Regulation of Bacterial Gene Expression by Riboswitches

Wade C. Winkler[1] and Ronald R. Breaker[2]

[1]Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390; email: wade.winkler@utsouthwestern.edu

[2]Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520; email: ronald.breaker@yale.edu

Annu. Rev. Microbiol.
2005. 59:487–517

The *Annual Review of Microbiology* is online at micro.annualreviews.org

doi: 10.1146/annurev.micro.59.030804.121336

0066-4227/05/1013-0487$20.00

Key Words

aptamer, RNA switch, ribozyme, transcription attenuation, translation initiation

Abstract

Riboswitches are structured domains that usually reside in the noncoding regions of mRNAs, where they bind metabolites and control gene expression. Like their protein counterparts, these RNA gene control elements form highly specific binding pockets for the target metabolite and undergo allosteric changes in structure. Numerous classes of riboswitches are present in bacteria and they comprise a common and robust metabolite-sensing system.

Contents

INTRODUCTION

Microorganisms experience a wide variety of fluctuating conditions caused by changes in their surroundings or by shifting metabolic demands. Cells must be able to quantify these changes and concordantly alter expression of gene subsets in a measured manner. A great variety of mechanisms for controlling expression in response to extracellular and intracellular conditions have been identified (66). Most well-characterized among these are the numerous proteins that exert regulatory control over transcription initiation (9). The large-scale sequencing of bacterial genomes has led to the identification of hundreds of putative transcription factors per genome. For example, more than 200 transcription factors are predicted to be encoded by the *Bacillus subtilis* genome (79). Therefore, regulation of transcription initiation is a major mechanism by which bacteria orchestrate gene expression.

However, gene control at the level of transcription initiation constitutes only a portion of the regulatory mechanisms used by bacteria. Many different posttranscriptional control mechanisms have been identified during the past few decades, including those that regulate transcription elongation, transcription termination, translation initiation, translation termination, and mRNA stability (3, 11, 12, 20, 56, 85, 86, 107). Within the past several years, a number of discoveries have revealed that regulatory RNA structures are often used for posttranscriptional control of essential genes in bacteria (**Figure 1**). Some of these are small RNAs called riboregulators that function in *trans*, while other *cis*-acting RNA elements are located within the noncoding portions of mRNAs. Several comprehensive reviews on *trans*-acting RNAs have been published recently (21, 85). We focus our discussion primarily on *cis*-acting regulatory RNAs and, in particular, we highlight some of the advances made in the study of metabolite-sensing riboswitches (50, 65, 104).

GENE CONTROL BY PROTEIN AND RNA FACTORS

Although metabolic changes can influence gene expression by affecting the expression or function of protein factors, there are an increasing number of reports describing the roles played by noncoding RNA factors (21) and natural RNA aptamers (50). Furthermore, it appears that numerous types of these RNAs might be present in many bacteria (21, 95) and their precise roles in cellular function await further characterization. *cis*-acting regulatory RNAs range in structural sophistication from short RNA sequences to intricately folded riboswitches that carry high-affinity receptors

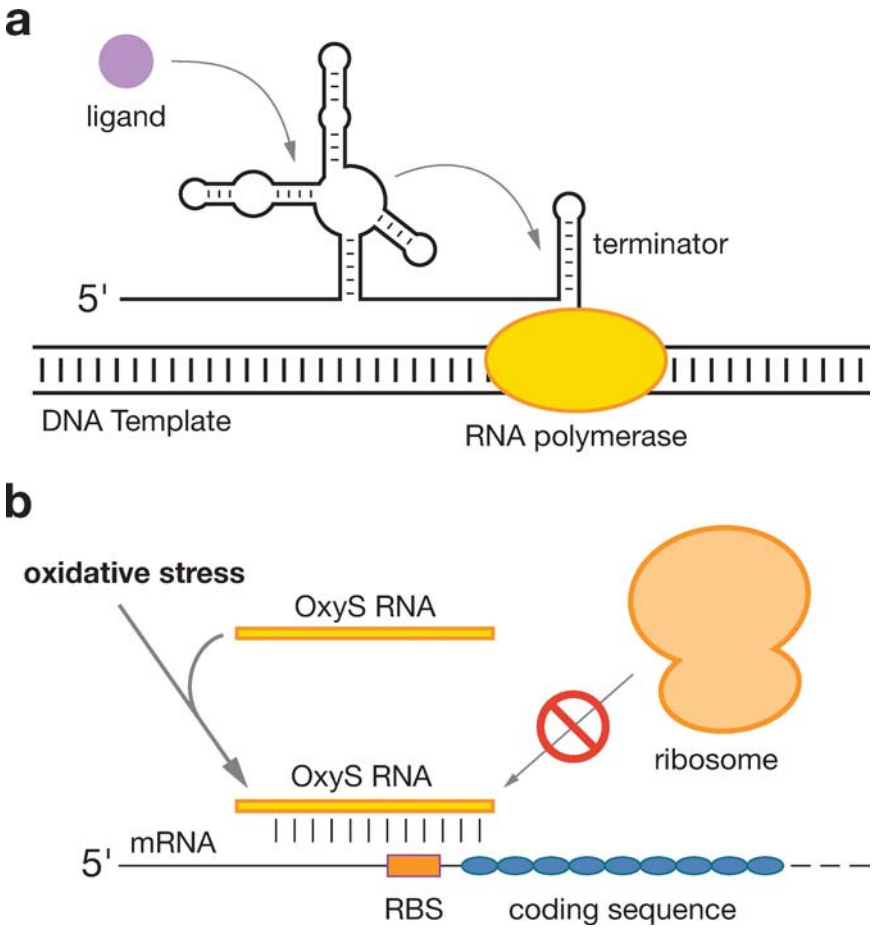

Figure 1

Examples of gene expression regulation by *cis*- and *trans*-acting RNAs. (*a*) A SAM-dependent riboswitch located within the 5′ UTR of an mRNA acts in *cis* to inhibit transcription of the full-length mRNA upon ligand binding (16, 52, 108). (*b*) During conditions of oxidative stress the OxyS riboregulator, a *trans*-acting RNA, interacts with *fhlA* transcripts and inhibits translation initiation (2). RBS, ribosome-binding site.

for small organic molecules. In general, regulatory RNAs located within the 5′ UTRs of transcripts receive intracellular metabolic signals that stimulate specific conformational changes. In turn, these structural changes exert control over associated genes.

A survey of *cis*-acting RNAs in *B. subtilis* reveals that greater than 4% of its genes are regulated, at least in part, by the action of genetic control mechanisms involving RNA (103) (**Table 1**). In each instance, the RNA elements residing within mRNAs have been classified by three categories based on the effector that is bound: (*a*) protein, (*b*) RNA, and (*c*) metabolite. Typically, the *cis*-acting RNA elements that bind to protein factors are simple in sequence and structure because it is the protein factor that forms the binding sites for the RNA and responds allosterically to various ligands. In contrast, riboswitches bind metabolites directly and therefore form more complex sequence and structural features.

What is striking about this assessment is that, for mechanisms that operate after transcription has been initiated, there are more examples of genetic control systems that do not involve proteins than those that do. Given that no systematic genome-wide assessment of RNA-based genetic control systems has been made, it is likely that this list is far from comprehensive. Although it is difficult to project from current data how substantial the role is for RNA in posttranscriptional gene control, surprisingly it is not clear whether gene control factors made of RNA or protein will dominate.

UTR: untranslated region

Table 1 RNA-mediated genetic control in *B. subtilis*[a]

Effector molecule	Number of regulated transcriptional units [total number of genes; percent of *B. subtilis* ORFs]	Gene categories
Survey of RNA-mediated genetic control in *Bacillus subtilis*		
Protein	21 [45; 1.1%]	Tryptophan, folate, glycerol, histidine, and pyrimidine metabolism; sugar catabolism; Rho synthesis; Cold shock response
RNA	19 [33; 0.8%]	Aminoacyl-tRNA synthetases; amino acid biosynthesis and transport
Metabolite	36 [89; 2.2%]	TPP, FMN, adenosylcobalamin, SAM, lysine, guanine, adenine, glycine, and GlcN6P

[a]Data are discussed in the text and are extracted and combined from Reference 103. *cis*-acting regulatory RNAs can receive metabolic cues via protein, RNA, and metabolite effectors. The number of regulated transcripts, number of genes encoded within these transcripts, and their general functional categories are indicated.

r-protein: ribosomal protein

ThrS: threonyl-tRNA synthetase

Riboswitches that sense metabolites and control gene expression via allosteric changes in RNA structure are remarkably widespread throughout bacteria and are used for regulating many fundamental biochemical pathways. Recent reports of new riboswitch classes (4, 48) increase the repertoire and the functional sophistication for these RNA gene control elements and provide validation that RNAs have a large intrinsic capacity for novel function that also could be harnessed as promising tools and targets for biomedical applications (8, 62, 88). In addition, the characteristics of modern riboswitches hint at the structural sophistication that RNAs may have exhibited several billion years ago during the hypothetical "RNA world" (39).

Metabolic Sensing via Protein Intermediates

Examples of proteins that regulate bacterial gene expression through RNA binding have been well described for several decades (37). The RNA ligands forming these protein-binding sites vary in structural complexity from short, single-stranded sequences (58) to magnesium-dependent tertiary structures that are required for high-affinity binding (73, 74). RNA-binding regulatory proteins can be placed into two overall categories: those that associate directly to their RNA-binding sites and those that exhibit signal-responsive binding ability. Numerous examples have been described and reviewed elsewhere (75, 86).

Most examples of the former category are of proteins that self-regulate expression through feedback repression. For example, regulation of r-protein operons is often achieved through the use of RNA structures that, when bound by an encoded r-protein, reduce expression of themselves and the remaining genes within the operon (82). This regulatory process ensures that excess r-proteins are not needlessly produced. There is growing experimental evidence that mRNA-binding sites for r-proteins exhibit structural mimicry to rRNA-binding sites, and competitive binding between these RNA elements is believed to form the basis of regulatory control (32, 53, 83). Similarly, structural mimicry between a functional RNA-binding site and a regulatory-binding site has been proposed for regulation of *Escherichia coli* ThrS, wherein a tRNA-like structure forms at the 5′ end of *thrS* mRNA and is bound by ThrS to repress translation (71).

There are also many proteins whose RNA-binding activity is controllable in a ligand-responsive manner. Metabolite binding can induce the protein factor to either increase or decrease its affinity for RNA. As an example of the former, 11 monomers of the *B. subtilis* TRAP (trp RNA-binding attenuation

protein) protein associate to form a ring-like structure that binds to a series of (G/U)AG repeats at the 5′ end of certain mRNAs (3). As intracellular tryptophan concentration increases, a tryptophan molecule binds to each protein subunit and stimulates association of the complex to the trinucleotide repeat sequences in the target mRNA, with one trinucleotide sequence associating with each protein subunit. Binding of the complex alters Watson-Crick base-paring interactions within the mRNA 5′ UTR, which ultimately leads to decreased expression of the associated genes.

Some RNA-binding proteins respond to cellular signals by binding to other proteins or by undergoing phosphorylation. For example, several operons in gram-positive bacteria encoding for sugar catabolism genes are regulated by RNA-binding proteins that contain two domains (PRD-I and PRD-II) subjected to independent phosphorylation events (86). PRD-II phosphorylation is stimulated by the presence of glucose (the preferred substrate), and phosphorylation of PRD-I is stimulated by the presence of the alternative sugar substrate. RNA-binding activity and subsequent activation of expression is stimulated only when PRD-I becomes phosphorylated and PRD-II is unphosphorylated.

Metabolic Sensing via RNA-RNA Interactions

Several examples of *trans*-acting oligonucleotides affect bacterial gene expression through controlled formation of antisense interactions to portions of 5′ UTRs of transcripts. A unique variation on this theme is the use of T box RNAs, which serve to regulate expression of aminoacyl-tRNA synthetases and amino acid biosynthesis and transport genes (27). T box RNAs, found within the 5′ UTRs of mRNAs, interact directly with uncharged tRNAs to promote expression of genes located immediately downstream of these elements (25). Presumably a conformational change develops upon binding to the correct uncharged tRNA, and this ultimately results in gene activation. For example, as intracellular tyrosine decreases, the ratio of charged-to-uncharged $tRNA^{Tyr}$ also decreases, which in turn leads to induction of expression for tyrosine-related genes through increased binding of $tRNA^{Tyr}$ to T box RNAs.

