Critical Reviews in Biochemistry and Molecular Biology, 41:329–338, 2006

ISSN: 1040-9238 print / 1549-7798 online
DOI: 10.1080/10409230600914294

informa
healthcare

From Ribosome to Riboswitch: Control of Gene Expression in Bacteria by RNA Structural Rearrangements

Frank J. Grundy
and Tina M. Henkin
Department of Microbiology,
The Ohio State University,
Columbus, Ohio, USA

ABSTRACT Structural elements in the 5′ region of a bacterial mRNA can have major effects on expression of downstream coding sequences. Folding of the nascent RNA into the helix of an intrinsic transcriptional terminator results in premature termination of transcription and in failure to synthesize the full-length transcript. Structure in the translation initiation region of an mRNA blocks access of the translation initiation complex to the ribosome binding site, thereby preventing protein synthesis. RNA structures can also affect the stability of an RNA by altering sensitivity to ribonucleases. A wide variety of mechanisms have been uncovered in which changes in mRNA structure in response to a regulatory signal are used to modulate gene expression in bacteria. These systems allow the cell to recognize an impressive array of signals, and to monitor those signals in many different ways.

KEYWORDS transcription, translation, attenuation, antitermination, regulation

INTRODUCTION

RNA structures play many crucial roles in bacterial cells. A simple helix in a nascent RNA is sufficient to pause an elongating RNA polymerase (RNAP), and in conjunction with a U-A RNA-DNA hybrid can result in transcription termination. Helical elements can also serve as substrates for cleavage by endoribonucleases. Both simple helices and more complicated structures allow specific binding of ligands ranging from proteins to small molecules, and complex RNAs play crucial roles in a number of cellular processes, while some RNAs even possess enzymatic activities (*e.g.*, the peptidyltransferase activity of rRNA, or tRNA processing by RNase P). Structural elements within mRNAs can also impact gene expression by modulating the synthesis of the full-length mRNA or its post-transcriptional fate (reviewed by Henkin, 1996; Henkin & Yanofsky, 2002; Grundy & Henkin, 2004). These regulatory RNA elements can interact directly with the transcriptional machinery (as in the bacteriophage HK022 *put* system) or they can act as binding sites for regulatory factors that then interact with RNAP (as in the bacteriophage lambda N antitermination system). The formation of the regulatory RNA element can also be determined by

Address correspondence to Tina M. Henkin, Department of Microbiology, The Ohio State University, 484 W. 12th Avenue, Columbus, OH 43210, USA. E-mail henkin.3@osu.edu

competition between alternate structural forms (Fig. 1). These regulatory elements are usually located upstream of the regulated coding sequences, within what are generally called "leader regions" of the mRNA; while in most cases these regions can also be termed "5′-untranslated regions" (5′-UTRs), this nomenclature is inappropriate for systems in which the 5′ region is in fact translated, so that leader region is the more general term.

The analysis of systems like the *Escherichia coli trp* biosynthetic operon provided the first example of a regulatory mechanism that employs competing RNA elements to control gene expression. In systems of this type, the regulated coding sequences are preceded by a short peptide coding region, and the processivity of the ribosome during translation of the leader peptide coding sequence determines whether the 5′ region of the mRNA folds into a terminator helix or a competing antiterminator helix. Leader peptide translation is also used in certain antibiotic resistance genes to control an RNA structural switch that regulates accessibility of the ribosome binding site of the downstream gene. Numerous systems use RNA binding proteins to mediate RNA structural switches, either through direct measurement of the availability of the protein itself (*e.g.*, for ribosomal protein operons) or by changing the RNA binding activity of a regulatory protein in response to an effector (*e.g.*, *E. coli* BglG or *Bacillus subtilis* TRAP). Leader RNAs can also interact with other RNAs, including small RNAs that interact by complementary base pairing or tRNAs, and binding of the regulatory RNA can affect transcription termination, translation initiation or RNA degradation. Leader RNAs that directly monitor a variety of small molecules, including nucleotides, cofactors and amino acids, have recently been described in bacteria, while "thermosensor" RNAs monitor temperature by simple melting of helical domains, without a requirement for binding of anything at all. It is apparent that bacterial cells are adept at exploiting changes in RNA structure to regulate gene expression, and this review will provide an introduction to the variations on this theme that have been uncovered so far.

RIBOSOMES AS SENSORS

The processivity of a ribosome as it traverses an upstream coding sequence can influence the fate of the downstream portion of the mRNA. Interruption of translation can lead to structural changes in the mRNA that affect accessibility of the ribosome binding site of the downstream gene through translational coupling, and can also determine whether the complete mRNA is synthesized through transcriptional polarity. In addition, loss of protection of an mRNA by translating ribosomes often leads to faster degradation of the mRNA. These general effects can be directed to a more specific regulatory response in systems in which the activity of the ribosome translating the upstream coding region responds to a physiological signal, so that the translating ribosome in some way monitors the signal and impacts downstream gene expression. Several systems of this type have been identified, and these can be categorized according to the type of signal sensed and the type of regulatory response.

