol Rev*. 2009; 73:36–61. [PubMed: 19258532]
4. Vitreschak AG, Mironov AA, Lyubetsky VA, Gelfand MS. Comparative genomic analysis of T-box regulatory systems in bacteria. *RNA*. 2008; 14:717–735. [PubMed: 18359782]
5. Babitzke P, Gollnick P. Posttranscription initiation control of tryptophan metabolism in *Bacillus subtilis* by the trp RNA-binding attenuation protein (TRAP), anti-TRAP, and RNA structure. *J Bacteriol*. 2001; 183:5795–5802. [PubMed: 11566976]
6. Kovaleva GY, Gelfand MS. Transcriptional regulation of the methionine and cysteine transport and metabolism in streptococci. *FEMS Microbiol Lett*. 2007; 276:207–215. [PubMed: 17956428]
7. Henkin TM, Yanofsky C. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays*. 2002; 24:700–707. [PubMed: 12210530]
8. Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, Perumov DA, Nudler E. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell*. 2002; 111:747–756. [PubMed: 12464185]
9. Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR. Genetic control by a metabolite binding mRNA. *Chem Biol*. 2002; 9:1043–1049. [PubMed: 12323379]
10. Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*. 2002; 419:952–956. [PubMed: 12410317]
11. Breaker, RR. Riboswitches and the RNA World. In: Gesteland, RF.; Cech, TR.; Atkins, JF., editors. *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor; 2006. p. 89–108.
12. Barrick JE, Breaker RR. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol*. 2007; 8:R239. [PubMed: 17997835]
13. Fuchs RT, Grundy FJ, Henkin TM. The S_{MK} box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat Struct Mol Biol*. 2006; 13:226–233. [PubMed: 16491091]
14. Bocobza SE, Aharoni A. Switching the light on plant riboswitches. *Trends Plant Sci*. 2008; 13:526–533. [PubMed: 18778966]
15. Corbino KA, Barrick JE, Lim J, Welz R, Tucker BJ, Puskarz I, Mandal M, Rudnick ND, Breaker RR. Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol*. 2005; 6:R70. [PubMed: 16086852]

16. Epshtein V, Mironov AS, Nudler E. The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. U S A.* 2003; 100:5052–5056. [PubMed: 12702767]
17. McDaniel BA, Grundy FJ, Artsimovitch I, Henkin TM. Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. *Proc. Natl. Acad. Sci. U S A.* 2003; 100:3083–3088. [PubMed: 12626738]
18. Winkler WC, Cohen-Chalamish S, Breaker RR. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. U S A.* 2002; 99:15908–15913. [PubMed: 12456892]
19. Winkler WC, Nahvi A, Sudarsan N, Barrick JE, Breaker RR. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol.* 2003; 10:701–707. [PubMed: 12910260]
20. Regulski EE, Moy RH, Weinberg Z, Barrick JE, Yao Z, Ruzzo WL, Breaker RR. A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Mol Microbiol.* 2008; 68:918–932. [PubMed: 18363797]
21. Wang JX, Lee ER, Morales DR, Lim J, Breaker RR. Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling. *Mol Cell.* 2008; 29:691–702. [PubMed: 18374645]
22. Kim JN, Roth A, Breaker RR. Guanine riboswitch variants from *Mesoplasma florum* selectively recognize 2'-deoxyguanosine. *Proc Natl Acad Sci U S A.* 2007; 104:16092–16097. [PubMed: 17911257]
23. Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell.* 2003; 113:577–586. [PubMed: 12787499]
24. Mandal M, Breaker RR. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat Struct Mol Biol.* 2004; 11:29–35. [PubMed: 14718920]
25. Roth A, Winkler WC, Regulski EE, Lee BW, Lim J, Jona I, Barrick JE, Ritwik A, Kim JN, Welz R, Iwata-Reuyl D, Breaker RR. A riboswitch selective for the queuosine precursor preQ₁ contains an unusually small aptamer domain. *Nat Struct Mol Biol.* 2007; 14:308–317. [PubMed: 17384645]
26. Sudarsan N, Lee ER, Weinberg Z, Moy RH, Kim JN, Link KH, Breaker RR. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science.* 2008; 321:411–413. [PubMed: 18635805]
27. Grundy FJ, Lehman SC, Henkin TM. The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc Natl Acad Sci U S A.* 2003; 100:12057–12062. [PubMed: 14523230]
28. Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL, Breaker RR. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science.* 2004; 306:275–279. [PubMed: 15472076]
29. Sudarsan N, Wickiser JK, Nakamura S, Ebert MS, Breaker RR. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.* 2003; 17:2688–2697. [PubMed: 14597663]
30. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature.* 2004; 428:281–286. [PubMed: 15029187]
31. Cromie MJ, Shi Y, Latifi T, Groisman EA. An RNA sensor for intracellular Mg²⁺. *Cell.* 2006; 125:71–84. [PubMed: 16615891]
32. Dann CE 3rd, Wakeman CA, Sieling CL, Baker SC, Irnov I, Winkler WC. Structure and mechanism of a metal-sensing regulatory RNA. *Cell.* 2007; 130:878–892. [PubMed: 17803910]
33. Borsuk P, Przykorska A, Blachnio K, Koper M, Pawlowicz JM, Pekala M, Weglenski P. L-arginine influences the structure and function of arginase mRNA in *Aspergillus nidulans*. *Biological chemistry.* 2007; 388:135–144. [PubMed: 17261076]
34. Blouin S, Mulhbach J, Penedo JC, Lafontaine DA. Riboswitches: ancient and promising genetic regulators. *Chembiochem.* 2009; 10:400–416. [PubMed: 19101979]
35. Coppins RL, Hall KB, Groisman EA. The intricate world of riboswitches. *Curr Opin Microbiol.* 2007; 10:176–181. [PubMed: 17383225]
36. Dambach MD, Winkler WC. Expanding roles for metabolite-sensing regulatory RNAs. *Curr Opin Microbiol.* 2009; 12:1–9. [PubMed: 19174324]