Two critical points of RNA-RNA interactions govern the specificity of this reaction. T box RNAs each contain a trinucleotide sequence (specifier sequence) that base-pairs with tRNA anticodons (28). Alteration of the specifier sequence by site-directed mutation alters the specificity that the T box exhibits for tRNA binding, which demonstrates that the tRNA being sensed is recognized at least in part by Watson-Crick base-pairing interactions between the mRNA and the anticodon loop of the tRNA ligand. A second base-pairing interaction is formed with the 3′-terminal sequence of tRNAs, which appears to stabilize a conformation that promotes expression of associated genes (30). T box RNAs are likely to recognize additional features of tRNA, including their overall three-dimensional shape (24), although the degree of structural preorganization of T box RNAs remains to be determined. Finally, the mechanism for T box regulation has been reconstituted in vitro in the absence of protein factors other than RNA polymerase (31, 67). This finding demonstrates that at least some T box RNAs do not have an obligate need for protein factors to serve as tRNA-sensing genetic switches.

Detection of Physical Cues by mRNAs

Organisms can also sense changes in their physical environment, such as temperature fluctuations, and make adjustments in gene expression. Indeed, RNA seems well suited to form meta-stable structures that can be shifted on the basis of even subtle changes in temperature. Such thermosensitive RNA structures have been implicated as one solution that bacteria use for temperature sensing

RBS: ribosome-binding site

TPP: thiamine pyrophosphate

FMN: flavin mononucleotide

SAM: *S*-adenosylmethionine

(61). For example, *Listeria monocytogenes*, a mammalian pathogen, upregulates virulence genes when its surroundings have reached the host's temperature of 37°C. In this instance, temperature changes lead to altered expression of PrfA, a transcriptional activator that in turn regulates virulence gene production (38). At low temperatures, the mRNA encoding this protein forms a secondary structure that prevents translation initiation, most likely by sequestering the RBS. Mutations that disrupt this structure and expose the RBS lead to increased expression. At 37°C, base-pairing within the secondary structure is disrupted in a manner that allows for improved translation initiation.

Similar temperature-sensing RNAs have been proposed to regulate phage lambda genes, *Escherichia coli* heat shock sigma factor, RpoH, as well as other genes (1, 61). Given that RNA structures can be highly dependent on temperature and on the ionic conditions under which they fold, it seems that many additional examples of these gene control systems might be awaiting discovery.

METABOLITE-SENSING RIBOSWITCHES

Riboswitch Organization

Bacteria use metabolite-sensing riboswitches to quantify intracellular metabolite concentrations and to make the appropriate changes in the expression of relevant genes. To date, riboswitches have been identified almost exclusively within the 5′ UTRs of bacterial mRNAs. However, one class of riboswitches that binds TPP is exceptionally common in bacteria and has also been identified in plants and fungi (44, 87).

Each riboswitch class carries an aptamer domain that is highly conserved in sequence and structure even among widely diverse organisms and likewise is conserved when multiple variants of a given riboswitch are present in a single organism. Upon binding of the target metabolite to this aptamer, conformational changes result that in turn modulate expression of genes carried by the mRNA. These conformational changes occur in a functional domain called an expression platform, which is located immediately downstream from the aptamer (102, 104) (**Figure 2**). Ligand binding is achieved solely by the aptamer, which usually does not require the presence of the expression platform or the action of protein factors. The lengths of the aptamer domains reported to date range from ~70 to ~200 nucleotides, and the expression platforms are far more diverse in sequence, structure, and size.

Riboswitches are almost always adjacent to genes involved in the synthesis or transport of their metabolite ligand. Many riboswitch candidates have been identified by computer-assisted searches for highly conserved RNA domains that reside in intergenic regions of bacterial genomes (4, 48, 68, 69, 88, 92, 93, 108). In some instances, these conserved domains appear to be associated with intrinsic transcription terminators that might serve as riboswitch expression platforms. That riboswitches almost always appear immediately upstream of the genes they control can help identify candidate metabolites that might trigger riboswitch function. Numerous uncharacterized genes with unknown functions appear to be controlled by riboswitches responsive to known metabolite effectors. The discovery of riboswitches has therefore served as a powerful tool for inferring functions of uncharacterized genes (68, 69, 92, 93). For example, an uncharacterized transport protein regulated by a TPP-sensing riboswitch likely functions as a thiamine transporter.

Confirming and Classifying Riboswitches

To date, reports of nine separate riboswitch classes have been made in which there is direct evidence of metabolite binding in the absence of protein factors (**Figure 3**) (**Table 2**). Varying levels of biochemical and genetic evidence for riboswitch classes that sense coenzyme B_{12} (59, 60), TPP (55, 102), FMN (55, 105), SAM

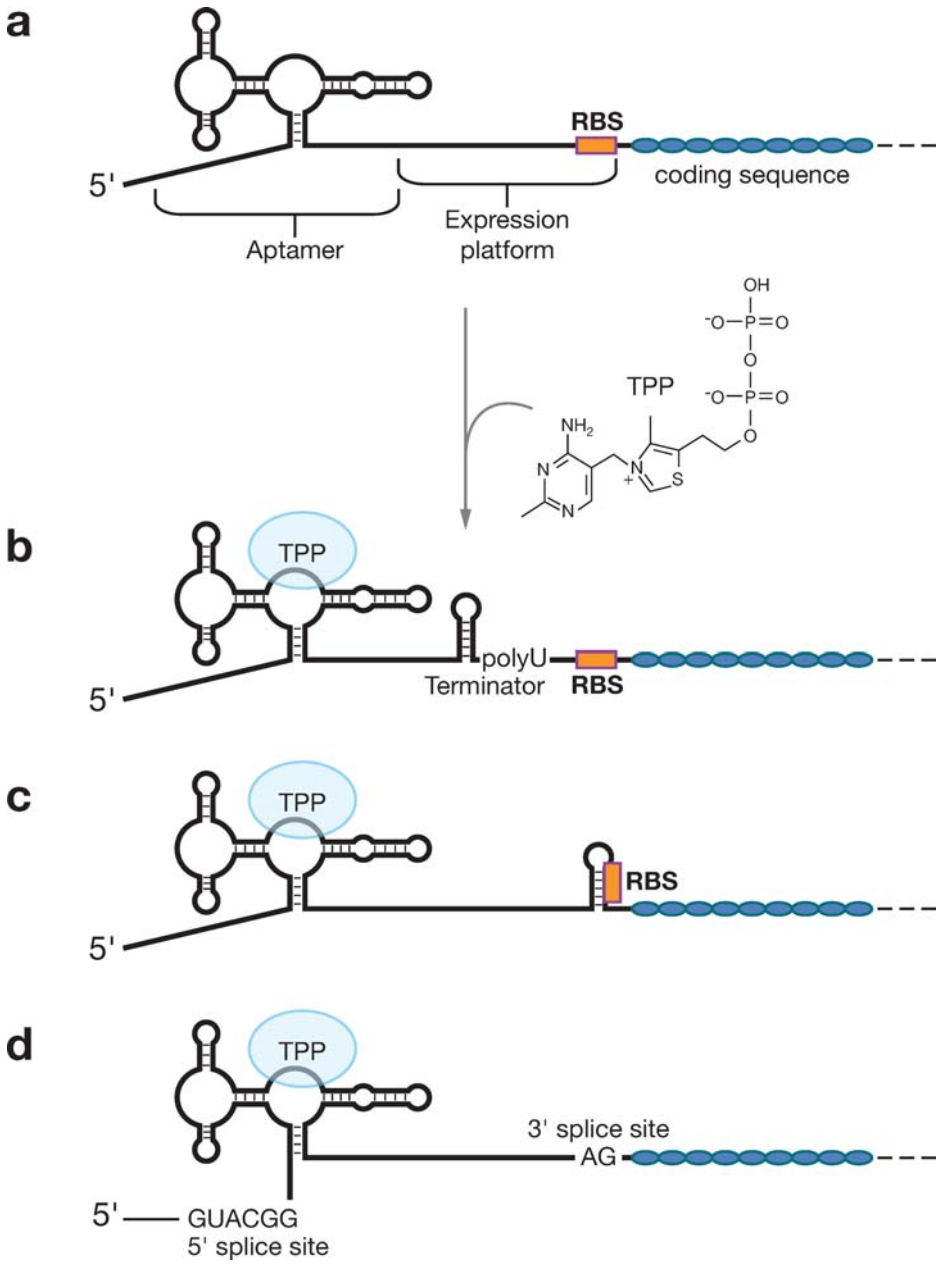

Figure 2

Organization of riboswitch RNAs. (*a*) General architecture of a TPP-sensing riboswitch. The highly conserved aptamer domain binds the ligand, and the expression platform resides downstream of the aptamer but upstream of the adjoining open reading frame. (*b*) Transcription termination. Some riboswitches induce formation of an intrinsic transcription terminator upon ligand binding. (*c*) Translation inhibition. Some riboswitches induce formation of a helix that sequesters the RBS, thereby reducing efficiency of translation initiation. (*d*) Splicing control. The aptamer domain of fungal TPP riboswitches is flanked by consensus splice site sequences (87), which suggests that TPP binding affects splicing efficiency.

GlcN6P: glucosamine-6-phosphate

(16, 52, 108), lysine (29, 88), guanine (48), adenine (49), GlcN6P (107), and glycine (51) have been reported. Evidence for the existence of at least two additional classes will be reported soon (R. Breaker et al., unpublished data). As mentioned above, the aptamer domains exhibit a high degree of sequence and structural conservation for each class, whereas expression platforms exhibit little to no primary sequence conservation (48, 59, 68, 69, 70, 88, 92, 93, 108). More historical overviews of the discovery and initial biochemical characterization of these motifs have been provided elsewhere (50, 65). However, it is worth briefly

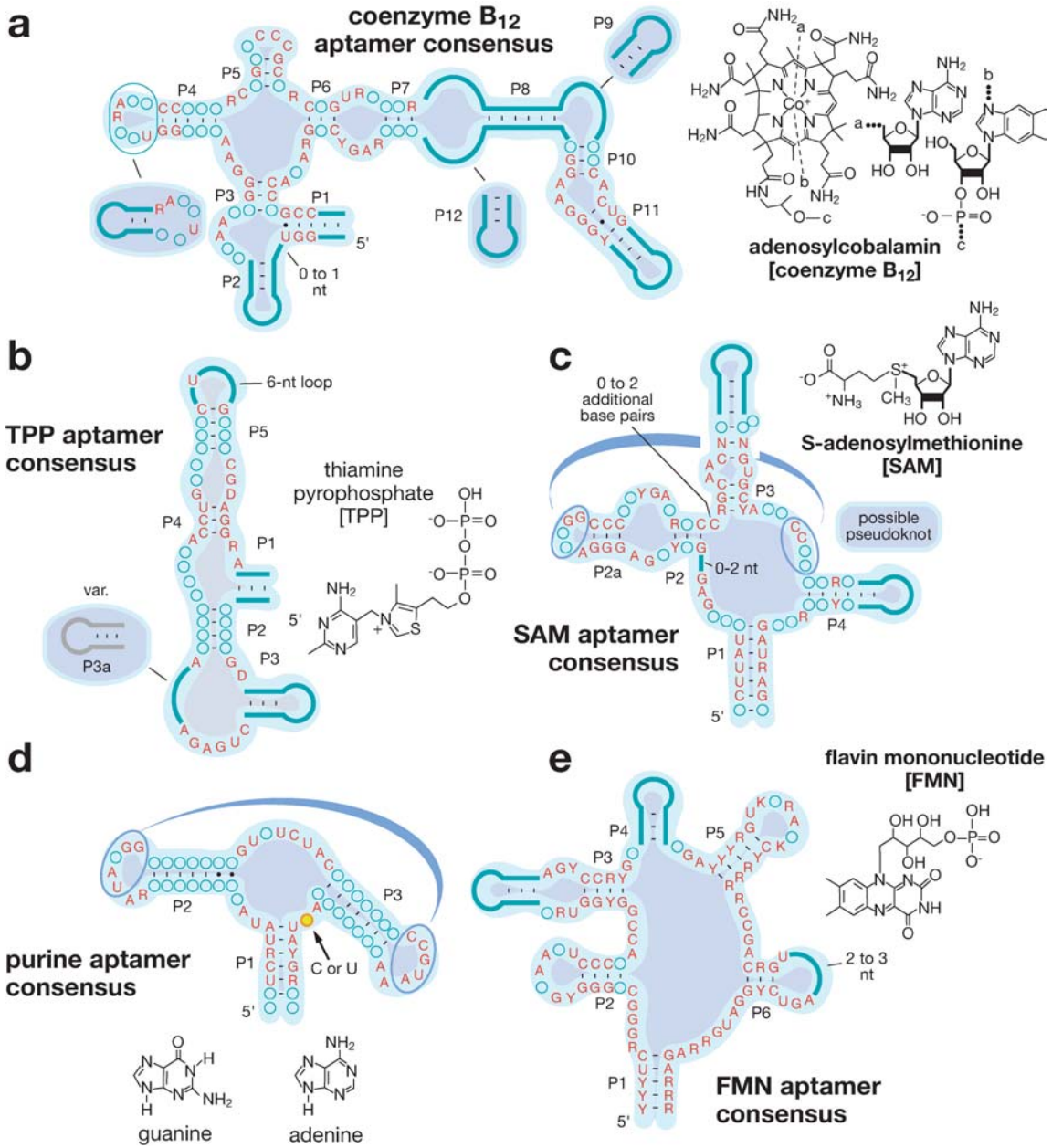

Figure 3

Consensus sequences and secondary structure models for riboswitch aptamer domains. Red letters signify nucleotides conserved in >90% of aligned sequences. Black letters signify nucleotides conserved in >80% of aligned sequences. (*a*) Coenzyme B_{12} aptamer. (*b*) TPP aptamer. (*c*) SAM aptamer. (*d*) Purine aptamer. (*e*) FMN aptamer. (*f*) Lysine aptamer. (*g*) Glycine aptamers. (*h*) GlcN6P-sensing ribozyme. Data were extracted and combined from References 4, 23, 48, 55, 59, 60, 68, 69, 88, 92, 93, and 108.