Leader Peptide Translation Monitors tRNA Charging

The classical example of regulation of gene expression *via* translation-dependent changes in RNA structure is represented by the *E. coli trp* operon (reviewed by Yanofsky, 1981; Landick *et al.*, 1996). In this system, the regulated *trp* biosynthesis genes are preceded by a leader region that contains a short peptide coding sequence followed by an intrinsic transcriptional terminator. RNAP pauses during transcription of the leader region, allowing translation of the leader peptide to begin. When tryptophan is abundant in the cell, the pool of $tRNA^{Trp}$ is efficiently aminoacylated, allowing efficient translation of the peptide coding sequence, which contains tandem tryptophan codons. Rapid progression of the ribosome releases RNAP from the pause site, and the nascent RNA folds into the helix of an intrinsic terminator, resulting in termination of transcription before RNAP reaches the start of the *trp* biosynthesis genes. Expression of the *trp* genes is therefore repressed when the cellular pool of aminoacylated $tRNA^{Trp}$ is high, indicating an adequate supply of tryptophan. In contrast, a decrease in aminoacylated $tRNA^{Trp}$ when tryptophan is limiting results in stalling of the ribosome during translation of the tandem tryptophan codons. The stalled ribosome triggers formation of an alternate RNA structure that serves as an antiterminator element, preventing formation of the terminator helix. RNAP resumes transcription and proceeds through the termination site, resulting in transcription of the full-length mRNA and synthesis of the biosynthetic enzymes. The translational activity of the ribosome translating the

leader peptide coding sequence therefore responds to the abundance of aminoacylated $tRNA^{Trp}$, and that information is transmitted to the transcription elongation complex (TEC) through a structural rearrangement of the nascent RNA. A similar mechanism is used for other amino acid biosynthesis and aminoacyl-tRNA synthetase genes, predominantly in Gram-negative organisms. The key *cis*-acting feature that specifies the response to a particular aminoacyl-tRNA is the presence of multiple codons for the cognate tRNA(s) within the leader peptide coding sequence. This type of mechanism exploits the coupling of transcription and translation in bacterial cells, since leader peptide translation must begin on the newly synthesized mRNA as the RNA emerges from the TEC.

Leader Peptide Translation Monitors an Amino Acid

Leader peptide translation is also exploited to directly measure tryptophan in the *E. coli tna* operon, which encodes tryptophanase. In this case, stalling of the ribosome is promoted by interaction of free tryptophan with the nascent leader peptide within the ribosome, and the stalled ribosome occludes a binding site for the Rho transcription termination factor, allowing the TEC to continue transcription of the downstream coding region (Konan & Yanofsky, 1997; Gong *et al.*, 2001). Tryptophan-induced stalling requires specific interactions between the leader peptide and the ribosome exit channel (Cruz-Vera *et al.*, 2005). This system does not involve an RNA structural rearrangement, and illustrates the ability of the ribosome, in conjunction with a nascent peptide, to directly sense a small molecule. It is remarkable that genes for both tryptophan biosynthesis and utilization are regulated using leader peptide translation in *E. coli*, although the molecular mechanism for sensing tryptophan abundance and the response to tryptophan are quite different in the two systems. The *tna* mechanism, like the *E. coli trp* system, depends on coupling of transcription and translation since processivity of the ribosome during leader peptide synthesis impacts transcription termination.

Ribosome Sensing of an Antibiotic

Since the biological role of the ribosome involves the incorporation of amino acids into protein, it is not surprising that mechanisms have evolved to sense amino acid availability using the translational activity of the ribosome. Similarly, expression of genes encoding enzymes that protect the cell from antibiotics that target the ribosome can be regulated by effects of the antibiotic on leader peptide translation. Examples of this type of mechanism are found in several Gram-positive *cat* (chloramphenicol acetyltransferase) and *erm* (erythromycin resistance) genes. In systems of this type, the regulated coding sequences are preceded by leader peptide coding regions, translation of which affects downstream mRNA structure. For example, in *cat* genes of this type, the Shine-Dalgarno (SD) sequence of the downstream *cat* coding sequence is sequestered in a structure that prevents translation initiation unless the structure is released by chloramphenicol-induced stalling of a ribosome synthesizing the leader peptide (Gu *et al.*, 1994; Lovett & Rogers, 1996). Stalling of the ribosome in the *cat* leader in response to chloramphenicol requires specific interactions between the nascent leader peptide and the ribosome, as was observed in the *tna* system. The major differences between the *tna* and *cat* systems involve the specificity of the effector response (tryptophan *vs.* chloramphenicol) and the mechanism for regulation of the downstream coding sequence (transcription termination *vs.* translation initiation). It is notable that while specific features of the nascent peptide are required in conjunction with the effector molecule to promote ribosome stalling in both the *tna* and *cat* systems, sensing of a ribosomally targeted antibiotic by a translating ribosome is uniquely suited to regulation of a gene that confers resistance to that antibiotic.

EFFECTOR SENSING BY RNA BINDING PROTEINS

Specific binding to RNA targets is a well-established feature of RNA binding proteins, and this property has been exploited in a number of regulatory systems. In many cases, binding of a protein to its mRNA target results in autogenous repression when the protein is available in excess of the primary cellular target for the protein. This type of mechanism is exemplified by the repression of ribosomal protein operons by individual ribosomal proteins when the amount of free protein exceeds the amount of free ribosomal RNA, signalling that no additional synthesis of the ribosomal proteins is required. These systems directly monitor the product of the regulated gene, with no other effector molecule. The activity of RNA binding proteins can also be controlled

by titration with noncoding RNAs that mimic the regulatory target site of the protein, as in the Csr system (Romeo, 1998; Weilbacker *et al.*, 2003; Majdalani *et al.*, 2005). Other RNA binding proteins exhibit changes in binding in response to the presence of an effector molecule, so that the protein serves as the sensor for the effector molecule. Binding of the protein to the RNA can have a variety of effects that modulate gene expression at the levels of transcription termination, translation initiation and RNA degradation, and can also cause RNA structural rearrangements that affect gene expression (reviewed by Romby & Springer, 2003). We will focus on specific, well-characterized examples that serve as paradigms for effector-dependent modulation of leader RNA structure.