37. Gelfand MS. Bacterial *cis*-regulatory RNA structures. *Molecular Biology*. 2006; 40:541–550.
38. Henkin TM. Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev*. 2008; 22:3383–3390. [PubMed: 19141470]
39. Roth A, Breaker RR. The structural and functional diversity of metabolite-binding riboswitches. *Annu Rev Biochem*. 2009
40. Serganov A, Patel DJ. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet*. 2007; 8:776–790. [PubMed: 17846637]
41. Narberhaus F, Waldminghaus T, Chowdhury S. RNA thermometers. *FEMS Microbiol Rev*. 2006; 30:3–16. [PubMed: 16438677]
42. Batey RT. Structures of regulatory elements in mRNAs. *Curr Opin Struct Biol*. 2006; 16:299–306. [PubMed: 16707260]
43. Edwards TE, Klein DJ, Ferre-D'Amare AR. Riboswitches: small-molecule recognition by gene regulatory RNAs. *Curr Opin Struct Biol*. 2007; 17:273–279. [PubMed: 17574837]
44. Montange RK, Batey RT. Riboswitches: emerging themes in RNA structure and function. *Annu Rev Biophys*. 2008; 37:117–133. [PubMed: 18573075]
45. Schwalbe H, Buck J, Furtig B, Noeske J, Wohnert J. Structures of RNA switches: insight into molecular recognition and tertiary structure. *Angew Chem Int Ed Engl*. 2007; 46:1212–1219. [PubMed: 17226886]
46. Serganov A. The long and the short of riboswitches. *Curr Opin Struct Biol*. 2009; 19:251–259. [PubMed: 19303767]
47. Serganov A, Patel DJ. Towards deciphering the principles underlying an mRNA recognition code. *Curr Opin Struct Biol*. 2008; 18:120–129. [PubMed: 18255277]
48. Wakeman CA, Winkler WC, Dann CE 3rd. Structural features of metabolite-sensing riboswitches. *Trends Biochem Sci*. 2007; 32:415–424. [PubMed: 17764952]
49. Kim JN, Breaker RR. Purine sensing by riboswitches. *Biol Cell*. 2008; 100:1–11. [PubMed: 18072940]
50. Wang JX, Breaker RR. Riboswitches that sense S-adenosylmethionine and S-adenosylhomocysteine. *Biochem Cell Biol*. 2008; 86:157–168. [PubMed: 18443629]
51. Garst AD, Heroux A, Rambo RP, Batey RT. Crystal structure of the lysine riboswitch regulatory mRNA element. *J Biol Chem*. 2008; 283:22347–22351. [PubMed: 18593706]
52. Serganov A, Huang L. Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature*. 2008; 455:1263–1267. [PubMed: 18784651]
53. Lipfert J, Das R, Chu VB, Kudaravalli M, Boyd N, Herschlag D, Doniach S. Structural transitions and thermodynamics of a glycine-dependent riboswitch from *Vibrio cholerae*. *J Mol Biol*. 2007; 365:1393–1406. [PubMed: 17118400]
54. Ataide SF, Wilson SN, Dang S, Rogers TE, Roy B, Banerjee R, Henkin TM, Ibba M. Mechanisms of resistance to an amino acid antibiotic that targets translation. *ACS Chem Biol*. 2007; 2:819–827. [PubMed: 18154269]
55. Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. *Nature biotechnology*. 2006; 24:1558–1564.
56. Blount KF, Wang JX, Lim J, Sudarsan N, Breaker RR. Antibacterial lysine analogs that target lysine riboswitches. *Nature chemical biology*. 2007; 3:44–49.
57. Belytsky, BR. *Bacillus subtilis* and Its Closest Relatives: From Genes to Cells. Sonenshein, AL.; Hoch, JA.; Losick, R., editors. Washington, DC: Am. Soc. Microbiol.; 2002. p. 203–231.
58. Forman M, Aronson A. Regulation of dipicolinic acid biosynthesis in sporulating *Bacillus cereus*. Characterization of enzymic changes and analysis of mutants. *Biochem J*. 1972; 126:503–513. [PubMed: 4627586]
59. Boy E, Borne F, Patte JC. Isolation and identification of mutants constitutive for aspartokinase III synthesis in *Escherichia coli* K 12. *Biochimie*. 1979; 61:1151–1160. [PubMed: 231461]
60. Patte JC, Akrim M, Mejean V. The leader sequence of the *Escherichia coli* *lysC* gene is involved in the regulation of LysC synthesis. *FEMS Microbiol Lett*. 1998; 169:165–170. [PubMed: 9851048]
61. Liao HH, Hseu TH. Analysis of the regulatory region of the *lysC* gene of *Escherichia coli*. *FEMS Microbiol Lett*. 1998; 168:31–36. [PubMed: 9812360]