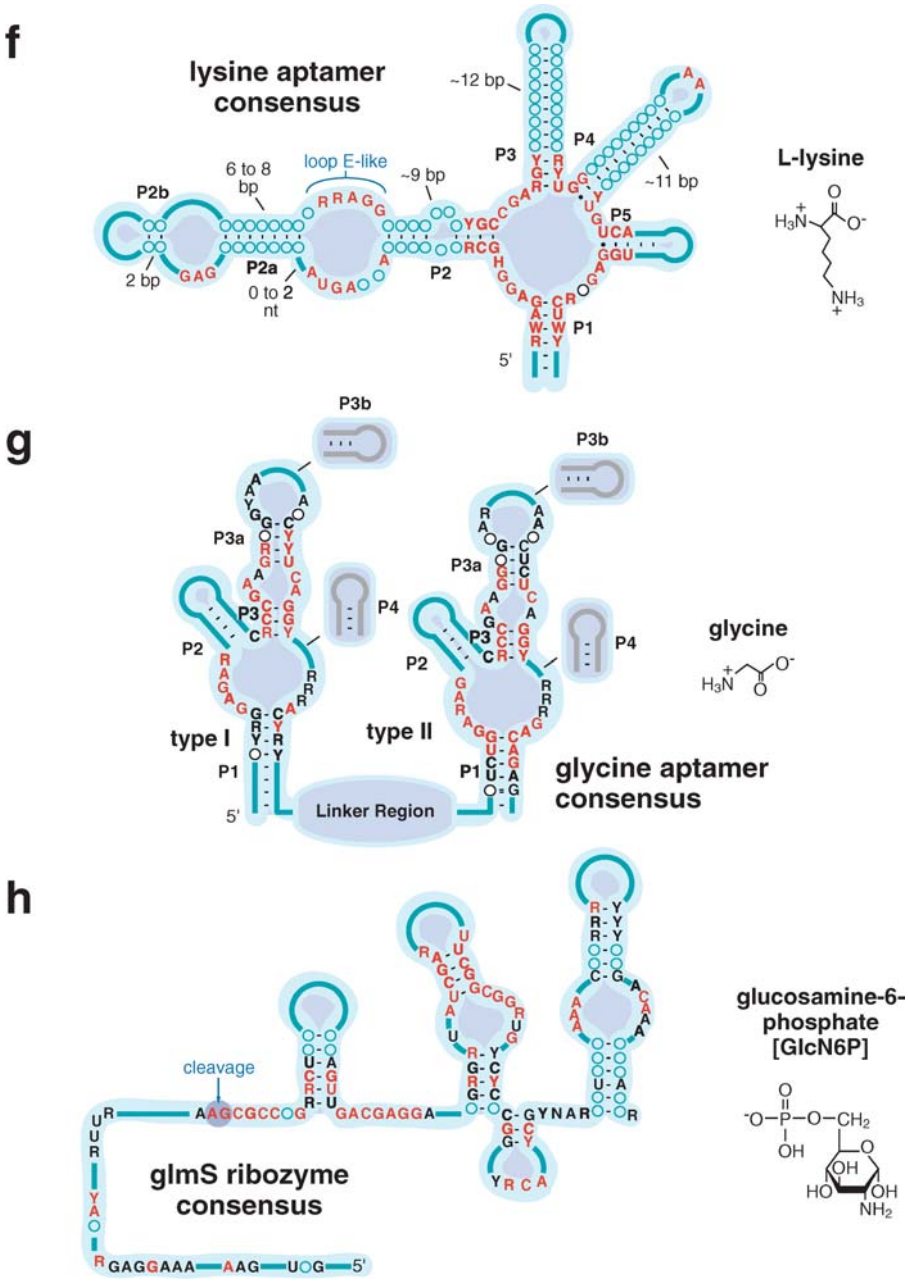

Figure 3

(*Continued*)

reviewing some of the biochemical experiments that established that metabolites are directly bound by natural RNA aptamers.

Evidence for the existence of several riboswitches, based on classical genetics studies, was reported as early as the 1970s with metabolic genes for lysine (6, 96) and soon after with metabolic genes for coenzyme B_{12} (17, 97), purines (10, 15) and FMN (43). However, it was not immediately evident or even suggested that RNA might be playing the dominant role during the gene control

Table 2 Biological distribution of riboswitch RNAs[a]

Gene categories	Representative genera
B_{12} riboswitch	
Cobalamin synthesis and transport Cobalt transport Aerobic and anaerobic ribonucleotide reductase Glutamate fermentation Succinate fermentation Uncharacterized genes	Proteobacteria: α-proteobacteria [*Mesorhizobium*, *Bradyrhizobium*, *Sinorhizobium*, *Brucella*, *Agrobacterium*, *Rhodopseudomonas*, *Rhodobacter*, *Caulobacter*] Proteobacteria: β-proteobacteria [*Nitrosomonas*, *Methylobacillus*, *Ralstonia*] Proteobacteria: γ-proteobacteria [*Escherichia*, *Salmonella*, *Klebsiella*, *Yersinia*, *Vibrio*, *Pasteurella*, *Pseudomonas*, *Shewanella*, *Xanthomonas*] Proteobacteria: δ-proteobacteria [*Geobacter*] *Deinococcales* [*Deinococcus*] *Bacillales* [*Bacillus*, *Listeria*] *Lactobacillales* [*Enterococcus*] *Clostridiales/Thermoanaerobacteriales* [*Clostridium*, *Thermoanaerobacter*] *Actinomycetales* [*Corynebacterium*, *Mycobacterium*, *Propionibacterium*, *Streptomyces*] Cyanobacteria [*Anabaena*, *Thermosynechococcus*] *Chlorobiales* [*Chlorobium*] *Bacteroidales* [*Bacteroides*] *Thermotogales* [*Thermotoga*] *Spirochaetales* [*Treponema*, *Leptospira*] Archaea [*Aeropyrum*, *Sulfolobus*, *Pyrococcus*]
FMN riboswitch	
Riboflavin biosynthesis and transport	Proteobacteria: α-proteobacteria [*Rhizobium*, *Mesorhizobium*, *Bradyrhizobium*, *Sinorhizobium*, *Brucella*, *Agrobacterium*] Proteobacteria: β-proteobacteria [*Bordetella*, *Chromobacterium*, *Ralstonia*] Proteobacteria: γ-proteobacteria [*Coxiella*, *Shewanella*, *Escherichia*, *Klebsiella*, *Pasteurella*, *Haemophilus*, *Vibrio*, *Xanthomonas*, *Salmonella*, *Yersinia*, *Shigella*, *Photorhabdus*, *Actinobacillus*, *Pseudomonas*] *Deinococcales* [*Deinococcus*] *Fusobacterales* [*Fusobacterium*] *Bacillales* [*Bacillus*, *Oceanobacillus*, *Staphylococcus*, *Listeria*] *Lactobacillales* [*Enterococcus*, *Streptococcus*, *Lactococcus*, *Lactobacillus*] *Clostridiales/Thermoanaerobacteriales* [*Clostridium*, *Thermoanaerobacter*] *Actinomycetales* [*Corynebacterium*, *Mycobacterium*, *Propionibacterium*, *Streptomyces*] *Thermotogales* [*Thermotoga*]
TPP riboswitch	
Thiamine synthesis, phosphorylation, and transport Uncharacterized genes	Proteobacteria: α-proteobacteria [*Mesorhizobium*, *Bradyrhizobium*, *Rhizobium*, *Sinorhizobium*, *Brucella*, *Agrobacterium*, *Rhodopseudomonas*, *Caulobacter*] Proteobacteria: β-proteobacteria [*Bordetella*, *Neisseria*, *Chromobacterium*, *Ralstonia*] Proteobacteria: γ-proteobacteria [*Pasteurella*, *Haemophilus*, *Aeromonas*, *Coxiella*, *Shewanella*, *Vibrio*, *Xanthomonas*, *Thiocystis*, *Escherichia*, *Salmonella*, *Shigella*, *Erwinia*, *Photorhabdus*, *Yersinia*, *Pseudomonas*] Proteobacteria: δ-proteobacteria [*Geobacter*, *Bdellovibrio*] Proteobacteria: ε-proteobacteria [*Helicobacter*, *Campylobacter*] *Deinococcales* [*Deinococcus*] *Bacillales* [*Bacillus*, *Listeria*, *Staphylococcus*, *Oceanobacillus*] *Lactobacillales* [*Enterococcus*, *Streptococcus*, *Lactococcus*, *Lactobacillus*]

(*Continued*)

Table 2 (*Continued*)

Gene categories	Representative genera
	Clostridiales/Thermoanaerobacteriales [*Clostridium*, *Thermoanaerobacter*]
	Actinomycetales [*Corynebacterium*, *Mycobacterium*, *Bifidobacterium*, *Streptomyces*]
	Cyanobacteria [*Prochlorococcus*, *Nostoc*, *Synechocystis*, *Thermosynechococcus*, *Gloeobacter*]
	Chlorobiales [*Chlorobium*]
	Bacteroidales [*Bacteroides*]
	Thermotogales [*Thermotoga*]
	Fusobacterales [*Fusobacterium*]
	Spirochaetales [*Treponema*, *Leptospira*]
	Archaea [*Thermoplasma*]
	Eukaryotes [*Fusarium*, *Aspergillus*, *Neurospora*, *Poa*, *Triticum*, *Oryza*, *Arabidopsis*]
Lysine riboswitch	
Lysine sythesis and transport	Proteobacteria: γ-proteobacteria [*Shewanella*, *Pasteurella*, *Haemophilus*, *Vibrio*, *Escherichia*, *Shigella*]
Lysine catabolism	*Bacillales* [*Bacillus*, *Staphylococcus*, *Oceanobacillus*]
	Clostridiales/Thermoanaerobacteriales [*Clostridium*, *Thermoanaerobacter*]
	Thermotogales [*Thermotoga*]
SAM riboswitch	
Methionine biosynthesis	Proteobacteria: γ-proteobacteria [*Xanthomonas*]
Cysteine biosynthesis	Proteobacteria: δ-proteobacteria [*Geobacter*]
Methionine recycling	*Bacillales* [*Bacillus*, *Listeria*, *Staphylococcus*, *Oceanobacillus*]
Methylene tetrahydrafolate reductase	*Clostridiales/Thermoanaerobacteriales* [*Clostridium*, *Thermoanaerobacter*]
SAM synthesis	*Actinomycetales* [*Streptomyces*]
Metabolite transport	*Cyanobacteria* [*Gloeobacter*]
Uncharacterized genes	*Chlorobiales* [*Chlorobium*]
	Fusobacterales [*Fusobacterium*]
Purine riboswitches	
Purine synthesis and transport	Proteobacteria: γ-proteobacteria [*Vibrio*]
	Proteobacteria: δ-proteobacteria [*Bdellovibrio*]
Uncharacterized genes	*Bacillales* [*Bacillus*, *Listeria*, *Staphylococcus*, *Oceanobacillus*, *Geobacillus*]
	Lactobacillales [*Enterococcus*, *Streptococcus*, *Lactococcus*]
	Clostridiales/Thermoanaerobacteriales [*Clostridium*, *Thermoanaerobacter*]
	Fusobacterales [*Fusobacterium*]
Glycine riboswitch	
Glycine catabolism and efflux	Proteobacteria: α-proteobacteria [*Mesorhizobium*, *Bradyrhizobium*, *Sinorhizobium*, *Brucella*, *Agrobacterium*, *Rhodopseudomonas*, *Caulobacter*]
	Proteobacteria: β-proteobacteria [*Ralstonia*]
	Proteobacteria: γ-proteobacteria [*Vibrio*]
	Bacillales [*Bacillus*, *Listeria*, *Staphylococcus*, *Oceanobacillus*]
	Lactobacillales [*Streptococcus*]
	Clostridiales/Thermoanaerobacteriales [*Clostridium*]
	Actinomycetales [*Mycobacterium*, *Streptomyces*]

(*Continued*)

Table 2 (*Continued*)

Gene categories	Representative genera
GlcN6P riboswitch	
Synthesis of GlcN6P	*Deinococcales* [*Deinococcus*] *Bacillales* [*Bacillus*, *Listeria*, *Staphylococcus*, *Oceanobacillus*] *Lactobacillales* [*Lactobacillus*, *Enterococcus*] *Clostridiales/Thermoanaerobacteriales* [*Clostridium*, *Thermoanaerobacter*] *Fusobacterales* [*Fusobacterium*]

[a]Categories of riboswitch-regulated genes are shown for each riboswitch class. Representative genera for which riboswitches have been identified are shown in the second column. Data were extracted and combined from References: 4, 18, 23, 48, 59, 68, 69, 88, 92, 93, 94, and 108.

processes until more recent studies on the control of these genes (18, 42, 64) and on new gene control mysteries involving methionine (26) and thiamine (54, 84) were published. Specifically, mutations in these *cis*-acting regions abrogated regulation of the downstream genes, yet the nature of their effector signals (sometimes assumed to be protein mediated) remained elusive.