Effector-Dependent Repression by RNA Binding Proteins

As in *E. coli*, expression of genes involved in tryptophan biosynthesis in *B. subtilis* is repressed when tryptophan is abundant. Regulation once again occurs primarily at the level of premature termination of transcription, but tryptophan is sensed by the TRAP protein rather than during leader peptide translation (reviewed by Gollnick *et al.*, 2005). TRAP forms an 11-mer ring structure, and binding of tryptophan to the TRAP complex is required for RNA binding activity (Antson *et al.*, 1999). Binding of the TRAP-Trp complex to the 5′ region of *trp* operon mRNA prevents formation of an antiterminator element, allowing formation of the competing, less stable terminator helix. Sensing of $tRNA^{Trp}$ charging is superimposed on the TRAP regulatory mechanism by the AT (anti-TRAP) protein, an antagonist of TRAP activity (Valbuzzi *et al.*, 2001), which is regulated by $tRNA^{Trp}$ via the T box transcription antitermination mechanism (Sarsero *et al.*, 2000; see below). This dual regulation, which allows sensing of both free tryptophan and the $tRNA^{Trp}$ aminoacylation reaction, is analogous to *E. coli trp* operon regulation which occurs both at the level of transcription initiation, where free tryptophan is sensed by the TrpR DNA binding protein, and at the level of premature translation, where $tRNA^{Trp}$ charging is sensed by the ribosome (reviewed by Yanofsky, 2004).

TRAP (in complex with tryptophan) can also regulate translation initiation of some of its target genes, either by direct occlusion of the SD region, or by promoting an RNA rearrangement that sequesters the SD (reviewed by Babitzke, 2004). Some targets of TRAP are regulated only at the level of translation, while others are regulated at both levels, resulting in a second opportunity to monitor tryptophan and turn off gene expression after the TEC has passed the termination site. Translational control has the potential to allow reversibility of the regulatory decision, as release of the TRAP-Trp complex can allow a previously inactive mRNA to be translated. Transcription termination control, in contrast, is essentially irreversible, since once the TEC has released from the termination site, or has moved beyond the termination site, the decision has been made for that TEC.

Effector-Dependent Activation by RNA Binding Proteins

The *E. coli bgl* operon, which is involved in utilization of β-glucoside sugars, is regulated by the BglG RNA binding protein, which binds as a dimer to an antiterminator element in the 5′ region of the *bgl* mRNA, stabilizing the antiterminator and preventing formation of a competing transcriptional terminator (Amster-Choder & Wright, 1993, 1997). BglG RNA binding activity is controlled by a second protein, BglF, which also serves as the transporter for the substrate sugars. When the substrate sugar is available, BglF phosphorylates the sugar during transport; when the sugar is absent, BglF instead phosphorylates BglG, which inhibits BglG dimerization and causes loss of RNA binding activity. BglG therefore responds indirectly to the effector, via signalling from BglF, in contrast to TRAP which directly senses its molecular effector. In addition, binding of the protein to the target RNA has the opposite effect on gene expression, as BglG promotes antitermination while TRAP promotes termination.

TRANS-ACTING RNAs

Regulation of gene expression by small non-coding RNAs is emerging as a common mechanism in both prokaryotes and eukaryotes. While *cis*-encoded antisense RNAs generally operate by extensive base-pairing, because they are completely complementary to their target RNA, *trans*-encoded regulatory RNAs are only partially complementary to their targets, and target binding can have a variety of effects on target RNA structure, translation and stability. Binding of *trans*-encoded regulatory RNAs often requires the RNA chaperone protein Hfq. Most small RNAs negatively affect gene

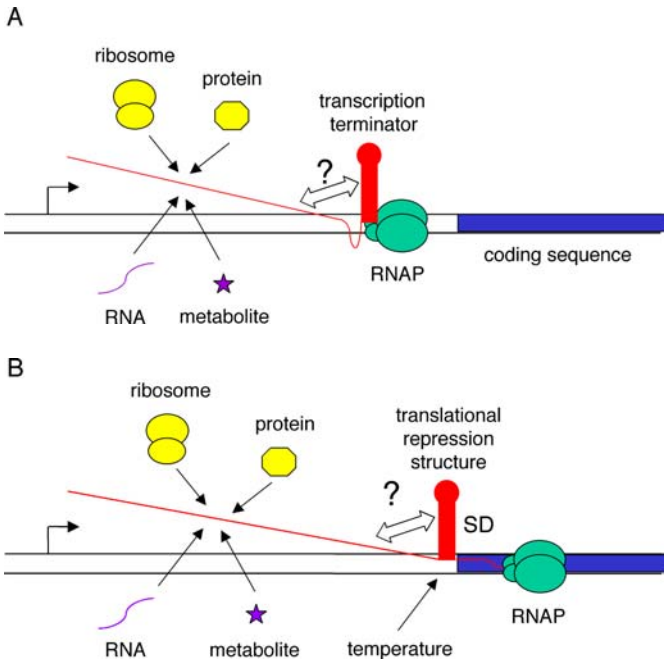

FIGURE 1 **Models for sensing of regulatory signals by leader RNAs. *A.* Transcription termination control. Interaction of a regulatory molecule (a translating ribosome, an RNA binding protein, a small RNA or a small molecule) modulates the structure of the nascent RNA as it emerges from RNAP. This structural modulation determines whether the transcript folds into the helix of an intrinsic terminator, resulting in premature termination of transcription, or a competing antiterminator that sequesters sequences necessary for formation of the terminator, resulting in continued transcription and expression of the downstream coding sequence. *B.* Translational control. Interaction of a regulatory molecule, or changes in temperature, determine whether the RNA transcript folds into a structure that sequesters the SD sequence of the downstream coding region. Sequestration of the SD results in inhibition of translation, while release of the SD, either directly or by folding of the RNA into a competing structure that allows access of the translational machinery to the SD, allows expression of the downstream coding sequence.**