62. Mader U, Homuth G, Scharf C, Buttner K, Bode R, Hecker M. Transcriptome and proteome analysis of *Bacillus subtilis* gene expression modulated by amino acid availability. *J Bacteriol.* 2002; 184:4288–4295. [PubMed: 12107147]
63. Vold B, Szulmajster J, Carbone A. Regulation of dihydrodipicolinate synthase and aspartate kinase in *Bacillus subtilis*. *J Bacteriol.* 1975; 121:970–974. [PubMed: 163819]
64. Lu Y, Chen NY, Paulus H. Identification of *aecA* mutations in *Bacillus subtilis* as nucleotide substitutions in the untranslated leader region of the aspartokinase II operon. *J Gen Microbiol.* 1991; 137:1135–1143. [PubMed: 1907638]
65. Lu Y, Shevtchenko TN, Paulus H. Fine-structure mapping of *cis*-acting control sites in the *lysC* operon of *Bacillus subtilis*. *FEMS Microbiol Lett.* 1992; 92:23–27. [PubMed: 1624109]
66. Kochhar S, Paulus H. Lysine-induced premature transcription termination in the *lysC* operon of *Bacillus subtilis*. *Microbiology.* 1996; 142(Pt 7):1635–1639. [PubMed: 8757727]
67. Chen NY, Hu FM, Paulus H. Nucleotide sequence of the overlapping genes for the subunits of *Bacillus subtilis* aspartokinase II and their control regions. *J Biol Chem.* 1987; 262:8787–8798. [PubMed: 3036830]
68. Stragier P, Richaud F, Borne F, Patte JC. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. I. Identification of a *lysR* gene encoding an activator of the *lysA* gene. *J Mol Biol.* 1983; 168:307–320. [PubMed: 6411928]
69. Stragier P, Danos O, Patte JC. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. II. Nucleotide sequence of the *lysA* gene and its regulatory region. *J Mol Biol.* 1983; 168:321–331. [PubMed: 6350601]
70. Schendel FJ, Flickinger MC. Cloning and nucleotide sequence of the gene coding for aspartokinase II from a thermophilic methylotrophic *Bacillus* sp. *Appl Environ Microbiol.* 1992; 58:2806–2814. [PubMed: 1444390]
71. Rosner A. Control of lysine biosynthesis in *Bacillus subtilis*: inhibition of diaminopimelate decarboxylase by lysine. *J Bacteriol.* 1975; 121:20–28. [PubMed: 234936]
72. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J Biol Chem.* 2002; 277:48949–48959. [PubMed: 12376536]
73. Vitreschak AG, Rodionov DA, Mironov AA, Gelfand MS. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* 2002; 30:3141–3151. [PubMed: 12136096]
74. Vitreschak AG, Rodionov DA, Mironov AA, Gelfand MS. Regulation of the vitamin B₁₂ metabolism and transport in bacteria by a conserved RNA structural element. *RNA.* 2003; 9:1084–1097. [PubMed: 12923257]
75. Sudarsan N, Barrick JE, Breaker RR. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA.* 2003; 9:644–647. [PubMed: 12756322]
76. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res.* 2003; 31:6748–6757. [PubMed: 14627808]
77. Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, Lee M, Roth A, Sudarsan N, Jona I, Wickiser JK, Breaker RR. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. U S A.* 2004; 101:6421–6426. [PubMed: 15096624]
78. Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, Bork P, Hugenholtz P, Rubin EM. Comparative metagenomics of microbial communities. *Science.* 2005; 308:554–557. [PubMed: 15845853]
79. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO. Environmental genome shotgun sequencing of the Sargasso Sea. *Science.* 2004; 304:66–74. [PubMed: 15001713]
80. Kazanov MD, Vitreschak AG, Gelfand MS. Abundance and functional diversity of riboswitches in microbial communities. *BMC Genomics.* 2007; 8:347. [PubMed: 17908319]