Several lines of biochemical investigation were then undertaken to test the hypothesis that representatives of each riboswitch class could function as direct receptors for small molecules without the need for proteins. An assay to determine structural features of RNA, called in-line probing, was employed as a useful method for determining whether structural changes occur when specific metabolites are provided (80, 81). This method relies upon the comparative analysis of spontaneous cleavage events that occur at specific internucleotide linkages. Spontaneous cleavage, which occurs by internal phosphoester transfer, requires an in-line geometry between the attacking nucleophile (a 2′ oxyanion), the electrophile (the adjacent phosphorus center), and the leaving group (5′ oxyanion). If nucleotides are base-paired or otherwise constrained, the rate of cleavage of the adjoining phosphodiester linkage is slow. However, if the bases are unconstrained by secondary or tertiary structures, then the internucleotide linkages in this region are structurally more flexible and an in-line geometry that favors more rapid spontaneous cleavage frequently will be sampled. Analyses of the products of spontaneous RNA fragmentation by gel electrophoresis provide a measure of RNA structure, including a simple means by which to establish allosteric changes that are brought about when riboswitch RNAs bind their target metabolites (50).

For each of the riboswitch classes examined, in-line probing experiments have established that distinct conformational changes occur only when the corresponding metabolite is mixed with its natural RNA aptamer (60, 102, 105). Also important is the observation that these ligand-induced conformational changes do not occur with unrelated metabolites or when the target metabolite is added to unrelated RNA sequences. Similarly, these studies demonstrate that most riboswitch RNAs exhibit a high level of molecular discrimination by rejecting many analogs of their target metabolites.

A nuclease cleavage assay also was used to demonstrate that certain metabolites preclude the binding of DNA oligonucleotide probes to representative FMN-dependent (55) and SAM-dependent (16, 52) riboswitches. In this approach, RNase H cleaves riboswitch RNAs only if a complementary DNA probe base-pairs with the RNA. However, metabolite-induced changes in oligonucleotide affinity cause a loss of RNase H cleavage activity, suggesting that conformational changes that prevent oligonucleotide binding are stabilized by metabolite binding.

Further biochemical evidence for protein-free metabolite binding by RNA was obtained by equilibrium dialysis. Two chambers

of the equilibrium dialysis system are separated by a membrane with a 5000-Da molecular mass cut-off (60, 102, 108). Riboswitch RNAs are added to one chamber and a sample of the corresponding radiolabeled metabolite is added to the other. Upon equilibration, the small metabolite becomes unequally distributed in favor of the chamber that holds the metabolite-binding RNA. However, if a disruptive mutation is present in the RNA used, or if an excess of unlabeled metabolite is added to the system, then the radiolabeled metabolites distribute equally between the two chambers.

Fluorescence techniques also have demonstrated that FMN is bound by the aptamer domain of certain FMN-responsive riboswitches. FMN is intrinsically fluorescent with a maximum near 520 nm, and its fluorescence intensity becomes quenched upon association with corresponding riboswitch RNAs (55). This quenching effect has demonstrated ligand binding in the complete absence of proteins and similarly established the kinetics of FMN binding to various riboswitch constructs based on the *ribD* riboswitch from *B. subtilis* (101).

Metabolite Binding by Natural Aptamers

In-line probing and equilibrium dialysis binding assays have been adapted to establish apparent K_d values for the RNA-ligand interactions, which range from approximately 1 nM (48) to 30 μM (51) for the various riboswitch aptamers. Thus, natural aptamers can bind tightly to their target compounds with K_d values that are more than sufficient for their molecular-sensing roles. When the molecular recognition characteristics of riboswitches were examined in more detail, each aptamer was shown to discriminate strongly against chemicals that are structurally similar to the target metabolites (48, 88, 108). In the sections below, we briefly describe the structural and functional features of riboswitch classes that have been reported to date.

Guanine and adenine riboswitches. The largest amount of data exists for two closely related classes of purine-specific riboswitches that selectively bind guanine or adenine. The demonstrations that guanine (48) and adenine (49) are directly bound by RNA in the absence of protein factors addressed a long-standing mystery regarding the control of various purine transport and biosynthesis genes in *B. subtilis* (10). The PurR protein is a genetic control factor that senses adenine and controls the expression of certain purine metabolic genes (98). It also was known that the 12-gene operon, encoding genes necessary for the de novo biosynthesis of purines, was repressed by high concentrations of guanine or related compounds such as xanthine and hypoxanthine. Although PurR represses this operon when adenine concentrations are high, the action of this protein factor did not explain this operon's dependence on guanine.

Biochemical studies using the 5′ UTR of the *xpt-pbuX* operon (48) and with the 5′ UTR of the *ydhL* gene (49), both from *B. subtilis*, demonstrate that riboswitches can form highly selective binding pockets that discriminate between guanine and adenine. The basis for this tight and selective binding has recently been established for two purine-binding riboswitches. Structural models based on X-ray crystallography data have been generated for the *xpt-pbuX* riboswitch from *B. subtilis* when bound to guanine (77) and hypoxanthine (5). These models depict a binding site that entirely engulfs the ligand (**Figure 4*a***). Numerous hydrogen-bonding contacts are formed with the target purine with functional groups from specific nucleotides and even a ribose of the aptamer (**Figure 4*b***). The architecture proposed for this ligand-binding site also is consistent with structural data generated by NMR (63).

Another interesting feature of the purine-binding riboswitches is that a single C-to-U change in the aptamer core (nucleotide 74 of the *xpt-pbuX* RNA) results in a change in specificity from guanine to adenine (49). This nucleotide forms a standard

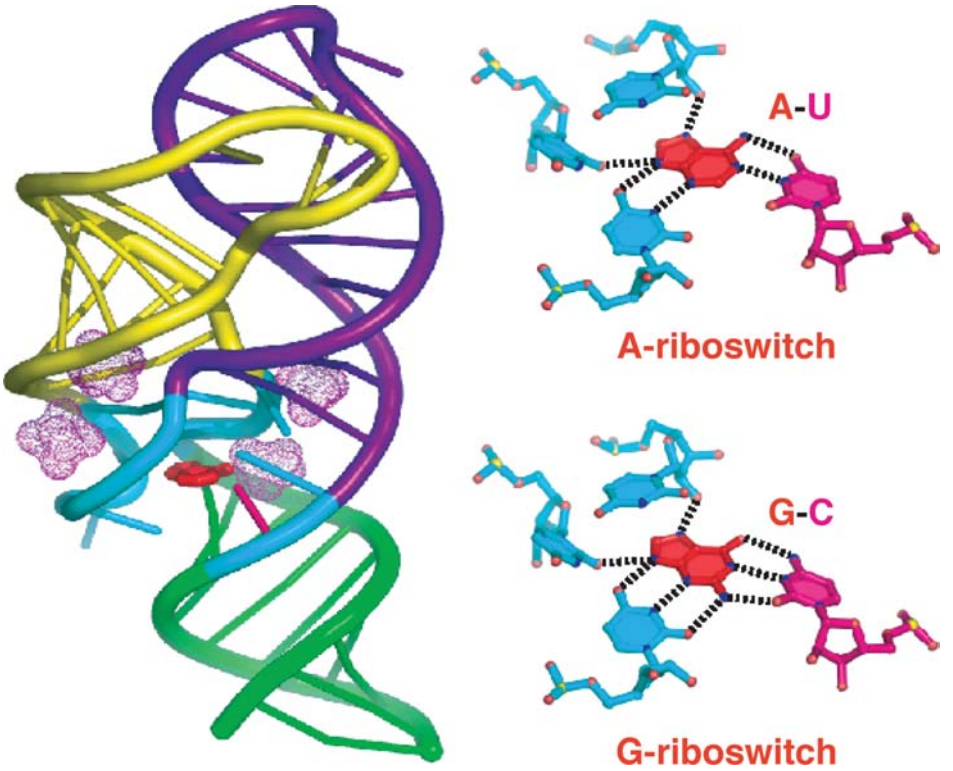

Figure 4

Structural model for a guanine-specific riboswitch and the formation of purine-binding pockets. (*a*) Model for the aptamer of the *xpt-pbuX* riboswitch from *B. subtilis* bound to hypoxanthine (5). (*b*) Comparison of the adenine- and guanine-specific-binding sites from the *Vibrio vulnificus add* and *B. subtilis xpt-pbuX* riboswitches (77). Nucleotides depicted form direct contacts with the adenine or guanine ligands (*red*). The U and C nucleotides depicted in magenta form Watson-Crick base-pairing interactions with the nucleobase ligands.

Watson-Crick base pair between the aptamer and the corresponding purine ligand. Although the guanine- and adenine-specific aptamers that were examined by X-ray crystallography share only 59% sequence identity, they form nearly identical tertiary structures (45) in which discrimination between guanine and adenine is established exclusively by this single nucleotide.

Similarly, the removal or addition of single functional groups to these purines typically has strong adverse effects on ligand binding, which is consistent with the restrictions that appear to be placed on the ligand by the binding pocket. Other small riboswitch ligands such as lysine (88) and glycine (51) also are selectively bound by their corresponding riboswitches (see below), and any variations made to these compounds typically cause a loss of binding affinity. Therefore, complexity and the intricacy of the ligand-binding site of the purine riboswitches will likely be representative of the level of complexity present in the aptamers of other riboswitch classes.

FMN riboswitch. Another riboswitch that binds to FMN appears to be structurally more complex and more widespread in bacteria than the guanine and adenine riboswitches described above (**Figure 3*e***) (**Table 2**). The aptamer component of this riboswitch, originally termed the RFN element, was identified

and modeled by comparative sequence analysis (18). This element was present near genes that encode for riboflavin biosynthetic or transport genes of certain bacteria. As was the case with several other riboswitch classes, no protein factors were involved in the regulation of these genes. Various biochemical experiments detailed above demonstrated that the RNA undergoes structural changes when incubated in the presence of FMN (55, 105) and that the RNA selectively binds FMN in the absence of proteins (105).

Part of this increase in structural complexity might be required for the RNA to make productive interactions with the phosphate moiety of FMN. The riboswitch located in the 5′ UTR of the *ribDEATH* operon of *B. subtilis* binds nearly 100-fold more tightly than the related compound riboflavin, which differs from FMN by the absence of a single phosphate group (105). Other riboswitches also exhibit high affinity and exquisite selectivity on the basis of the presence of phosphate groups in the target ligands. Specifically, the removal of phosphate groups from TPP (102) and GlcN6P (107) leads to dramatic reductions in the affinity that their respective riboswitches exhibit for these derivative ligands.

Coenzyme B_{12} riboswitch. Coenzyme B_{12}- or adenosylcobalamin-specific riboswitches (**Figure 3*a***) are widespread throughout diverse bacterial species (**Table 2**), ranging from as few as 1 to as many as 13 representatives per genome (59, 93). This riboswitch is primarily used for regulation of genes expressing cobalamin synthesis proteins, but it may also regulate expression of porphyrin and cobalt transport genes. In isolated instances, this RNA regulates expression of proteins for glutamate fermentation, succinate fermentation, and coenzyme B_{12}-independent ribonucleotide reductases.

Only limited data have been acquired on the determinants for coenzyme B_{12} recognition, although a series of analogs has been used to reveal that the RNA uses molecular recognition determinants on the adenosyl moiety, the dimethylbenzimidazole moiety, and the cobinamide ring (60). Also, Scatchard analysis reveals that the riboswitch likely forms a single ligand-binding pocket (59). However, representatives of this class of riboswitches typically carry aptamer domains that are the largest of all natural aptamers identified thus far, suggesting that the structural demands are high for an RNA that is selectively recognizing this large and chemically sophisticated metabolite.