expression, either by directly sequestering the translation initiation region of target mRNA, as in the case of repression of the *hns* global regulator by the DsrA RNA in *E. coli*, or by causing rapid degradation of their target mRNAs, which can occur as a consequencce of creating a recognition determinant for cellular endoribonucleases, or by loss of protection of the mRNA by translating ribosomes. There are also examples of positive effects on gene expression, including the activation of *rpoS* translation by binding of DsrA RNA to a region of the mRNA that would otherwise sequester the ribosome binding site. As noted above, small RNAs can also modulate gene expression by titrating an RNA binding protein, as in the CsrA system. Small RNA regulators have been reviewed recently (Majdalani *et al.*, 2005; Storz *et al.*, 2005) and will not be discussed further here. Regulation generally operates at the level of small RNA synthesis; use of tRNA as a regulator in the T box mechanism represents a special case where the activity of the RNA is modulated by aminoacylation, and this system is considered separately below.

RIBOSWITCH RNAs

A number of systems have recently been described in which sequence elements in the 5′ region of a transcript can directly monitor a physiological signal. Recognition of the signal results in a transition in the structural arrangement of the leader RNA that impacts expression of the downstream coding sequence(s) (reviewed by Grundy & Henkin, 2004; Tucker & Breaker, 2005). The most common effects occur at the level of premature termination of transcription or translation initiation; one system exhibits RNA cleavage in response to the effector, which probably affects mRNA stability. In each of these systems, the leader RNAs have in common the ability to sense their signal in the absense of any additional regulatory factor, such as an RNA binding protein or a translating ribosome. The signals that can be sensed range from physical parameters, such as temperature, to cellular metabolites, such as nucleotides, amino acids and cofactors, to tRNA. Each class of riboswitch RNA responds specifically to its cognate signal, and with the exception of the RNA thermosensors, each class exhibits a set of conserved sequence and structural elements that are involved in signal recognition and the regulatory response. The ability of these RNA elements to respond directly to their signal without a requirement for *trans*-acting cellular components is likely to be responsible for their widespread dissemination by horizontal gene transfer. Artificial versions of these regulatory elements have also been designed to allow controlled gene expression *in vivo* by molecules such as tetracycline (Suess *et al.*, 2003; Hanson *et al.*, 2005).

RNA Thermosensors

The RNA thermosensors represent the simplest class of riboswitch RNAs. These regulatory elements form structures that sequester the translation initiation region of the target gene under standard growth conditions, and unfold to allow translation in response to increased temperature. Regulatory elements of this type are used to control the heat shock response in *E. coli*, as well as a key regulator of pathogenesis in *Listeria* (Morita *et al.*, 1999; Johansson *et al.*, 2002). These RNA

elements have no specific recognition properties, and appear instead to have their intrinsic stability calibrated to allow the default state of the mRNA to maintain folding under growth conditions where the regulated gene is repressed; increased temperature is sufficient to destabilize the RNA structure so that the ribosome binding site of the downstream coding sequence becomes accessible to the translational machinery. The simplicity of these systems is beautifully suited to the regulatory signal being monitored, and allows a very rapid response to changes in growth temperature as the mRNA is already present in the cell. It also seems likely that this type of mechanism has the potential to be reversible, since a reduction in growth temperature can result in refolding of the mRNA into the inhibitory structure, preventing further translation. It remains possible that similar thermosensitive RNA structures could act at levels of gene expression other than translation initiation, such as by the presence of a temperature-sensitive RNase binding element or intrinsic terminator helix, but no system of this type has been reported to date.

Metabolite-Sensing Riboswitches

Recent studies have revealed a growing list of systems in which elements within the 5′ region of a transcript can directly and specifically bind a small molecule within the cell, resulting in an RNA structural rearrangement that impacts gene expression. Each of these systems represents a family of genes that are usually involved in biosynthesis or transport of a particular metabolite which feedback regulates the expression of those genes. Expression of genes involved in utilization of a compound can also be induced by that compound, so that these systems can act either positively or negatively, although negative systems appear to predominate.

Nou and Kadner (2000) provided the first clear evidence of a direct interaction of an effector molecule (adenosylcobalamin) with a target mRNA (*E. coli btuB*), and also showed that binding of the effector to the RNA inhibited ribosome binding to the SD sequence. Subsequent studies showed that the conserved B_{12} box identified in the initial studies is part of a larger element conserved in a number of genes involved in cobalamin biosynthesis and transport (Vitreschak *et al.*, 2003), and that RNAs in this family exhibit structural alterations in response to effector binding (Nahvi *et al.*, 2002). Similar studies of genes involved in biosynthesis and transport of thiamine (Miranda-Rios *et al.*, 2001; Rodionov *et al.*, 2002; Winkler *et al.*, 2002b), flavin mononucleotide (Gelfand *et al.*, 1999; Vitreschak *et al.*, 2002; Mironov *et al.*, 2002; Winkler *et al.*, 2002), guanine (Johansen *et al.*, 2003; Mandal *et al.*, 2003), adenine (Johansen *et al.*, 2003; Mandal & Breaker, 2004), *S*-adenosylmethionine (SAM; Grundy *et al.*, 1998; McDaniel *et al.*, 2003; Ephstein *et al.*, 2003; Winkler *et al.*, 2003), lysine (Grundy *et al.*, 2003; Sudarsen *et al.*, 2003), glucosamine-6-phosphate (Winkler *et al.*, 2004), glycine (Mandal *et al.*, 2004), and magnesium (Cromie *et al.*, 2006) have quickly revealed that mechanisms of this type are widespread in bacteria, and can regulate gene expression at several different levels. Three different classes of SAM-binding riboswitch RNAs have been identified: the S box, which is found predominantly in *Bacillus*, *Clostridium* and *Staphylococcus* sp. (Grundy and Henkin, 1998, 2002); the S_{MK} box, which is found in SAM synthetase (*metK*) genes in members of the Lactobacillales (Fuchs *et al.*, 2006); and the SAM-II box, which is in alpha-proteobacteria (Corbino *et al.*, 2005).