81. Morris RM, Rappe MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*. 2002; 420:806–810. [PubMed: 12490947]
82. Ito M, Guffanti AA, Zemsky J, Ivey DM, Krulwich TA. Role of the *nhaC*-encoded Na⁺/H⁺ antiporter of alkaliphilic *Bacillus firmus* OF4. *J Bacteriol*. 1997; 179:3851–3857. [PubMed: 9190799]
83. Krulwich TA, Ito M, Guffanti AA. The Na⁺-dependence of alkaliphily in *Bacillus*. *Biochim Biophys Acta*. 2001; 1505:158–168. [PubMed: 11248197]
84. Chang CH, Frey PA. Cloning, sequencing, heterologous expression, purification, and characterization of adenosylcobalamin-dependent D-lysine 5, 6-aminomutase from *Clostridium sticklandii*. *J Biol Chem*. 2000; 275:106–114. [PubMed: 10617592]
85. Tripp HJ, Schwalbach MS, Meyer MM, Kitner JB, Breaker RR, Giovannoni SJ. Unique glycine-activated riboswitch linked to glycine-serine auxotrophy in SAR11. *Environ Microbiol*. 2009; 11:230–238. [PubMed: 19125817]
86. Sudarsan N, Hammond MC, Block KF, Welz R, Barrick JE, Roth A, Breaker RR. Tandem riboswitch architectures exhibit complex gene control functions. *Science*. 2006; 314:300–304. [PubMed: 17038623]
87. Welz R, Breaker RR. Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA*. 2007; 13:573–582. [PubMed: 17307816]
88. Kwon M, Strobel SA. Chemical basis of glycine riboswitch cooperativity. *RNA*. 2008; 14:25–34. [PubMed: 18042658]
89. Edelstein SJ. Cooperative interactions of hemoglobin. *Annu Rev Biochem*. 1975; 44:209–232. [PubMed: 237460]
90. Royer WE Jr, Knapp JE, Strand K, Heaslet HA. Cooperative hemoglobins: conserved fold, diverse quaternary assemblies and allosteric mechanisms. *Trends Biochem Sci*. 2001; 26:297–304. [PubMed: 11343922]
91. Correll CC, Freeborn B, Moore PB, Steitz TA. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell*. 1997; 91:705–712. [PubMed: 9393863]
92. Leontis NB, Westhof E. The 5S rRNA loop E: chemical probing and phylogenetic data versus crystal structure. *RNA*. 1998; 4:1134–1153. [PubMed: 9740131]
93. Klein DJ, Schmeing TM, Moore PB, Steitz TA. The kink-turn: a new RNA secondary structure motif. *EMBO J*. 2001; 20:4214–4221. [PubMed: 11483524]
94. Blouin S, Lafontaine DA. A loop loop interaction and a K-turn motif located in the lysine aptamer domain are important for the riboswitch gene regulation control. *RNA*. 2007; 13:1256–1267. [PubMed: 17585050]
95. Nissen P, Ippolito JA, Ban N, Moore PB, Steitz TA. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. *Proc Natl Acad Sci U S A*. 2001; 98:4899–4903. [PubMed: 11296253]
96. Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, Wilkinson AC, Finn RD, Griffiths-Jones S, Eddy SR, Bateman A. Rfam: updates to the RNA families database. *Nucleic Acids Res*. 2009; 37:D136–D140. [PubMed: 18953034]
97. Chevalier C, Geissmann T, Helfer AC, Romby P. Probing mRNA structure and sRNA-mRNA interactions in bacteria using enzymes and lead(II). *Methods Mol Biol*. 2009; 540:215–232. [PubMed: 19381563]
98. Chiaruttini C, Allem F, Springer M. Structural probing of RNA thermosensors. *Methods Mol Biol*. 2009; 540:233–245. [PubMed: 19381564]
99. Soukup GA, Breaker RR. Relationship between internucleotide linkage geometry and the stability of RNA. *RNA*. 1999; 5:1308–1325. [PubMed: 10573122]
100. Wakeman CA, Winkler WC. Analysis of the RNA backbone: structural analysis of riboswitches by in-line probing and selective 2'-hydroxyl acylation and primer extension. *Methods Mol Biol*. 2009; 540:173–191. [PubMed: 19381560]
101. Wilkinson KA, Merino EJ, Weeks KM. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat Protoc*. 2006; 1:1610–1616. [PubMed: 17406453]

102. Tullius TD, Greenbaum JA. Mapping nucleic acid structure by hydroxyl radical cleavage. *Curr Opin Chem Biol.* 2005; 9:127–134. [PubMed: 15811796]
103. Ryder SP, Strobel SA. Nucleotide analog interference mapping. *Methods.* 1999; 18:38–50. [PubMed: 10208815]
104. Soukup JK, Soukup GA. Identification of metabolite-riboswitch interactions using nucleotide analog interference mapping and suppression. *Methods Mol Biol.* 2009; 540:193–206. [PubMed: 19381561]
105. Lang K, Rieder R, Micura R. Ligand-induced folding of the *thiM*TPP riboswitch investigated by a structure-based fluorescence spectroscopic approach. *Nucleic Acids Res.* 2007; 35:5370–5378. [PubMed: 17693433]
106. Lemay JF, Penedo JC, Tremblay R, Lilley DM, Lafontaine DA. Folding of the adenine riboswitch. *Chem Biol.* 2006; 13:857–868. [PubMed: 16931335]
107. Rieder R, Lang K, Graber D, Micura R. Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control. *Chembiochem.* 2007; 8:896–902. [PubMed: 17440909]
108. Buck J, Furtig B, Noeske J, Wohnert J, Schwalbe H. Time-resolved NMR methods resolving ligand-induced RNA folding at atomic resolution. *Proc Natl Acad Sci U S A.* 2007; 104:15699–15704. [PubMed: 17895388]
109. Noeske J, Buck J, Furtig B, Nasiri HR, Schwalbe H, Wohnert J. Interplay of 'induced fit' and preorganization in the ligand induced folding of the aptamer domain of the guanine binding riboswitch. *Nucleic Acids Res.* 2007; 35:572–583. [PubMed: 17175531]
110. Greenleaf WJ, Frieda KL, Foster DA, Woodside MT, Block SM. Direct observation of hierarchical folding in single riboswitch aptamers. *Science.* 2008; 319:630–633. [PubMed: 18174398]
111. Batey RT, Gilbert SD, Montange RK. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature.* 2004; 432:411–415. [PubMed: 15549109]
112. Serganov A, Yuan YR, Pikovskaya O, Polonskaia A, Malinina L, Phan AT, Hobartner C, Micura R, Breaker RR, Patel DJ. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* 2004; 11:1729–1741. [PubMed: 15610857]
113. Lipfert J, Herschlag D, Doniach S. Riboswitch conformations revealed by small-angle X-ray scattering. *Methods Mol Biol.* 2009; 540:141–159. [PubMed: 19381558]
114. Edwards TE, Ferre-D'Amare AR. Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure.* 2006; 14:1459–1468. [PubMed: 16962976]
115. Serganov A, Polonskaia A, Phan AT, Breaker RR, Patel DJ. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature.* 2006; 441:1167–1171. [PubMed: 16728979]
116. Thore S, Leibundgut M, Ban N. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science.* 2006; 312:1208–1211. [PubMed: 16675665]
117. Yamauchi T, Miyoshi D, Kubodera T, Nishimura A, Nakai S, Sugimoto N. Roles of Mg²⁺ in TPP-dependent riboswitch. *FEBS Lett.* 2005; 579:2583–2588. [PubMed: 15862294]
118. Serganov A, Huang L, Patel DJ. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature.* 2009; 458:233–237. [PubMed: 19169240]
119. Noeske J, Schwalbe H, Wohnert J. Metal-ion binding and metal-ion induced folding of the adenine-sensing riboswitch aptamer domain. *Nucleic Acids Res.* 2007; 35:5262–5273. [PubMed: 17686787]
120. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science.* 1996; 273:1678–1685. [PubMed: 8781224]
121. Ennifar E, Walter P, Ehresmann B, Ehresmann C, Dumas P. Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat Struct Biol.* 2001; 8:1064–1068. [PubMed: 11702070]
122. Woodson SA. Metal ions and RNA folding: a highly charged topic with a dynamic future. *Curr Opin Chem Biol.* 2005; 9:104–109. [PubMed: 15811793]