TPP riboswitch. The TPP class of riboswitches (**Figure 3*b***) appears to be the most widespread of all metabolite-sensing RNAs (**Table 2**). This motif has been identified in most of the major bacterial taxonomic groups, as well as for a single archaeal genus (*Thermoplasma*). In addition, RNAs that match the TPP riboswitch consensus sequence and structure have been identified in several plant and fungal species (87). These latter RNAs also bind to TPP when examined in a protein-free environment in vitro.

As mentioned above, TPP riboswitches are one of three classes that associate with ligands containing phosphate groups. Removal of the phosphate groups from TPP typically results in a dramatic decrease in binding affinity (102). Given the highly anionic character of RNA, one might assume that riboswitches would struggle to form structures that made productive interactions with phosphate groups. However, the selective recognition of phosphate groups by TPP riboswitches and by at least two other riboswitch classes indicates that RNA can readily address this challenge.

Comparative analyses of thiamine-induced and TPP-induced structural modulations (via in-line probing) revealed that the internal loop between P4 and P5 (**Figure 3**) is a candidate site for phosphate recognition (102). This conclusion is drawn from the observation that the P4/P5 internal loop is structurally altered only when the phosphate groups are included on the thiamine molecule. These structural changes could also be caused

P: pairing element

AEC: aminoethylcysteine, or thiosine

by the binding of phosphate at a distal site, and therefore further structural analysis is needed to determine whether this region of the RNA directly serves as the binding pocket for phosphate.

SAM riboswitch. An RNA motif called the S box (26) (**Figure 3*c***) had been identified to exist almost exclusively in Gram-positive bacteria, although few representatives are present in Gram-negative organisms (**Table 2**). In many instances, S box RNAs are represented multiple times within individual genomes in which they are located upstream of genes involved in sulfur metabolism and in the biosynthesis and transport of important metabolites such as cysteine, methionine and SAM. This widespread and well-conserved structural motif has since proven to serve as a SAM-responsive riboswitch (16, 52, 108).

The structure of this riboswitch has features indicative of a complex tertiary structure involving at least five base-paired elements (**Figure 3**). For example, there is phylogenetic evidence that the RNA forms a pseudoknot between the hairpin loop of P2a and the junction between P3 and P4. In addition, the bulge separating P2 from P2a appears to form a k-turn element. This latter structure is a common motif often used by complex-folded RNAs to introduce a sharp angle between helices (41, 106). Most likely, the k-turn allows the disparate elements of the proposed pseudoknot to colocalize and form a more compact aptamer structure. Moreover, the possible existence of a k-turn in the SAM riboswitch would be consistent with the proposal that there is a finite suite of structural motifs small enough to be comprehensively identified (57). Continued three-dimensional model building of riboswitch structures therefore promises to further illuminate the different types of common RNA building blocks and their relative frequency of use in structured RNAs.

The aptamer domain isolated from the SAM riboswitch of the *B. subtilis yitJ* gene exhibits strong affinity for SAM (K_d of ~4 nM) and considerable discrimination against closely related compounds such as *S*-adenosylhomosysteine (100-fold) and *S*-adenosylcysteine (nearly 10,000-fold) (108). As with most other riboswitches, the SAM riboswitch binds to its target with a 1:1 aptamer:ligand stoichiometry, and this binding is strongly dependent on the presence of Mg^{2+} ions.

These molecular recognition skills are extensively used in organisms that carry this riboswitch. For example, nearly half of all known riboswitch representatives present in *Bacillus anthracis* match the consensus sequence and structure of the SAM riboswitch class based on the S box motif. The remaining representatives are distributed among six other riboswitch classes. These observations indicate that SAM is a critical intracellular indicator for diverse aspects of sulfur-related metabolism and that the control of numerous important metabolic pathways in Gram-positive organisms is entrusted to a metabolite-sensing RNA.

Lysine riboswitch. The k-turn motif also appears to be present in another class of riboswitches that selectively binds to L-lysine (29, 88). Again, this structural element, as well as another common RNA motif called loop E, are likely important for this RNA to fold extended stem-bulge elements into a more compact RNA structure.

Although the affinity measured for a minimal lysine-binding aptamer from the *B. subtilis lysC* mRNA is modest (K_d = 1 μM), the large size and extensive nucleotide conservation of lysine aptamers might be needed to form a structure that excludes the many close analogs of lysine that naturally occur in most cells. For example, the aptamer discriminates between the L and D forms of lysine and rejects compounds such as ornithine, homolysine, and 5-hydroxylysine, all of which differ from L-lysine in the length or structure of the side chain (88).

The compound AEC is one analog of L-lysine that is bound by the riboswitch with reasonable affinity (~30 μM). This finding

is intriguing because AEC is an antimetabolite that is toxic to many bacteria and because AEC-resistant strains of *B. subtilis* (47, 96) carry mutations that disrupt function of the riboswitch. Although the incorporation of AEC into proteins during translation (as a replacement for lysine) might cause the antimicrobial effects that are seen with bacteria, that AEC also triggers riboswitch function suggests that at least part of the antimicrobial action of this compound could be due to the repression of genes involved in lysine biosynthesis. Similarly, other riboswitches might serve as novel drug targets for the development of new classes of antibiotics.

GlcN6P riboswitch. An RNA structure consisting primarily of four stem-loop regions (**Figure 3*f***) was discovered upstream of the *glmS* gene in many Gram-positive bacteria (**Table 2**). Because GlmS protein is an enzyme that catalyzes the conversion of fructose-6-phosphate to GlcN6P, the RNA was tested for its ability to directly interact with GlcN6P (107). Similar to other riboswitch classes, GlcN6P selectively associates with the *glmS* RNA. Genetic fusions between DNAs corresponding to the *glmS* control region and a *lacZ* reporter gene revealed that the *glmS* RNA element represses gene expression. Unexpectedly, the RNA undergoes rapid self-cleavage upon binding GlcN6P and serves as a novel metabolite-responsive ribozyme. Mutation of the *glmS-lacZ* reporter constructs that cause disruption of the ribozyme also cause derepression of gene expression. This unprecedented molecular mechanism for control of mRNA stability is discussed in greater detail below.

Glycine riboswitch. Two strikingly similar RNA structures (**Figure 3*g***), connected by a short linker region, were conserved upstream of glycine catabolic and efflux genes in a wide variety of bacteria (4). Typically, these domains appear in the 5′ UTR of genes encoding the glycine cleavage system, which breaks down glycine when excess of this amino acid is available. Biochemical tests using constructs derived from *B. subtilis* and *Vibrio cholerae* subsequently revealed that each conserved domain associates with a single glycine molecule (51). Remarkably, these aptamers exclude similarly structured chemicals such as alanine, serine, and the dipeptide glycylglycine.

Even more striking is the observation that, upon binding of glycine to one aptamer domain, the second domain exhibits a cooperative increase in affinity for glycine by at least 1000-fold. This cooperative mechanism allows expression of the gene to be controlled throughout its entire range by changes in metabolite concentration that are far less than that observed with other riboswitches. Because this particular class is usually used for control of glycine catabolism (**Table 2**), in which overexpression would have obvious deleterious consequences, organisms that make use of a cooperative riboswitch for glycine sensing might have a distinct evolutionary advantage. With such a digital riboswitch, cells can maximally express protein components of the glycine cleavage system when a modest excess of glycine is available, and likewise maximally suppress the expression of these proteins when glycine concentrations approach the point of limitation for its critical use in protein synthesis.

Orphan riboswitch candidates. Of the nine classes of riboswitches reported, evidence for the existence of seven of these elements had appeared in the literature before definitive proof of metabolite binding in the absence of proteins had been established. In nearly all of these instances, more classical genetics studies provided the first clues to the existence of a riboswitch. Perhaps these riboswitches have been uncovered first because they carry large aptamer domains and therefore make larger genetic targets. Also, many of these riboswitches are exceedingly widespread in bacteria (**Table 2**), and therefore the probability that they would have been encountered

by researchers over the years is increased. Regardless, it is highly unlikely that these examples constitute all (or even a majority) of the riboswitches that exist in bacteria. If true, then there might be a considerable number of smaller or rarer riboswitches that remain to be discovered.

To this end, a bioinformatics-based approach was used to seek out additional evolutionarily conserved regulatory RNAs within *B. subtilis* (4). This approach makes use of computer algorithms that search for sequence similarity among intergenic regions and couples this information with predictions of the functions of the products derived from neighboring genes. This approach successfully rediscovered most known riboswitch classes. More importantly, eight new candidate regulatory RNAs were identified (**Table 3**). As with the known riboswitches, these new motifs exhibit significant evolutionary conservation, are present in diverse bacteria, are associated with similar classes of genes, and are adjacent to structures that resemble expression platforms (see below). Two of these candidates, *glmS* and *gcvT*, have subsequently been proven to function as riboswitches for GlcN6P (107) and glycine, respectively (51). These initial findings suggest that many other classes of metabolite-sensing RNAs await discovery and analysis.

Riboswitch Genetic Control Mechanisms

Although feedback repression of genes is by far the dominant mode of operation, several examples of riboswitches activate gene expression as metabolite concentrations rise (49, 51, 94). It is not due to any limitations in RNA structure or deficiencies in function that have caused this asymmetry in the direction of gene control that riboswitches manage. Rather, there is a frequent need for bacteria to sense excesses of a wide diversity of metabolites and to deactivate the expression of biosynthetic or transport genes that would otherwise generate more of these compounds. Less common is the need for a cell to sense these same compounds and activate gene expression as these metabolites accumulate.

Two predominant mechanisms of genetic regulation are used by riboswitches: control of transcription termination and control of translation initiation (**Figure 5**). However, both processes can exploit similar changes in RNA folding that involve Watson-Crick

Table 3 Orphan riboswitch candidates[a]

Orphan class (downstream gene in *Bacillus subtilis*)	Metabolite ligand	Associated genes
1 (*glmS*)	GlcN6P	*glmS* (GlcN6P synthase)
2 (*gcvTHP*)	Glycine	Glycine catabolism, metabolite transport
3 (*ykoK*)	?	Divalent metal transport systems
4 (*yybP/ykoY*)	?	Cation transport systems
5 (*ykkC/yxkD*)	?	Nitrate/sulfonate/bicarbonate transport systems
6 (*ydaO/yuaA*)	?	Amino acid transport, K^+ transport, metalloendopeptidases, cell-wall-associated hydrolases
7 (*ykvJKLM*)	?	?
8 (*ylbHI*)	?	?

[a]Rows indicate eight classes of conserved RNAs discovered through computer-assisted search methods (4). Column 1 indicates the *B. subtilis* genes associated with each RNA class. Column 2 indicates the effector molecule, if known. Column 3 indicates the predicted classes of associated genes.