Each system of this type is characterized by a conserved RNA element required for specific effector binding, and binding results in an RNA rearrangement that affects formation of a terminator helix or a helix that sequesters the SD sequence; the glucosamine-6-phosphate system is unique so far in that effector binding stimulates self-cleavage of the RNA, which may affect mRNA stability. In general, systems that regulate at the level of premature termination of transcription tend to predominate in low G+C Gram-positive bacteria, while systems that regulate at the level of translation initiation are more common in Gram-negative and high G+C Gram-positive bacteria, although there are exceptions to this pattern. Thiamine-binding RNA elements have also been identified in eukaryotic organisms, where it has been suggested that effector binding may inhibit mRNA splicing (Kubodera *et al.*, 2003).

The most remarkable feature of the metabolite binding riboswitch RNAs is the very high specificity of each RNA for its cognate effector molecule. High-resolution structural information is currently available for purine and thiamine binding riboswitches (Batey *et al.*, 2004; Serganov *et al.*, 2004; Serganov *et al.*, 2006; Thore *et al.*, 2006), and these structures demonstrate binding of the effector within a highly structured pocket formed at a junction between three helices, with extensive interactions with the surface of the ligand, consistent with the high specificity of ligand recognition. Discrimination between guanine and adenine by their respective

RNA elements was shown to depend on a single C to U substitution (Mandal & Breaker, 2004; Serganov *et al.*, 2004; Noeske *et al.*, 2005). The basis for the specificity and selectivity of more complex riboswitch RNAs remains to be determined, but all of them have been shown to effectively differentiate between the natural ligand and related compounds.

Another crucial feature of the metabolite binding RNAs is their affinity for their respective ligands, which must be matched to the intracellular concentration of the metabolite for an appropriate regulatory response. For example, the S box SAM-binding riboswitches exhibit efficient SAM-dependent transcription termination *in vitro* at low micromolar concentrations of SAM (McDaniel *et al.*, 2003). Total SAM concentrations in *B. subtilis*, in which the S box system is found, have been measured at 80 to 400 μM (Wabiko *et al.*, 1988); since a large proportion of the SAM is likely to be bound to other cellular components, the effective SAM concentration *in vitro* is probably within the range of the free SAM pool *in vivo*. In contrast, the *B. subtilis* L box riboswitch requires approximately 3 mM lysine for efficient lysine-dependent transcription termination *in vitro* (Grundy *et al.*, 2003), consistent with the intracellular lysine pool of 2 to 4 mM (Tempest *et al.*, 1970). Modulation of effector concentration *in vivo* (*e.g.*, by mutation or overproduction of SAM synthetase for S box genes) has the predicted effect on gene expression (McDaniel *et al.*, 2003, 2006), consistent with the hypothesis that these regulatory RNAs are sensitive to variations in effector pools.

A tRNA-Sensing Riboswitch: The T Box Mechanism

Analysis of the *B. subtilis tyrS* gene, encoding tyrosyl-tRNA synthetase, revealed that readthrough of a leader region termination site requires binding of uncharged $tRNA^{Tyr}$ to the nascent transcript (Grundy & Henkin, 1993). Specific recognition of $tRNA^{Tyr}$ depends on base-pairing of the tRNA anticodon with a tyrosine codon, designated the "specifier sequence," located in an unpaired domain of the leader RNA; a second base-pairing interaction between the acceptor end of the tRNA and residues within an antiterminator element stabilizes the antiterminator and prevents formation of the competing terminator helix. Codon-anticodon pairing occurs with either uncharged or charged tRNA, while tRNA acceptor end-antiterminator pairing, and antitermination, occurs only with uncharged tRNA, allowing the system to monitor the relative amounts of charged and uncharged tRNA (Grundy *et al.*, 1994). The T box mechanism is utilized for a wide range of amino acid-related genes, with each gene responding specifically to the cognate tRNA. All T box genes exhibit a common pattern of sequence and structural elements in the leader region of the mRNA, although primary sequence conservation is low. Variation of the codon found at the position of the specifier sequence in *tyrS* dictates the identity of the regulatory tRNA for each gene in the family.

Biochemical analyses demonstrated that tRNA-dependent antitermination can take place in a minimal *in vitro* transcription system using purified components (Grundy *et al.*, 2002; Putzer *et al.*, 2002). Antitermination requires a match between the specifier sequence in the leader RNA and the anticodon of the tRNA, so that the nascent transcript mediates codon-anticodon recognition in the absence of a ribosome. T box leader RNA transcribed *in vitro* by T7 RNAP can specifically bind the cognate tRNA in solution, allowing mapping of structural alterations in each partner in the complex (Yousef *et al.*, 2005). Multiple changes were detected in the leader RNA in response to binding of the uncharged tRNA; binding of an RNA that is a mimic of charged tRNA confers protection of the specifier loop but no other changes, indicating that the remaining structural changes require the acceptor end-antiterminator interaction that is blocked by charging of the tRNA. tRNA-leader RNA binding results in protection of the anticodon loop of the tRNA, as expected, as well as protection of G19 in the D loop. The tRNA-dependent structural switch between the terminator and antiterminator forms of the leader RNA was also demonstrated. Together, these studies demonstrated that T box leader RNAs can specifically interact with the cognate uncharged tRNA in the absence of any other cellular components, and that they can also discriminate against non-cognate tRNA and differentiate between charged and uncharged tRNA.