123. Cochrane JC, Lipchock SV, Strobel SA. Structural investigation of the GlmS ribozyme bound to its catalytic cofactor. *Chem. Biol.* 2007; 14:97–105. [PubMed: 17196404]
124. Klein DJ, Ferre-D'Amare AR. Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science.* 2006; 313:1752–1756. [PubMed: 16990543]
125. Xiao H, Edwards TE, Ferre-D'Amare AR. Structural basis for specific, high-affinity tetracycline binding by an in vitro evolved aptamer and artificial riboswitch. *Chem Biol.* 2008; 15:1125–1137. [PubMed: 18940672]
126. Klein DJ, Edwards TE, Ferre-D'Amare AR. Cocystal structure of a class I preQ1 riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nat Struct Mol Biol.* 2009; 16:343–344. [PubMed: 19234468]
127. Kang M, Peterson R, Feigon J. Structural Insights into riboswitch control of the biosynthesis of queuosine, a modified nucleotide found in the anticodon of tRNA. *Mol Cell.* 2009; 33:784–790. [PubMed: 19285444]
128. Spitale RC, Torelli AT, Krucinska J, Bandarian V, Wedekind JE. The structural basis for recognition of the PreQ0 metabolite by an unusually small riboswitch aptamer domain. *J Biol Chem.* 2009; 284:11012–11016. [PubMed: 19261617]
129. Rieder U, Lang K, Kreutz C, Polacek N, Micura R. Evidence for pseudoknot formation of class I preQ1 riboswitch aptamers. *Chembiochem.* 2009; 10:1141–1144. [PubMed: 19382115]
130. Adams PL, Stahley MR, Kosek AB, Wang J, Strobel SA. Crystal structure of a self-splicing group I intron with both exons. *Nature.* 2004; 430:45–50. [PubMed: 15175762]
131. Basu S, Rambo RP, Strauss-Soukup J, Cate JH, Ferre-D'Amare AR, Strobel SA, Doudna JA. A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nat Struct Biol.* 1998; 5:986–992. [PubMed: 9808044]
132. Batey RT, Doudna JA. Structural and energetic analysis of metal ions essential to SRP signal recognition domain assembly. *Biochemistry.* 2002; 41:11703–11710. [PubMed: 12269812]
133. Serganov A, Polonskaia A, Ehresmann B, Ehresmann C, Patel DJ. Ribosomal protein S15 represses its own translation via adaptation of an rRNA-like fold within its mRNA. *EMBO J.* 2003; 22:1898–1908. [PubMed: 12682022]
134. Furia, TE. Sequestrants in foods. In: Furia, TE., editor. *CRC Handbook of food additives*. Vol. vol. I. Palo Alto: CRC Press; 1972. p. 271-296.
135. Feig, AL.; Uhlenbeck, OC. The role of metal ions in RNA biochemistry. In: Gesteland, RF.; Cech, TR.; Atkins, JF., editors. *The RNA World Second Edition*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor; 1999. p. 287-320.
136. Wang W, Guffanti AA, Wei Y, Ito M, Krulwich TA. Two types of *Bacillus subtilis* tetA(L) deletion strains reveal the physiological importance of TetA(L) in K⁺ acquisition as well as in Na⁺, alkali, and tetracycline resistance. *J Bacteriol.* 2000; 182:2088–2095. [PubMed: 10735849]
137. Dinnbier U, Limpinsel E, Schmid R, Bakker EP. Transient accumulation of potassium glutamate and its replacement by trehalose during adaptation of growing cells of *Escherichia coli* K-12 to elevated sodium chloride concentrations. *Arch Microbiol.* 1988; 150:348–357. [PubMed: 3060036]
138. Whatmore AM, Reed RH. Determination of turgor pressure in *Bacillus subtilis*: a possible role for K⁺ in turgor regulation. *J Gen Microbiol.* 1990; 136:2521–2526. [PubMed: 2127801]
139. Outten CE, O'Halloran TV. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science.* 2001; 292:2488–2492. [PubMed: 11397910]
140. Moncany ML, Kellenberger E. High magnesium content of *Escherichia coli* B. *Experientia.* 1981; 37:846–847. [PubMed: 7026272]
141. Lu C, Smith AM, Fuchs RT, Ding F, Rajashankar K, Henkin TM, Ke A. Crystal structures of the SAM-III/S_{MK} riboswitch reveal the SAM-dependent translation inhibition mechanism. *Nat Struct Mol Biol.* 2008; 15:1076–1083. [PubMed: 18806797]
142. Breaker RR. Complex riboswitches. *Science.* 2008; 319:1795–1797. [PubMed: 18369140]
143. Wickiser JK, Cheah MT, Breaker RR, Crothers DM. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry.* 2005; 44:13404–13414. [PubMed: 16201765]