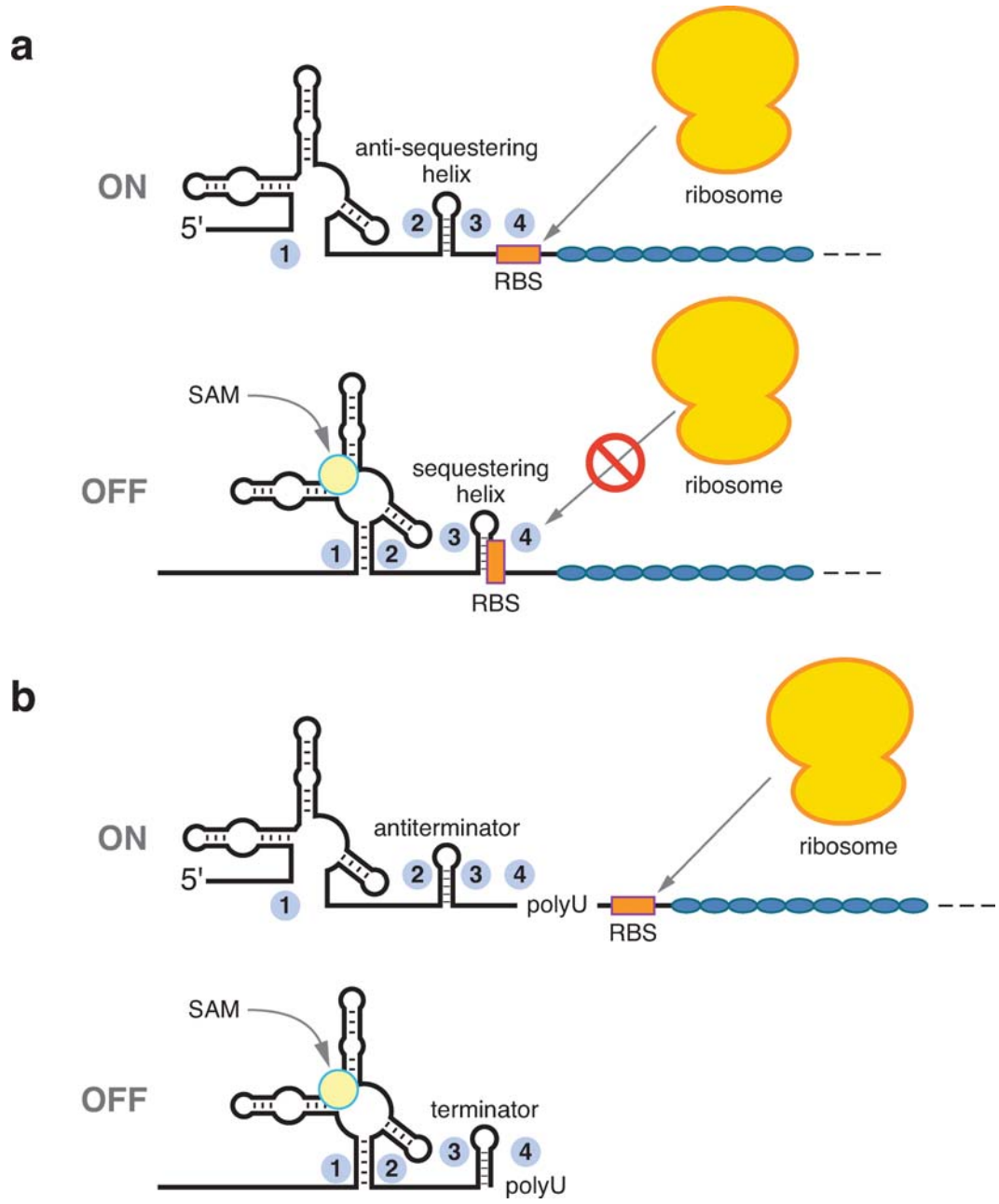

Figure 5

Common structural interchanges used by riboswitches to regulate transcription termination and translation initiation. (*a*) Numbers 1 through 4 represent different sequence elements. In the absence of the metabolite ligand a helix is formed between 2 and 3, allowing access for translation initiation. Upon metabolite binding, helices are created between 1 and 2 and between 3 and 4. The latter hairpin occludes ribosome association, reducing translation. (*b*) Same as in (*a*), except the 3 + 4 hairpin creates an intrinsic transcription terminator, which aborts transcription elongation, thereby reducing expression of downstream genes.

base-pairing. Natural aptamers, like their engineered counterparts (35), can undergo substantial structural reorganization upon ligand binding. When bound to ligand, each aptamer typically forms a stable assembly of secondary-structure and tertiary-structure elements. In the absence of ligand, some of the nucleotides of the aptamer become structurally heterogeneous and are available to form alternative structures with portions of the expression platform.

This metabolite-dependent folding diversity can be harnessed in numerous ways to control the expression of adjacent coding regions. However, note that the general mechanisms discussed herein are oversimplifications of the detailed molecular processes required for riboswitch function. These mechanisms will likely require precise coordination of many molecular processes such as the kinetics of RNA folding, transcription elongation, metabolite binding, helical formation, ribosome binding, and mRNA processing.

Transcription Attenuation

Intrinsic transcription terminators are GC-rich stem-loops followed by a polyuridyl tract of five to nine nucleotides (33, 110). These structural elements destabilize elongation complexes, resulting in cessation of transcription. There is considerable precedence for regulation of transcription termination, termed attenuation (20, 34). The typical attenuation mechanism involves controlling formation of a terminator hairpin by optional formation of a mutually exclusive antiterminator RNA helix created from the left half of the terminator and an upstream sequence (**Figure 5**).

RNA elements that exhibit the structural features of known intrinsic terminator stems reside downstream of many riboswitch aptamers (68, 69, 92, 93, 108). These representative riboswitches are predicted to function through transcription attenuation, based upon the identification of aptamer, terminator, and antiterminator structures. This speculation has been confirmed for several riboswitch representatives by more detailed genetic studies. For example, transcriptional fusions between riboswitches and *lacZ* reporter genes have been monitored under various ligand concentrations. There is a dramatic change in expression with increasing ligand concentrations, ranging between ~7-fold and ~1,200-fold for purine-sending and SAM-sensing RNAs (48, 52, 108). Mutations that disrupt either the aptamer or terminator structures ablate this regulatory response in vivo, whereas compensatory mutations restoring structural integrity also re-establish regulation.

Additional support for this metabolite-induced attenuation mechanism comes from biochemical studies. For example, data from in-line probing assays are consistent with the formation of terminator hairpins (105). Furthermore, in vitro transcription of DNA templates for riboswitches carrying intrinsic transcription terminators yields truncated RNA transcripts when the corresponding ligand is present (16, 29, 51, 52, 55, 88, 105, 108). For the FMN riboswitch from the *B. subtilis ribD* gene, the ligand-induced transcription termination site was mapped to the final two to three U residues of the putative terminator stem's polyuridine tract (105).

Metabolite-induced transcription termination products are not produced if either the aptamer or the terminator stem is rendered defective by mutation. Also, accessory protein factors are not required for metabolite binding or for transcription termination, although it still remains to be determined if specialized protein factors might be important for influencing aspects of riboswitch regulation in vivo. Typically, concentrations of metabolite in the low micromolar range are required to induce half-maximal termination of transcription in vitro. This is in approximate agreement with intracellular metabolite concentrations that have been reported in literature. However, note that these latter values vary immensely depending on the methodology used for measurement. In addition,

published reports for intracellular metabolites are often a closer reflection of the total amount present, as opposed to the concentration of a given metabolite that might be free to interact with a riboswitch. Therefore, the concentrations of metabolites determined to yield half-maximal termination of riboswitch transcripts in vitro are not likely to produce accurate estimates of the concentrations of metabolites that are present inside cells.

A further complication is that the concentrations of metabolite needed to approach saturation of riboswitch aptamers, measured by in-line probing or equilibrium dialysis assays, are typically several orders of magnitude lower than that required for half-maximal termination of transcription in vitro. One possible explanation for this discrepancy is that some riboswitches do not reach equilibrium with intracellular ligands before genetic control decisions are made. Recent results indicate that, at least for one representative of an FMN-binding riboswitch, the speed of aptamer-ligand equilibration is far too slow to reach equilibrium given the time needed for RNA polymerase to reach the transcription terminator. Thus, concentrations of FMN that are much higher than the measured K_d for the aptamer-ligand interaction are necessary to trigger transcription termination in vitro (101).

Note that there may be variations to the attenuation mechanism described herein. These exceptions may use transcription termination signals different from the classical intrinsic terminator. For example, the *E. coli thiC* riboswitch exerts control over translation initiation and transcription, even though it does not contain a recognizable intrinsic terminator (102).

Control of Translation Initiation

Within the expression platform of some riboswitches is a helical element of varying length and character in the region of the SD sequence (59, 68, 69, 92, 93). This structure is called a sequestering helix because it appears capable of base-pairing to the SD element. Alternatively, a portion of the helix (usually the left half) is capable of pairing with an oligonucleotide tract within the aptamer domain (usually overlapping the right half of the P1 helix), thus forming an anti-sequestering helix. This architecture suggests that these RNAs regulate the efficiency of translation initiation by occluding the RBS under certain conditions.

SD: Shine-Dalgarno

There is precedence for regulation of translation initiation by SD-sequestering helices. For example, association of tryptophan-bound TRAP to the 5′ UTRs of *B. subtilis trpEDCFBA*, *pabA*, and *trpP* mRNAs promotes formation of a SD-sequestering helix for each transcript (3, 109). Several lines of investigation have provided support for SD-sequestration by prokaryotic riboswitches. In-line probing of riboswitches with SD-sequestering helices revealed that the SD region is structurally altered in response to ligand binding. Specifically, upon association of TPP to *E. coli thiM* transcripts, the SD sequence adopts a more highly structured state (102), consistent with formation of a sequestering helix (**Figure 6**). Similar ordering of the SD region has been observed for the *B. subtilis ypaA* transcript upon association of FMN (105). More direct evidence for riboswitch-mediated inhibition of translation initiation has been obtained via analysis of the *E. coli btuB* gene. The presence of coenzyme B_{12} inhibits binding of 30S ribosomal subunits to mRNAs in vitro (64).

Several variations to this common sequestering and anti-sequestering hairpin architecture have been identified (68, 92, 108). In these cases, the aptamer resides immediately upstream of the coding portion of the mRNA and therefore the SD sequence serves an integral role in forming the ligand-binding structure. Presumably, metabolite binding directly sequesters the SD element and the resulting increase in overall structure leads to diminished access for the ribosome.

Interestingly, the biological distribution of riboswitches that employ transcription

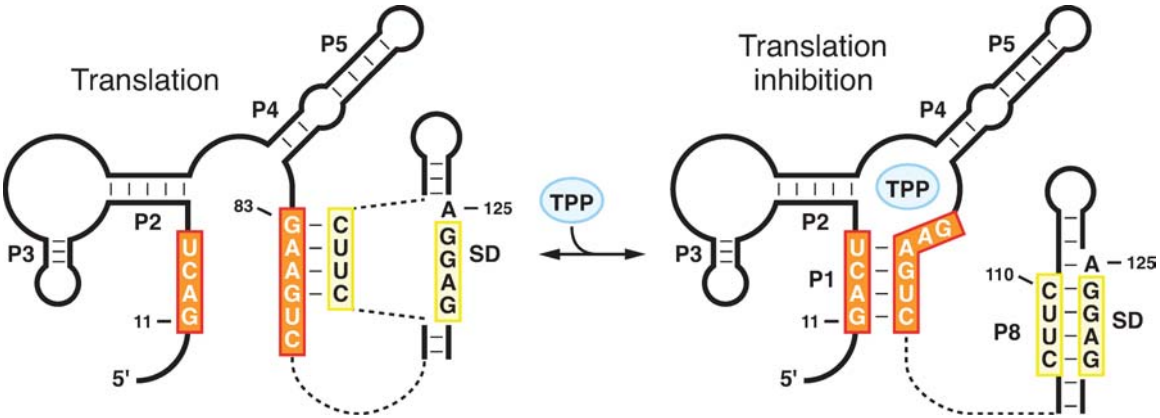

Figure 6

Proposed mechanism for the regulation of *E. coli thiM* translation initiation by TPP binding. In the presence of TPP the P1 helix is formed, thereby allowing for formation of the P8 hairpin (102). The latter structure incorporates the SD site, thereby reducing translation efficiency. In the absence of TPP, an alternative helix is formed between the right half of P1 and the left half of P8, thereby increasing availability of the SD for ribosome binding.

attenuation versus translation initiation mechanisms is nonrandom. Gram-positive organisms preferentially use transcription attenuation mechanisms, and Gram-negative organisms employ SD-sequestering helices (94). For example, of the FMN-dependent riboswitches recently identified from Gram-positive organisms (92), ~83% are predicted to use an attenuation mechanism. Similarly, ~80% of Gram-negative sequences employ SD-sequestering helices. It is not clear why certain bacterial lineages retain this bias in expression platform mechanisms.

Control of Antisense RNA

Bioinformatics analyses have identified a small subset of riboswitch RNAs that appear to regulate production of antisense transcripts. For example, the *Clostridium acetobutylicum ubiG-yrhAB* operon retains a T box regulatory RNA within the 5′ UTR that is predicted to respond to $tRNA^{Cys}$ (70). However, an RNA motif that conforms to the consensus sequence and structure of a SAM-dependent riboswitch class is located immediately downstream of the operon, where it appears to be transcribed in the opposite direction relative to the operon. This architecture suggests that increasing concentrations of SAM might trigger the production of an antisense transcript for the *ubiG-yrhAB* operon, although experimental evidence for this mechanism is lacking.

Control of mRNA Processing

As described above, a bioinformatics approach revealed the existence of a conserved RNA structure located upstream from the *glmS* gene in Gram-positive bacteria (4). GlmS catalyzes the conversion of fructose-6-phosphate to GlcN6P, a metabolic precursor for cell wall biosynthesis. In-line probing revealed that GlcN6P interacts with the *glmS* leader RNA and appears to subtly modulate its structure (107). However, these tests also revealed that binding of GlcN6P to the RNA stimulates self-processing activity at the 5′ end of the conserved region in a magnesium-dependent manner. In the presence of excess GlcN6P, this self-cleavage event occurs at a rate that is 10^6-fold higher than the background rate for spontaneous RNA cleavage (46, 107), indicating the RNA is a novel, metabolite-sensing natural ribozyme. Comparative analysis of wild-type and mutant *glmS-lacZ* fusions

indicate that the ribozyme represses *glmS* expression. However, the molecular details for this unprecedented regulatory mechanism remain to be uncovered. Only four small self-cleaving ribozymes had been previously discovered (14). It remains to be determined whether the *glmS* regulatory RNA is an anomaly or is representative of a more widespread method of genetic control.