The T box mechanism is found in all groups of Gram-positive bacteria, and is especially prominent in the Firmicutes (*e.g.*, *Bacillus* and *Clostridium*). The system is found more rarely in high G+C Gram-positive bacteria, and has been identified sporadically in Gram-negative organisms (Grundy & Henkin, 2003). In the high G+C Gram-positive and Gram-negative T box genes, it appears that regulation occurs at the level of translation

initiation rather than transcription termination, as the intrinsic terminator helix is replaced by a helical element that sequesters the SD sequence of the downstream coding sequence. In genes of this type, binding of uncharged tRNA to the leader RNA is predicted to stabilize an element analogous to the antiterminator element of T box leaders that are regulated by transcription termination, preventing formation of the helix that sequesters the SD sequence. No biochemical studies have yet been carried out to verify this prediction.

The most remarkable aspect of the T box mechanism is its adaptability to genes of different amino acid classes by minor changes in the leader sequence (primarily the specifier sequence itself, although other determinants contribute to leader RNA-tRNA recognition). This adaptability is analogous to systems (like *E. coli trp*) that monitor tRNA charging through leader peptide translation. In that case, a switch in amino acid specificity is mediated by changing the peptide coding sequence, so that ribosome stalling occurs in response to limitation for the appropriate charged tRNA. The high conservation of tRNA structure is likely to have facilitated dispersion of the T box system by horizontal gene transfer, as the regulatory element is *cis*-encoded and the signal being sensed exhibits low variability in different bacterial systems.

EFFECTOR SENSING BY RNAP

The regulatory system for the *B. subtilis pyrG* gene, which encodes CTP synthase, represents a special example of gene regulation by alternate RNA structures. The *pyrG* leader RNA contains an intrinsic terminator but no competing antiterminator, and only the first 4 nt of the transcript (GGGC) are required in conjunction with the terminator for readthrough of the terminator in response to limitation for CTP (Meng & Switzer, 2002). The leader region terminator is active when CTP is high, resulting in premature termination and repression of expression. The readthrough transcript is generated when CTP is limiting by pausing of RNAP before addition of the C residue at the +4 position; pausing results in the nontemplated addition of multiple G residues by reiterative transcription, and the resulting poly(G) sequence in the transcript pairs with a run of pyrimidines present in the 5′ side of the terminator, preventing formation of the terminator helix (Meng *et al.*, 2004). The antiterminator element is therefore not encoded in the DNA template, and is instead generated by "misbehavior" of RNAP during transcription of the initial portion of the gene. The transcriptional machinery itself senses the effector molecule (CTP), resulting in the synthesis of two different RNA species (terminator *vs.* antiterminator) that differ in their 5′ sequence, depending on the concentration of the effector. This unique system should serve as a warning that not all RNA elements can be predicted based on genomic sequence, and that the gene expression machinery can play clever tricks on those of us who like to make such predictions!

CONCLUSIONS AND PERSPECTIVES

Recent studies have revealed that the ability of RNA sequences to act in *cis* to modulate gene expression extends beyond a few isolated examples to a very large number of genes in many experimental systems. Many of these types of systems can be recognized by searching for conserved patterns in raw genomic sequence, while others (like the *pyrG* system) are more difficult to predict *a priori*. Nevertheless, our ability to identify likely regulatory patterns by comparison with known systems provides new tools for predicting regulatory mechanisms. In some cases, recognition of a regulatory RNA pattern can provide important insight not only in terms of gene regulation but also in terms of the function of the downstream gene. For example, identification of a leader peptide with a particular set of codons signals regulation in response to the matching tRNA, as does identification of the specifier sequence in a T box leader; similarly, identification of the S box RNA pattern upstream of a set of genes of unknown function in *B. subtilis* was used to predict roles in a previously unknown methionine recycling pathway (Grundy & Henkin, 2002; Murphy *et al.*, 2002). Improvements in bioinformatics approaches and recognition of new classes of RNA patterns will provide a continually growing set of tools and paradigms that will be enhanced by the wealth of genomic sequence data. It is clear that biological systems have explored a large variety of RNA arrangements to control gene expression, and our understanding of the roles of these arrangements is limited only by our ability to recognize them and determine how they work.

ACKNOWLEDGMENTS

This work was supported by NIH grants GM47823 and GM63615.

REFERENCES

Amster-Choder, O. and Wright, A. 1993. Transcriptional regulation of the *bgl* operon of *Escherichia coli* involves phosphotransferase system-mediated phosphorylation of a transcriptional antiterminator. *J Cell Biochem* 51:83–90.

Amster-Choder, O. and Wright, A. 1997. BglG, the response regulator of the *Escherichia coli bgl* operon, is phosphorylated on a histidine residue. *J Bacteriol* 179:5621–5624.

Antson, A.A., Dodson, E.J., Dodson, G.G., Greaves, R.B., Chen, Z., and Gollnick, P. 1999. Structure of the *trp* RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* 401:235–242.

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* 7:132–139.

Batey, R.T., Gilbert, S.D., and Montange, R.K. 2004. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* 432:411–415.

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

Cromie, M.H., Shi, Y., Latifi, T., and Groisman, E.A. 2006. An RNA sensor for intracellular Mg_2^+. *Cell* 125:71–84.