144. Wickiser JK, Winkler WC, Breaker RR, Crothers DM. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell.* 2005; 18:49–60. [PubMed: 15808508]
145. Hermann T, Tor Y. RNA as a target for small-molecule therapeutics. *Expert Opin. Ther. Pat.* 2005; 15:49–62.
146. Foloppe N, Matassova N, Aboul-Ela F. Towards the discovery of drug-like RNA ligands? *Drug Discov Today.* 2006; 11:1019–1027. [PubMed: 17055412]
147. Franceschi F, Duffy EM. Structure-based drug design meets the ribosome. *Biochem Pharmacol.* 2006; 71:1016–1025. [PubMed: 16443192]
148. Yonath A. Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation. *Annu Rev Biochem.* 2005; 74:649–679. [PubMed: 16180279]
149. Sudarsan N, Cohen-Chalamish S, Nakamura S, Emilsson GM, Breaker RR. Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chem. Biol.* 2005; 12:1325–1335. [PubMed: 16356850]
150. Lee ER, Blount KF, Breaker RR. Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. *RNA Biol.* 2009; 6:187–194. [PubMed: 19246992]
151. Ott E, Stolz J, Lehmann M, Mack M. The RFN riboswitch of *Bacillus subtilis* is a target for the antibiotic roseoflavin produced by *Streptomyces davawensis*. *RNA Biol.* 2009; 6
152. Anastassiadis S. L-lysine fermentation. *Recent Pat Biotechnol.* 2007; 1:11–24. [PubMed: 19075830]
153. Hutton CA, Perugini MA, Gerrard JA. Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Mol Biosyst.* 2007; 3:458–465. [PubMed: 17579770]
154. Hutton CA, Southwood TJ, Turner JJ. Inhibitors of lysine biosynthesis as antibacterial agents. *Mini Rev Med Chem.* 2003; 3:115–127. [PubMed: 12570844]
155. Becker D, Selbach M, Rollenhagen C, Ballmaier M, Meyer TF, Mann M, Bumann D. Robust *Salmonella* metabolism limits possibilities for new antimicrobials. *Nature.* 2006; 440:303–307. [PubMed: 16541065]
156. Gerdes SY, Scholle MD, Campbell JW, Balazsi G, Ravasz E, Daugherty MD, Somera AL, Kyrpides NC, Anderson I, Gelfand MS, Bhattacharya A, Kapatral V, D'Souza M, Baev MV, Grechkin Y, Mseeh F, Fonstein MY, Overbeek R, Barabasi AL, Oltvai ZN, Osterman AL. Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J Bacteriol.* 2003; 185:5673–5684. [PubMed: 13129938]
157. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK, Arnaud M, Asai K, Ashikaga S, Aymerich S, Bessieres P, Boland F, Brignell SC, Bron S, Bunai K, Chapuis J, Christiansen LC, Danchin A, Debarbouille M, Dervyn E, Deuerling E, Devine K, Devine SK, Dreesen O, Errington J, Fillinger S, Foster SJ, Fujita Y, Galizzi A, Gardan R, Eschevins C, Fukushima T, Haga K, Harwood CR, Hecker M, Hosoya D, Hullo MF, Kakeshita H, Karamata D, Kasahara Y, Kawamura F, Koga K, Koski P, Kuwana R, Imamura D, Ishimaru M, Ishikawa S, Ishio I, Le Coq D, Masson A, Mauel C, Meima R, Mellado RP, Moir A, Moriya S, Nagakawa E, Nanamiya H, Nakai S, Nygaard P, Ogura M, Ohanan T, O'Reilly M, O'Rourke M, Pragai Z, Pooley HM, Rapoport G, Rawlins JP, Rivas LA, Rivolta C, Sadaie A, Sadaie Y, Sarvas M, Sato T, Saxild HH, Scanlan E, Schumann W, Seegers JF, Sekiguchi J, Sekowska A, Seror SJ, Simon M, Stragier P, Studer R, Takamatsu H, Tanaka T, Takeuchi M, Thomaides HB, Vagner V, van Dijk JM, Watabe K, Wipat A, Yamamoto H, Yamamoto M, Yamamoto Y, Yamane K, Yata K, Yoshida K, Yoshikawa H, Zuber U, Ogasawara N. Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A.* 2003; 100:4678–4683. [PubMed: 12682299]
158. Zwolschen JH, Bhattacharjee JK. Genetic and biochemical properties of thialysine-resistant mutants of *Saccharomyces cerevisiae*. *J Gen Microbiol.* 1981; 122:281–287. [PubMed: 6798161]
159. Kalinowski J, Bachmann B, Thierbach G, Puhler A. Aspartokinase genes *lysCa* and *lysC β* overlap and are adjacent to the aspartate β -semialdehyde dehydrogenase gene *asd* in *Corynebacterium glutamicum*. *Mol Gen Genet.* 1990; 224:317–324. [PubMed: 1980002]

160. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* 2005; 33:D121–D124. [PubMed: 15608160]
161. Di Girolamo M, Busiello V, Di Girolamo A, Foppoli C, De Marco C. Aspartokinase III repression in a thialysine-resistant mutant of *E. coli*. *Biochem Int.* 1988; 17:545–554. [PubMed: 2849443]
162. Hirshfield IN, Tomford JW, Zamecnik PC. Thiosine-resistant mutants of *Escherichia coli* K-12 with growth-medium-dependent lysyl-tRNA synthetase activity. II. Evidence for an altered lysyl-tRNA synthetase. *Biochim Biophys Acta.* 1972; 259:344–356. [PubMed: 4552090]
163. Hirshfield IN, Zamecnik PC. Thiosine-resistant mutants of *Escherichia coli* K-12 with growth-medium-dependent lysyl-tRNA synthetase activity. I. Isolation and physiological characterization. *Biochim Biophys Acta.* 1972; 259:330–343. [PubMed: 4552089]
164. Di Girolamo M, Busiello V, Coccia R, Foppoli C. Aspartokinase III repression and lysine analogs utilization for protein synthesis. *Physiol Chem Phys Med NMR.* 1990; 22:241–245. [PubMed: 2129308]
165. Di Girolamo M, Coccia R, Blarzino C, Di Girolamo A, Busiello V. Degradation of thialysine- or selenalysine-containing abnormal proteins in *E. coli*. *Biochem Int.* 1988; 16:1033–1040. [PubMed: 3140818]
166. Phan TT, Schumann W. Development of a glycine-inducible expression system for *Bacillus subtilis*. *J Biotechnol.* 2007; 128:486–499. [PubMed: 17208325]
167. Kikuchi G, Motokawa Y, Yoshida T, Hiraga K. Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proc Jpn Acad Ser B Phys Biol Sci.* 2008; 84:246–263.
168. Shi Y. Histone lysine demethylases: emerging roles in development, physiology and disease. *Nat Rev Genet.* 2007; 8:829–833. [PubMed: 17909537]
169. Torres AG. The cad locus of *Enterobacteriaceae*: more than just lysine decarboxylation. *Anaerobe.* 2009; 15:1–6. [PubMed: 18572426]
170. Wieland M, Benz A, Klauser B, Hartig JS. Artificial ribozyme switches containing natural riboswitch aptamer domains. *Angew Chem Int Ed Engl.* 2009; 48:2715–2718. [PubMed: 19156802]

Fig. 1.

The general mechanism of riboswitch function. The nascent mRNA molecule alternates its conformation in response to the cognate metabolite. If a ligand is present, most riboswitches adopt a stable metabolite-bound conformation of the sensing domain and a stable conformation of the expression platform shown as a transcription terminator. In this case, gene transcription is prematurely terminated. If the ligand is absent, mRNA forms an anti-terminator conformation which allows the transcription to proceed through the entire gene. Complementary regions of mRNA are shown by colored boxes. The figure depicts transcription termination/antitermination mechanism of gene expression control used by majority of riboswitches. In addition to the repression of transcription in the presence of metabolites, riboswitches utilize several other means to modulate gene expression. For example, riboswitches can control translation initiation through occlusion of ribosome-binding site in bacteria or alternative splicing through regulation of accessibility to splicing elements in eukaryotes.

Fig. 2. The metabolic pathways controlled by lysine and glycine riboswitches. (A) Diaminopimelate pathway of lysine biosynthesis and lysine import pathways in various bacteria. Gene names are indicated in accordance to *E. coli* nomenclature. *B. subtilis* gene names are underlined. Putative lysine transport genes are boxed. Lysine riboswitches that repress expression of single genes or are located upstream of the first gene of the operon in any bacteria are shown with red lines. Most bacterial species, including *E. coli*, convert tetrahydrodipicolinate to diaminopimelate via succinylated DAP derivatives. Some species, including *B. subtilis*, use acetylated intermediates. *Bacillus sphaericus* perform conversion in one step using a unique DAP dehydrogenase encoded by the *ddh* gene. The figure is modified from refs. [56, 76]. (B) Simplified schematic of glycine metabolism and TCA of *Cand. P. ubique*. The glycine riboswitches that activate gene expression are shown by green arrows. The figure is adapted from ref. [85].

Fig. 3. Distribution of amino acid-responsive riboswitches in bacteria. A phylogenetic tree is shown on the left. The numbers indicate the total occurrences of each riboswitch in the bacterial group. The figure is modified from ref. [12] and updated using Rfam database [160] except the metagenome distribution which is based on ref. [12]. Note that available sequence information varies greatly for different taxonomic divisions and the number of predicted riboswitches in the environmental metagenomes in ref. [12] exceeds the number in ref. [80].