Examples of RNA elements that correspond in sequence and structure to aptamers of bacterial TPP riboswitches have been identified in several eukaryote genomes by comparative sequence analysis (87). These are presumably homologs of bacterial TPP riboswitches that reside in the 5′ and 3′ UTRs of TPP metabolic genes in various fungi and plant species. Versions of these RNAs, synthesized in vitro, bind TPP with high affinity, which provides some evidence that the role of these RNAs is metabolite sensing. Additionally, these RNAs are located adjacent to splice site junctions, which suggests that metabolite binding might exert control over mRNA processing. Deletion of the riboswitch-containing region indeed results in a loss of thiamine-responsive splicing, whereas excess thiamine results in diminished splicing activity in vivo (44). Taken together, these findings suggest that some eukaryotic cells employ riboswitches to control gene expression and that the control of RNA splicing might be a preferred mechanism for riboswitch function in higher organisms.

RIBOSWITCH APPLICATIONS

Molecular Engineering of RNA-Based Genetic Elements

Well-characterized regulatory mechanisms, such as the system that regulates expression of the *lac* operon (40), have been adapted as common laboratory tools for controllable expression of gene constructs. Similarly, the harnessing of natural or engineered riboswitches would expand the collection of gene control systems that researchers could integrate with their transgenic constructs. In some ways, RNA-based gene control systems might provide a simpler and more versatile architecture for expanding gene control capabilities than that provided by protein-based systems. The gene control systems most widely used by genetic engineers typically involve a regulatory protein whose function is modified by a chemical effector that must be permeable to the cellular host. Their use is also limited to hosts that produce appropriate levels of the regulatory protein either naturally or from an accessory gene that is inserted along with the desired transgene. In addition, each regulatory protein is responsive to a single chemical, and therefore a variety of protein factors and genetic elements are needed if the user desires to control several genes differentially. The use of natural riboswitches for user-defined gene control applications is complicated because they typically sense fundamental (and usually indispensable) metabolites whose concentrations might be difficult to control at will. However, RNA engineers might be able to coopt the mechanisms used by natural riboswitches to create a collection of RNA elements that respond to a diversity of chemical effectors.

Considerable progress already has been made by those who seek to validate this type of technology. For example, aptamers generated through in vitro evolution have been used to regulate gene expression by placing them into the 5′ UTR of eukaryotic mRNAs (22, 78, 90, 99). One possible mechanism for how ligand binding reduces gene expression is that complex formation might disrupt scanning by the translational apparatus. Similar approaches are also being used to control the expression of bacterial genes. For example, a theophylline-sensing RNA switch has been engineered to reduce access to an adjacent SD sequence when theophylline is added (78). Similarly, the theophylline aptamer has also been used to activate gene expression in *E. coli* (13).

In addition, allosteric ribozymes that also could be used to modulate gene expression

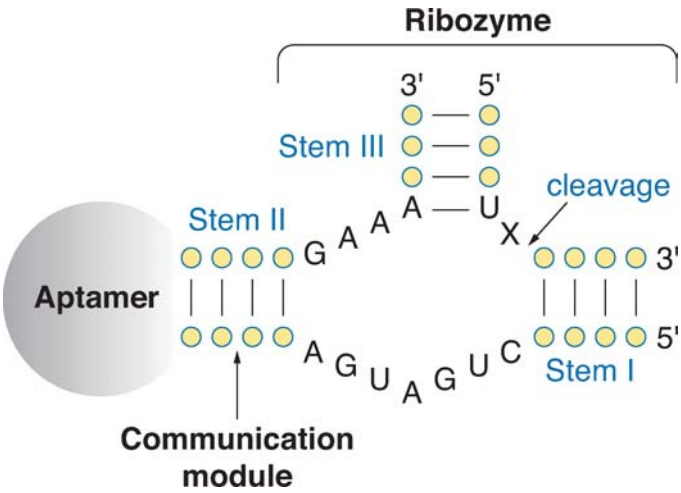

Figure 7

Common organization of engineered allosteric hammerhead ribozymes. The function of the self-cleaving ribozyme is regulated by ligand binding to the aptamer by allosteric changes in RNA structure facilitated by the communication module. Secondary structures and conserved nucleotides required for hammerhead ribozyme activity are depicted.

by controlling RNA processing have been created (7, 80). Specifically, small self-cleaving ribozymes, such as the hammerhead ribozyme, have been generated by in vitro evolution to respond allosterically to exogenous ligand effectors (72). These RNAs typically are modular in construction (**Figure 7**) and carry an aptamer for ligand sensing, a reporter domain (often a ribozyme), and a bridge element (communication module) that functionally links the aptamer and reporter domains. Some of these features are similar to those found in natural riboswitches.

Engineered allosteric RNAs have been used as biosensor elements in vitro (36, 76) and are beginning to be adapted for in vivo applications. Similar designs have been used in bacterial cells, including the use of aptamers to control gene expression by modulating the activity of self-splicing ribozymes (89). More recently, chemical-mediated inhibition of cleavage by the hammerhead ribozyme has been exploited for genetic control of a mammalian reporter gene (111). Therefore, elucidating the underlying principles of riboswitch-mediated regulation may lead to improvements in the molecular engineering of designer riboswitches.

Riboswitches as Drug Targets

Riboswitches control many fundamental genes and metabolic pathways; therefore, they are potential targets for antimicrobial agents. Indeed, a hypothetical drug that bound to the SAM riboswitch and forcibly reduced expression of associated genes would simultaneously inhibit expression of genes involved in the biosynthesis or transport of numerous compounds such as methionine, cysteine, and SAM. Historically, considerable attention has been focused on RNAs as targets for drug development, motivated in part by the success of ribosome-targeting antimicrobial agents (91). In contrast to efforts that search for chemicals that fortuitously bind to a given RNA target, riboswitches offer an advantage in that they naturally bind to small molecules. Presumably, drug-like compounds that compete with metabolite binding could be identified and used to alter the expression of essential metabolic genes of bacterial pathogens.

Such efforts to develop riboswitches as targets for drug action are still in their infancy. However, precedence for this mechanism of antimicrobial drug action has been reported recently. The mode of action of the antimicrobial compound AEC appears to be due, at least in part, to this compound directly binding to lysine riboswitches and downregulating the expression of lysine biosynthesis genes (88). In addition, three-dimensional structural data, such as that recently acquired for purine riboswitches (5, 77), promise to establish a foundation for structure-based drug design of riboswitch-targeted chemicals. Similarly, the highly conserved antiterminator structure of T box RNAs, which is required for discrimination of amino acid charged versus uncharged tRNAs, has been structurally characterized via nuclear magnetic resonance (19). This structure is common to all T box RNAs

and therefore also represents a potential subject for drug development.

CONCLUSIONS

The roles of noncoding RNAs and untranslated portions of mRNAs in gene control processes appear to be far more extensive than originally believed. Some of these mechanisms primarily involve the expression of small RNAs that selectively interact with mRNAs and cause changes in gene expression as a result. These RNA gene control factors could easily emerge through typical evolutionary processes, given that fortuitous base-pairing interactions between mRNAs and other cellular RNAs are likely to influence gene expression. RNA seems particularly adept at making new and changing existing gene control factors, and so it is not surprising that noncoding RNAs are widespread in many organisms and that numerous new examples are being discovered at a rapid pace.

In contrast to most noncoding RNAs, riboswitches must form binding pockets for specific metabolites and carry accessory elements that allosterically respond to ligand binding. These functional requirements place greater demands on RNA structure, which in turn suggests that riboswitches carry sequence and structural elements that remain highly conserved through evolution. Unlike many noncoding RNAs, whose function as binding partners with other RNAs are especially suited to a nucleic acid polymer, riboswitches are competing with proteins as a medium for forming metabolite-binding pockets. Although natural and engineered aptamers have the ability to bind target compounds with high affinity and specificity, protein is most likely the superior medium for forming such structures. In other words, a protein will most likely be tapped by a cell if a new metabolite-sensing factor is needed for gene control purposes.

If this is true, then perhaps many of the riboswitches we see in modern cells are the distant relatives of metabolite-binding RNAs that might have been present before the evolutionary emergence of proteins. Given that many riboswitch ligands appear to be ancient in origin (100), this possibility cannot easily be ruled out. Also, these ligands have not changed over the last several billion years of evolution, and so the RNAs that bind them also might have changed little despite the enormous amount of evolution that has taken place all around them.

Regardless of the evolutionary history of riboswitches, it is now clear that they serve key roles in maintaining genetic and biochemical homeostasis in many organisms. In addition, the recent biochemical and structural analyses of riboswitches have expanded our understanding of how RNA polymers can fold into functional shapes and achieve complex biochemical tasks. Further investigation of the biological distribution of riboswitches and their molecular mechanisms undoubtedly will help elucidate the roles of RNA in bacterial genetic circuitry and allow us to build a more complete picture of the structures and functions of complex RNAs.

ACKNOWLEDGMENTS

We thank members of the Breaker laboratory for helpful discussions. Riboswitch research in the Breaker laboratory is supported by grants from NIH, NSF, and DARPA.

LITERATURE CITED

1. Altuvia S, Kornitzer D, Teff D, Oppenheim AB. 1989. Alternative mRNA structures of the cIII gene of bacteriophage λ determine the rate of its translation initiation. *J. Mol. Biol.* 210:265–80

2. Argaman I, Altuvia S. 2000. *fhlA* repression by OxyS RNA: Kissing complex formation at two sites results in a stable antisense-target RNA complex. *J. Mol. Biol.* 300:1101–12
3. Babitzke P. 2004. Regulation of transcription attenuation and translation initiation by allosteric control of an RNA-binding protein: the *Bacillus subtilis* TRAP protein. *Curr. Opin. Microbiol.* 7:132–39
4. Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, et al. 2004. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc. Natl. Acad. Sci. USA* 101:6421–26
5. Batey RT, Gilbert SD, Montange RK. 2004. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* 432:411–15
6. Boy E, Borne F, Patte JC. 1979. Isolation and identification of mutants constitutive for aspartokinase III synthesis in *Escherichia coli* K12. *Biochemie* 61:1151–60
7. Breaker RR. 2002. Engineered allosteric ribozymes as biosensor components. *Curr. Opin. Biotechnol.* 13:31–39
8. Breaker RR. 2004. Natural and engineered nucleic acids as tools to explore biology. *Nature* 432:838–45
9. Browning DF, Busby SJ. 2004. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* 2:57–65
10. Christiansen LC, Schou S, Nygaard P, Saxild HH. 1997. Xanthine metabolism in *Bacillus subtilis*: characterization of the *xpt-pbuX* operon and evidence for purine- and nitrogen-controlled expression of genes involved in xanthine salvage and catabolism. *J. Bacteriol.* 179:2540–50
11. Condon C. 2003. RNA processing and degradation in *Bacillus subtilis*. *Microbiol. Mol. Biol. Rev.* 67:157–74
12. Copeland PR. 2003. Regulation of gene expression by stop codon recoding: selenocysteine. *Gene* 312:17–25
13. Desai SK, Gallivan JP. 2004. Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J. Am. Chem. Soc.* 126:13247–54
14. Doherty EA, Doudna JA. 2001. Ribozyme structures and mechanisms. *Annu. Rev. Biophys. Biomol. Struct.* 30:457–75
15. Ebbole DJ, Zalkin H. 1987. Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* encoding nine enzymes for de novo purine nucleotide synthesis. *J. Biol. Chem.* 262:8274–87
16. Epshtein V, Mironov AS, Nudler E. 2003. The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc. Natl. Acad. Sci. USA* 100:5052–56
17. Escalante-Semerena JC, Roth JR. 1987. Regulation of cobalamin biosynthetic operons in *Salmonella typhimurium*. *J. Bacteriol.* 169:2251–58
18. Gelfand MS, Mironov AA, Jomantas J, Kozlov YI, Perumov DA. 1999. A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* 15:439–42
19. Gerdeman MS, Henkin TM, Hines JV. 2003. Solution structure of the *Bacillus subtilis* T-box antiterminator RNA: seven nucleotide bulge characterized by stacking and flexibility. *J. Mol. Biol.* 326:189–201
20. Gollnick P, Babitzke P. 2002. Transcription attenuation. *Biochim. Biophys. Acta.* 1577:240–50
21. Gottesman S. 2004. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu. Rev. Microbiol.* 58:303–28