Cruz-Vera, L.R., Rajagopal, S., Squires, C., and Yanosky, C. 2005. Features of ribosome-peptidyl-tRNA interactions essential for tryptophan induction of *tna* operon expression. *Mol Cell* 19:333–343.

Epshtein, V., Mironov, A.S., and Nudler, E. 2003. The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc Natl Acad Sci USA* 100:5052–5056.

Fuchs, R.T., Grundy, F.J., and Henkin, T.M. 2006. The S_{MK} box is a new SAM-binding RNA for translational regulation of SAM synthetase. *Nat Struct Mol Biol* 13:226–233.

Gelfand, M.S., Mironov, A.A., Jomantas, J., Kozlov, Y.I., and Perumov, D.A. 1999. A conserved RNA structure element involved in the regulation of bacterial riboflavin biosynthesis genes. *Trends Genet* 15:439–442.

Gollnick, P., Babitzke, P., Antson, A., and Yanofsky, C. 2005. Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis*. *Annu Rev Genet* 39:47–68.

Gong, F., Ito, K., Nakamura, Y., and Yanofsky, C. 2001. The mechanism of tryptophan induction of tryptophanase operon expression: tryptophan inhibits release factor-mediated cleavage of TnaC-peptidyl-$tRNA^{Pro}$. *Proc Natl Acad Sci USA* 98:8997–9001.

Grundy, F.J. and Henkin, T.M. 1993. tRNA as a positive regulator of transcription antitermination in *B Subtilis Cell* 74:475–482.

Grundy, F.J. and Henkin, T.M. 1998. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Mol Microbiol* 30:737–749.

Grundy, F.J. and Henkin, T.M. 2002. Synthesis of serine, glycine, cysteine and methionine. In: *Bacillus subtilis and Its Closest Relatives: from Genes to Cells*. Sonenshein, A. L., Hoch, J. A., and Losick, R., Eds., American Society for Microbiology Press, Washington, DC, pp. 245–254.

Grundy, F.J. and Henkin, T.M. 2003. The T box and S box transcription termination control systems. *Front Biosci* 8:d20–31.

Grundy, F.J. and Henkin, T.M. 2004. Regulation of gene expression by effectors that bind to RNA. *Curr Opin Microbiol* 7:126–131.

Grundy, F.J., Lehman, S.C., and Henkin, T.M. 2003. The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc Natl Acad Sci USA* 100:12057–12062.

Grundy, F.J., Rollins, S.M., and Henkin, T.M. 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* 176:4518–4526.

Grundy, F.J., Winkler, W.C., and Henkin, T.M. 2002. tRNA-mediated transcription antitermination *in vitro*: codon-anticodon pairing independent of the ribosome. *Proc Natl Acad Sci USA* 99:11121–11126.

Gu, Z., Harrod, R., Rogers, E.J., and Lovett, P.S. 1994. Anti-peptidyltransferase leader peptides of attenuation-regulated chloramphenicol resistance genes. *Proc Natl Acad Sci USA* 91:5612–5616.

Hanson, S., Bauer, G., Fink, B., and Suess, B. 2005. Molecular analysis of a synthetic tetracycline-binding riboswitch. *RNA* 11:503–511.

Henkin, T.M. 1996. Control of transcription termination in prokaryotes. *Annu Rev Genet* 30:35–57.

Henkin, T.M. and Yanofsky, C. 2002. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *BioEssays* 24:700–707.

Johansen, L.E., Nygaard, P., Lassen, C., Agerso, Y., and Saxild, H.H. 2003. Definition of a second *Bacillus subtilis pur* regulon comprising the *pur* and *xpt-pbuX* operons plus *pbuG, nupG* (*yxjA*), and *pbuE* (*ydhL*). *J Bacteriol* 185:5200–5209.

Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., and Cossart, P. 2002. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes. Cell* 110:551–561.

Konan, K.V. and Yanofsky, C. 1997. Regulation of the *Escherichia coli tna* operon: nascent leader peptide control at the *tnaC* stop codon. *J Bacteriol* 179:1774–1779.

Kubodera, T., Watanabe, M., Yoshiuchi, K., Yamashita, N., Nishimura, A., Nakai, S., Gomi, K., and Hanamoto, H. 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* 555:516–520.

Landick, R., Turnbough, C.L., and Yanofsky, C. 1996. Transcription attenuation. In: *Escherichia coli and Salmonella: Cellular and Molecular Biology.* Neidhardt, F. C., et al., Eds. American Society for Microbiology, Washington, DC, pp. 1263–1286.

Lovett, P.S. and Rogers, E.J. 1996. Ribosome regulation by the nascent peptide. *Microbiol Rev* 60:366–385.

Majdalani, N., Vanderpool, C.K., and Gottesman, S. 2005. Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol* 40:93–113.

Mandal, M., Boese, B., Barrick, J.E., Winkler, W.C., and Breaker, R. R. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* 113:577–586.

Mandal, M. and Breaker, R.R. 2004. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat Struct Mol Biol* 11:29–35.

Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L., and Breaker, R.R. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* 306:275–279.

McDaniel, B.A.M., Grundy, F.J., Artsimovitch, I., and Henkin, T.M. 2003. Transcription termination control of the S box system: direct measurement of *S*-adenosylmethionine by the leader RNA. *Proc Natl Acad Sci USA* 100:3083–3088.

McDaniel, B.A., Grundy, F.J., Kurlekar, V.P., Tomsic, J., and Henkin, T.M. 2006. Identification of a mutation in the *Bacillus subtilis S*-adenosylmethionine synthetase gene that results in derepression of S-box gene expression. *J Bacteriol* 188:3674–3681.