Fig. 4.

Representative mechanisms of gene expression control by amino acid-responsive riboswitches. Lysine-induced inhibition of transcription of the *T. maritima asd* gene (A) and repression of translation of the *E. coli lysC* gene (B). Glycine-induced activation of transcription of the *B. subtilis gcvT* (C) and translation of the *Cand. P. ubique glcB* (D). Complementary sequences involved in formation of alternative conformations of the regulatory elements are shown in magenta. The ribosome-binding site (Shine-Dalgarno sequence, SD) is depicted by a blue line.

Fig. 5.
 The consensus secondary structure models based on sequence alignments of lysine (A) and glycine (B) riboswitches from ref. [12] and structural information for the junctional region of the lysine riboswitch from refs. [51, 52]. Nucleotide and stem-loop conservation is depicted according to the symbols defined in the inset.

Fig. 6. Molecular determinants of lysine and glycine recognition by riboswitches. (A) Lysine analogs capable of specific binding to the *B. subtilis* *lysC* riboswitch [29, 56]. Pink, grey and blue shadings highlight changes in the compounds. The three-dimensional structures are available for the underlined compounds [52]. (B) Summary of the molecular recognition characteristics of the lysine riboswitch derived from the binding experiments using lysine analogs. Modified from ref. [56] with an addition from ref. [29]. The summary is in a good agreement with the structural data and does not indicate hydrogen bonds to carboxylate observed in refs. [51, 52]. (C) Glycine analogs that bind to glycine riboswitches [28, 88]. (D) Summary of the molecular recognition characteristics of glycine riboswitches derived from analog binding experiments [28, 88].

Fig. 7.

Chemical probing data on the glycine riboswitch shown on the secondary structure predicted from phylogenetic analysis. Glycine-induced effects are color-coded as indicated in the inset. (A) Results of the in-line probing [28], hydroxyl radical footprinting [53] and NAIM experiments [88] on the *V. cholerae* gcvT glycine riboswitch. The riboswitch regions 1–31 and 210–225 were not assessed by hydroxyl radical footprinting. Only strong glycine-induced changes from the hydroxyl radical assay are shown in the figure. NAIM data are presented by symmetric interference sites between the tandem aptamers. Note that nucleotides with modified bases used in NAIM also carry phosphorothioate modifications. The base-specific interferences indicated in figure reflect the interferences which were not observed with control parental phosphorothioate-modified nucleotides. (B) Summary of the probing data and regions of the riboswitch potentially involved in the riboswitch function. The regions possibly involved in formation of the glycine binding pocket are shaded in magenta. Glycine position within the riboswitch is predicted based on the biochemical data and analogy with lysine riboswitch. Asymmetric interference sites observed between the tandem sensing domains in the NAIM experiments are depicted by colored circles as in panel A. These sites (grey shadings) are implicated in the tertiary contacts that are involved in the cooperativity between riboswitch domains.

Fig. 8.

In-line probing and SHAPE experiments on lysine riboswitches. Lysine- and magnesium-induced changes are color-coded as indicated in the inset. Secondary structure schematic is prepared in accordance with the three-dimensional structure [51, 52]. (A) *T. maritima asd* riboswitch. In-line probing and SHAPE data from refs. [52] and [51], respectively. The P1 helix was not visualized in in-line probing. (B) *B. subtilis lysC* riboswitch. In-line probing and SHAPE data are from refs. [29] and [51], respectively. Only the core region of the riboswitch is shown.

Fig. 9. Structural features of the lysine riboswitch. Overall structure of the lysine-bound *T. maritima* riboswitch in front (A) and side (B) views [52]. The bound lysine is in red. Black frames emphasize the kissing loop and “molecular staple” regions shown in details in panels (C) and (D), respectively. Hydrogen bonds are depicted in dashed lines.

Fig. 10.

Details of lysine riboswitch-ligand interactions. (A) Structural representation of the lysine-binding pocket. Hydrogen and coordination bonds formed by lysine are depicted in dashed lines. (B) Schematic representation of lysine recognition. Hydrogen and coordination bonds are shown as dashed black and green lines, respectively. (C) Cross-section through the surface view of the lysine-binding pocket with bound lysine. Openings next to lysine termini and a cavity next to the C4 atom are shown with blue arrows. (D) Interaction of homoarginine (HArg) with the riboswitch in vicinity of the ligand side chain. The view shows homoarginine (in cyan) from the homoarginine-bound structure superposed with lysine (in red) from the lysine-bound structure. The G163 nucleotide from the lysine-bound and homoarginine-bound structures is in light pink and magenta, respectively. Arrow indicates conformational adjustment of the G163 sugar. Dashed lines show hydrogen bonds formed by the terminal nitrogen atoms of HArg. (E) Superposition of the lysine-binding pockets from the lysine-bound structure (pink) and lysine-free structure (blue). Bound lysine is shown as a reference. The structures are practically identical with ~ 0.3 Å conformational adjustment observed for G12 [52]. (F) Stabilization of the P1 helix by the bound lysine. Formation of the purine quartet G80•A162•G163•G11 on top of the P1 helix is shown by hydrogen bonds (dashed lines).