22. Grate D, Wilson C. 2001. Inducible regulation of the *S. cerevisiae* cell cycle mediated by an RNA aptamer-ligand complex. *Bioorg. Med. Chem.* 9:2565–70
23. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. 2003. Rfam: an RNA family database. *Nucleic Acids Res.* 31:439–41
24. Grundy FJ, Collins JA, Rollins SM, Henkin TM. 2000. tRNA determinants for transcription antitermination of the *Bacillus subtilis tyrS* gene. *RNA* 6:1131–41
25. Grundy FJ, Henkin TM. 1993. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* 74:475–82
26. Grundy FJ, Henkin TM. 1998. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Mol. Microbiol.* 30:737–49
27. Grundy FJ, Henkin TM. 2003. The T box and S box transcription termination control systems. *Front. Biosci.* 8:d20–31
28. Grundy FJ, Hodil SE, Rollins SM, Henkin TM. 1997. Specificity of tRNA-mRNA interactions in *Bacillus subtilis tyrS* antitermination. *J. Bacteriol.* 179:2587–94
29. Grundy FJ, Lehman SC, Henkin TM. 2003. The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc. Natl. Acad. Sci. USA* 100:12057–62
30. Grundy FJ, Rollins SM, Henkin TM. 1994. Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the *Bacillus subtilis tyrS* gene: a new role for the discriminator base. *J. Bacteriol.* 176:4518–26
31. Grundy FJ, Winkler WC, Henkin TM. 2002. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. *Proc. Natl. Acad. Sci. USA* 99:11121–26
32. Guillier M, Allemand F, Raibaud S, Dardel F, Springer M, Chiaruttini C. 2002. Translational feedback regulation of the gene for L35 in *Escherichia coli* requires binding of ribosomal protein L20 to two sites in its leader mRNA: a possible case of ribosomal RNA-messenger RNA molecular mimicry. *RNA* 8:878–89
33. Gusarov I, Nudler E. 1999. The mechanism of intrinsic transcription termination. *Mol. Cell* 3:495–504
34. Henkin TM, Yanofsky C. 2002. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* 24:700–7
35. Hermann T, Patel DJ. 2000. Adaptive recognition by nucleic acids. *Science* 287:820–25
36. Hesselberth JR, Robertson MP, Knudsen SM, Ellington AD. 2003. Simultaneous detection of diverse analytes with an aptazyme ligase array. *Anal. Biochem.* 312:106–12
37. Houman F, Diaz-Torres MR, Wright A. 1990. Transcriptional antitermination in the *bgl* operon of *E. coli* is modulated by a specific RNA binding protein. *Cell* 62:1153–63
38. Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. 2002. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* 110:551–61
39. Joyce GF. 2002. The antiquity of RNA-based evolution. *Nature* 418:214–21
40. Kercher MA, Lu P, Lewis M. 1997. *lac* repressor-operator complex. *Curr. Opin. Struct. Biol.* 7:76–85
41. Klein DJ, Schmeing TM, Moore PB, Steitz TA. 2001. The kink-turn: a new RNA secondary structure motif. *EMBO J.* 20:4214–21
42. Kochhar S, Paulus H. 1992. Lysine-induced premature transcription termination in the *lys C* operon of *Bacillus subtilis*. *Microbiology* 142:1635–39

43. Kreneva RA, Perumov DA. 1990. Genetic mapping of regulatory mutations of *Bacillus subtilis* riboflavin operon. *Mol. Gen. Genet.* 222:467–69
44. Kubodera T, Watanabe M, Yoshiuchi K, Yamashita N, Nishimura A, et al. 2003. Thiamine-regulated gene expression of *Aspergillus oryzae thiA* requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR. *FEBS Lett.* 555:516–20
45. Lescoute A, Westhof E. 2005. Riboswitch structures: Purine ligands replace tertiary contacts. *Chem. Biol.* 12:10–13
46. Li Y, Breaker RR. 1999. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2′-hydroxyl group. *J. Am. Chem. Soc.* 121:5364–72
47. Lu Y, Shevtchenko TN, Paulus H. 1992. Fine structure mapping of *cis*-acting control sites in the *lys C* operon of *Bacillus subtilis*. *FEMS Microbiol. Lett.* 92:23–27
48. Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* 113:577–86
49. Mandal M, Breaker RR. 2004. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat. Struct. Mol. Biol.* 11:29–35
50. Mandal M, Breaker RR. 2004. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* 5:451–63
51. Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, et al. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* 306:275–79
52. McDaniel BA, Grundy FJ, Artsimovitch I, Henkin TM. 2003. Transcription termination control of the S box system: direct measurement of *S*-adenosylmethionine by the leader RNA. *Proc. Natl. Acad. Sci. USA* 100:3083–88
53. Merianos HJ, Wang J, Moore PB. 2004. The structure of a ribosomal protein S8/*spc* operon mRNA complex. *RNA* 10:954–64
54. Miranda-Rios J, Navarro M, Soberon M. 2001. A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc. Natl. Acad. Sci. USA* 98:9736–41
55. Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, et al. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 111:747–56
56. Mooney RA, Artsimovitch I, Landick R. 1998. Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *J. Bacteriol.* 180:3265–75
57. Moore PB. 1999. Structural motifs in RNA. *Annu. Rev. Biochem.* 68:287–300
58. Muto Y, Oubridge C, Nagai K. 2000. RNA-binding proteins: TRAPping RNA bases. *Curr. Biol.* 10:R19–21
59. Nahvi A, Barrick JE, Breaker RR. 2004. Coenzyme B_{12} riboswitches are widespread genetic control elements in prokaryotes. *Nucleic Acids Res.* 32:143–50
60. Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR. 2002. Genetic control by a metabolite binding mRNA. *Chem. Biol.* 9:1043
61. Narberhaus F. 2002. mRNA-mediated detection of environmental conditions. *Arch. Microbiol.* 178:404–10
62. Nimjee SM, Rusconi CP, Sullenger BA. 2005. APTAMERS: an emerging class of therapeutics. *Annu. Rev. Med.* 56:555–83
63. Noeske J, Richter C, Grundl MA, Nasiri HR, Schwalbe H, Wohnert J. 2005. An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs. *Proc. Natl. Acad. Sci. USA* 102:1372–77

64. Nou X, Kadner RJ. 2000. Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc. Natl. Acad. Sci. USA* 97:7190–95
65. Nudler E, Mironov AS. 2004. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* 29:11–17
66. Ptashne M, Gann A. 2002. *Genes and Signals*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
67. Putzer H, Condon C, Brechemier-Baey D, Brito R, Grunberg-Manago M. 2002. Transfer RNA-mediated antitermination in vitro. *Nucleic Acids Res.* 30:3026–33
68. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2002. Comparative genomics of thiamin biosynthesis in prokaryotes. New genes and regulatory mechanisms. *J. Biol. Chem.* 277:48949–59
69. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2003. Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Res.* 31:6748–57
70. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2004. Comparative genomics of the methionine metabolism in gram-positive bacteria: a variety of regulatory systems. *Nucleic Acids Res.* 32:3340–53
71. Romby P, Springer M. 2003. Bacterial translational control at atomic resolution. *Trends Genet.* 19:155–61
72. Roth A, Breaker RR. 2004. Selection in vitro of allosteric ribozymes. *Methods Mol. Biol.* 252:145–64
73. Schaak JE, Yakhnin H, Bevilacqua PC, Babitzke P. 2003. A Mg^{2+}-dependent RNA tertiary structure forms in the *Bacillus subtilis trp* operon leader transcript and appears to interfere with *trpE* translation control by inhibiting TRAP binding. *J. Mol. Biol.* 332:555–74
74. Schilling O, Langbein I, Muller M, Schmalisch MH, Stulke J. 2004. A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity. *Nucleic Acids Res.* 32:2853–64
75. Schlax PJ, Worhunsky DJ. 2003. Translational repression mechanisms in prokaryotes. *Mol. Microbiol.* 48:1157–69
76. Seetharaman S, Zivarts M, Sudarsan N, Breaker RR. 2001. Immobilized RNA switches for the analysis of complex chemical and biological mixtures. *Nat. Biotechnol.* 19:336–41
77. Serganov A, Yuan YR, Pikovskaya O, Polonskaia A, Malinina L, et al. 2004. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* 11:1729–41
78. Seuss B, Hanson S, Berens C, Fink B, Schroeder R, Hillen W. 2003. Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res.* 31:1853–58
79. Sonenshein AL. 2002. Appendix 2. In Bacillus subtilis *and Its Closest Relatives: From Genes to Cells*, ed. AL Sonenshein, JA Hoch, R Losick, pp. 602–6. Washington, DC: ASM Press
80. Soukup GA, Breaker RR. 1999. Relationship between internucleotide linkage geometry and the stability of RNA. *RNA* 5:1308–25
81. Soukup GA, DeRose EC, Koizumi M, Breaker RR. 2001. Generating new ligand-binding RNAs by affinity maturation and disintegration of allosteric ribozymes. *RNA* 7:524–36
82. Springer M, Portier C, Grunberg-Manago M.1997. RNA mimicry in the translational apparatus. In *RNA Structure and Function*, ed. RW Simons, M Grunberg-Manago, pp. 377–414. Cold Spring Harbor, NY: Cold Spring Harbor Lab. Press

83. Steltzl U, Zengel JM, Tovbina M, Walker M, Nierhaus KH, et al. 2003. RNA-structural mimicry in *Escherichia coli* ribosomal protein L4-dependent regulation of the *S10* operon. *J. Biol. Chem.* 278:28237–45
84. Stormo GD, Ji Y. 2001. Do mRNAs act as direct sensors of small molecules to control their expression? *Proc. Natl. Acad. Sci. USA* 98:9465–67
85. Storz G, Opdyke JA, Zhang A. 2004. Controlling mRNA stability and translation with small, noncoding RNAs. *Curr. Opin. Microbiol.* 7:140–44
86. Stulke J. 2002. Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures. *Arch. Microbiol.* 177:433–40
87. Sudarsan N, Barrick JE, Breaker RR. 2003. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* 9:644–47
88. Sudarsan N, Wickiser JK, Nakamura S, Ebert MS, Breaker RR. 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.* 17:2688–97
89. Thompson KM, Syrett HA, Knudsen SM, Ellington AD. 2002. Group I aptazymes as genetic regulatory switches. *BMC Biotechnol.* 2:21
90. Toulme JJ, Di Primo C, Boucard D. 2004. Regulating eukaryotic gene expression with aptamers. *FEBS Lett.* 567:55–62
91. Vicens Q, Westhof E. 2003. RNA as a drug target: the case of aminoglycosides. *Chembiochem* 4:1018–23
92. Vitreschak AG, Rodionov DA, Mironov AA, Gelfand MS. 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* 30:3141–51
93. Vitreschak AG, Rodionov DA, Mironov A A, Gelfand MS. 2003. Regulation of the vitamin B_{12} metabolism and transport in bacteria by a conserved RNA structural element. *RNA* 9:1084–97
94. Vitreschak AG, Rodionov DA, Mironov AA, Gelfand MS. 2004. Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet.* 20:44–50
95. Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jager JG, et al. 2003. RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* 31:6435–43
96. Vold B, Szulmajster J, Carbone A. 1975. Regulation of dihydrodipicolinate synthase and aspartate kinase in *Bacillus subtilis*. *J. Bacteriol.* 121:970–74
97. Wei BY, Bradbeer C, Kadner RJ. 1992. Conserved structural and regulatory regions in the *Salmonella typhimurium btuB* gene for the outer membrane vitamin B_{12} transport protein. *Res. Microbiol.* 143:459–66
98. Weng M, Nagy P, Zalkin H. 1995. Identification of the *Bacillus subtilis pur* operon repressor. *Proc. Natl. Acad. Sci. USA* 92:7455–59
99. Werstuck G, Green MR. 1998. Controlling gene expression in living cells through small molecule-RNA interactions. *Science* 282:296–98
100. White HB. 1976. Coenzymes as fossils of an earlier metabolic state. *J. Mol. Evol.* 7:101–4
101. Wickiser JK, Winkler WC, Breaker RR, Crothers DM. 2005. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell* 18:49–60
102. Winkler W, Nahvi A, Breaker RR. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419:952–56
103. Winkler WC. 2005. Metabolic monitoring by bacterial mRNAs. *Arch. Microbiol.* 183:151–59
104. Winkler WC, Breaker RR. 2003. Genetic control by metabolite-binding riboswitches. *Chembiochem* 4:1024–32

105. Winkler WC, Cohen-Chalamish S, Breaker RR. 2002. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl. Acad. Sci. USA* 99:15908–13
106. Winkler WC, Grundy FJ, Murphy BA, Henkin TM. 2001. The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. *RNA* 7:1165–72
107. Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428:281–86
108. Winkler WC, Nahvi A, Sudarsan N, Barrick JE, Breaker RR. 2003. An mRNA structure that controls gene expression by binding *S*-adenosylmethionine. *Nat. Struct. Biol.* 10:701–7
109. Yakhnin H, Zhang H, Yakhnin AV, Babitzke P. 2004. The trp RNA-binding attenuation protein of *Bacillus subtilis* regulates translation of the tryptophan transport gene *trpP* (*ybaG*) by blocking ribosome binding. *J. Bacteriol.* 186:278–86
110. Yarnell WS, Roberts JW. 1999. Mechanism of intrinsic transcription termination and antitermination. *Science* 284:611–15
111. Yen L, Svendsen J, Lee JS, Gray JT, Magnier M, et al. 2004. Exogenous control of mammalian gene expression through modulation of RNA self-cleavage. *Nature* 431:471–76