Meng, Q., Turnbough, C.L. Jr., and Switzer, R.L. 2004. Attenuation control of *pyrG* expression in *Bacillus subtilis* is mediated by CTP-sensitive reiterative transcription. *Proc Natl Acad Sci USA* 101:10943–10948.

Miranda-Rios, J., Navarro, M., and 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* 98:9736–9741.

Mironov, A.S., Gusarov, I., Rafikov, R., Lopez, L.E., Shatalin, K., Kreneva, R.A., Perumov, D.A., and Nudler, E. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 111:747–756.

Morita, M.T., Tanaka, Y., Kodama, T.T., Kyogoku, Y., Yanagi, H., and Yura, T. 1999. Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. *Genes Dev* 13:655–665.

Murphy, B.A., Grundy, F.J., and Henkin, T.M. 2002. Prediction of gene function in methylthioadenosine recycling from regulatory signals. *J Bacteriol* 184:2314–2318.

Nahvi, A., Sudarsan, N., Ebert, M.S., Zou, X., Brown, K.L., and Breaker, R.R. 2002. Genetic control by a metabolite binding mRNA. *Chem Biol* 9:1043–1049.

Noeske, J., Richter, C., Grundl, M.A., Nasiri, H.R., Schwalbe, H., and Wohnert, J. 2005. An intermolecular base triple as the basis of ligand specificity and affininty in the guanine- and adenine-sensing riboswitch RNAs. *Proc Natl Acad Sci USA* 102:1372–1377.

Nou, X. and Kadner, R.J. 2000. Adenosylcobalamin inhibits ribosome binding to *btuB* RNA. *Proc Natl Acad Sci USA* 97:7190–7195.

Putzer, H., Condon, C., Brechemier-Baey, D., Brito, R., and Grunberg-Manago, M. 2002. Transfer RNA-mediated antitermination *in vitro*. *Nucl Acids Res* 30:3026–3033.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. 2002. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J Biol Chem* 277:48949–48959.

Romby, P. and Springer, M. 2003. Bacterial translational control at atomic resolution. *Trends Genet* 19:155–161.

Romeo, T. 1998. Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol Microbiol* 29:1321–1330.

Sarsero, J.P., Merino, E., and Yanofsky, C. 2000. A *Bacillus subtilis* operon containing genes of unknown function senses $tRNA^{Trp}$ charging and regulates expression of the genes of tryptophan biosynthesis. *Proc Natl Acad Sci USA* 97:2656–2661.

Serganov, A., Yuan, Y.R., Pikofskaya, O., Polonskala, A., Malinina, L., Phan, A.T., Hobartner, C., Micura, R., Breaker, R.R., and Patel, D.J. 2004. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem Biol* 11:1729–1741.

Serganov, A., Polonskaia, A., Phan, A.T., Breaker, R.R., and Patel, D.J. 2006. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* 441:1167–1171.

Storz, G., Altuvia, S., and Wassarman, K.M. 2005. An abundance of RNA regulators. *Annu Rev Biochem* 74:199–217.

Sudarsen, N., Wickiser, J.K., Nakamura, S., Ebert, M.S., and Breaker, R.R. 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev* 17:2688–2697.

Suess, B., Hanson, S., Berens, C., Fink, B., Schroeder, R., and Hillen, W. 2003. Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res* 31:1853–1858.

Tempest, D.W., Meers, J.L., and Brown, C.M. 1970. Influence of environment on the content and composition of microbial free amino acid pools. *J Gen Microbiol* 64:171–185.

Thore, S., Leibungut, M., and Ban, N. 2006. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* 312:1208–1211.

Tucker, B.J. and Breaker, R.R. 2005. Riboswitches as versatile gene control elements. *Curr Opin Struct Biol* 15:342–348.

Valbuzzi, A. and Yanofsky, C. 2001. Inhibition of the *B. subtilis* regulatory protein TRAP by a TRAP-inhibitory protein, AT. *Science* 293:2057–2059.

Vitreschak, A.H., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res* 30:3141–3151.

Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. 2003. Regulation of the vitamin B_{12} metabolism and transport in bacteria by a conserved RNA structural element. *RNA* 9:1084–1097.

Wabiko, H., Ochi, K., Nguyen, D.M., Allen, E.R., and Freese, E. 1988. Genetic mapping and physiological consequences of *metE* mutations of *Bacillus Subtilis*. *J Bacteriol* 170:2705–2710.

Weilbacher, T., Suzuki, K., Dubey, A.K., Wang, X., Gudapati, S., Morozov, I., Baker, C.S., Georgellis, D., Babitzke, P., and Romeo, T. 2003. A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol Microbiol* 48:657–670.

Winkler, W.C., Cohen-Chalamis, S., and Breaker, R.R. 2002. An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci USA* 99:15908–15913.

Winkler, W.C., Nahvi, A., and Breaker, R.R. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* 419:952–956.

Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428:281–286.

Winkler, W.C., Nahvi, A., Surarsan, N., Barrick, J.E., and Breaker, R.R. 2003. An mRNA structure that controls gene expression by binding *S*-adenosylmethionine. *Nat Struct Biol* 10:701–707.

Yanofsky, C. 1981. Attenuation in the control of expression of bacterial operons. *Nature* 289:751–758.

Yanofsky, C. 2004. The different roles of tryptophan transfer RNA in regulating *trp* operon expression in *E. coli* versus *B. subtilis*. *Trends Genet* 20:367–374.

Yousef, M.R., Grundy, F.J., and Henkin, T.M. 2005. Structural transitions induced by the interaction between $tRNA^{Gly}$ and the *Bacillus subtilis glyQS* T box leader RNA. *J Mol Biol* 349:273–287.

Editor: James Imlay