Regulation of Bacterial Gene Expression by Transcription Attenuation

Charles L. Turnbough, Jr.*

*Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, USA

SUMMARY
A wide variety of mechanisms that control gene expression in bacteria are based on conditional transcription termination. Generally, in these mechanisms, a transcription terminator is located between a promoter and a downstream gene(s), and the efficiency of the terminator is controlled by a regulatory effector that can be a metabolite, protein, or RNA. The most common type of regulation involving conditional termination is transcription attenuation, in which the primary regulatory target is an essential element of a single terminator. The terminator can be either intrinsic or Rho dependent, with each presenting unique regulatory targets. Transcription attenuation mechanisms can be divided into five classes based primarily on the manner in which transcription termination is rendered conditional. This review summarizes each class of control mechanisms from a historical perspective, describes important examples in a physiological context and the current state of knowledge, highlights major advances, and discusses expectations of future discoveries.

KEYWORDS RNA-binding proteins, Rho-dependent termination, transcription attenuation, bacterial gene regulation, intrinsic termination, reiterative transcription, riboswitches
INTRODUCTION

With the establishment of the operon model in bacteria in the 1960s, influential scientists pressed the idea that all gene expression would be negatively regulated by a repressor protein that bound to the promoter and prevented transcription initiation by RNA polymerase (RNAP) (1). Strict adherence to this idea hindered the acceptance of the next major advancement in the study of bacterial gene regulation, namely, positive control, in which an activator protein bound near the promoter to stimulate transcription initiation. The lack of acceptance of positive control occurred even though it shared many features of negative regulation. It is not surprising, therefore, that the discovery of the next new general mechanism of bacterial gene regulation took nearly two more decades, because it did not resemble in any way previously discovered examples of negative or positive control. The new mechanism was based on conditional transcription termination at a site (the terminator) located between a promoter and a downstream gene(s), with the efficiency of termination controlled by a specific physiological effector that interacted with the nascent RNA transcript (2). As a consequence, premature termination occurs when downstream gene activities are not needed by the cell. Even more surprising was the discovery, over time, that the examples of this new type of gene regulation employed many different ways to make transcription termination conditional in response to a multitude of physiological effectors (3). This diversity allows the regulation of expression of genes involved in many aspects of cellular metabolism in a wide variety of bacteria.

The discovery of gene regulation by conditional transcription termination was closely linked to seminal studies in the 1970s on the mechanisms of transcription termination in bacteria (4, 5). These studies revealed two distinct mechanisms, now called intrinsic and Rho-dependent termination (6), both of which can participate in control of gene expression. The role of each termination mechanism is to disrupt a highly stable transcription elongation complex (EC) that includes RNAP, DNA that is partially unwound to form a roughly 12-residue transcription bubble, and approximately 14 nucleotides (nt) of newly transcribed RNA (6). Within the active cleft of RNAP, the 3'-most region of the nascent RNA transcript forms a 9- to 10-bp RNA-DNA hybrid with the template strand of the transcription bubble, while ~5 residues of adjacent transcript fill the RNA exit channel of RNAP (7). The stability of the EC is due in large part to contacts between RNAP and both the RNA-DNA hybrid and single-stranded RNA in the exit channel (8). Disruption of these interactions and the extraction of the RNA from the EC are required for both intrinsic and Rho-dependent transcription termination, although these objectives are accomplished differently in the two mechanisms. Understanding the molecular details of these mechanisms is critical to understanding how termination is rendered conditional in gene regulation.

In intrinsic termination, a canonical DNA sequence specifies two RNA elements that are sufficient to cause RNAP to immediately terminate elongation and release the nascent transcript (6). The upstream region of the conserved DNA sequence is an approximately 20-bp hyphenated dyad symmetry, which specifies a stem-loop structure called the terminator hairpin. The stem of the terminator hairpin is G+C rich and usually contains at least 8 bp. The hyphenated dyad symmetry is immediately followed by a tract of typically 8 T residues, or mostly T residues, in the nontemplate strand, which specifies U or mostly U residues in the RNA. The first step in intrinsic termination is transcription pausing at the end of the T/U tract. This pause provides time for the terminator hairpin to form in the RNAP exit channel, which disrupts RNAP-nucleic acid interactions that stabilize the EC and causes 3 to 4 bp of the upstream RNA-DNA hybrid to melt. With these (and perhaps other) structural perturbations and with the remainder of the RNA-DNA hybrid consisting of only exceptionally weak rU-dA base pairs (9), the nascent transcript is free to dissociate from the EC.

In contrast to intrinsic termination, Rho-dependent termination requires a factor called Rho, a homohexameric protein with both RNA translocase and helicase activities. In addition to Rho, there are two other elements needed for termination: an approxi-
mately 80- to 90-nucleotide-long Rho-binding site called the rut (Rho utilization) site in the nascent RNA transcript and a transcription pause site specified by downstream DNA (6). The first step in termination is the binding of Rho to the rut site, which is followed by translocation in the 5’ to 3’ direction on the nascent transcript via ATP hydrolysis. During translocation, Rho remains bound to the rut site and threads RNA through its central pore, a process called tethered tracking (10). When Rho encounters a paused RNAP, it apparently causes disruptive conformational changes in the EC and uses its helicase activity to shear the RNA-DNA hybrid, thereby extracting the nascent RNA transcript (6).

Most mechanisms of gene regulation based on conditional transcription termination fall into two categories called transcription attenuation (or simply attenuation) and antitermination (4, 11). The major differences between these two types of control mechanisms are the number of terminators regulated and the nature of the regulatory target. In attenuation, typically one transcription terminator is involved, and it is directly targeted by the control mechanism. In antitermination, a large number of terminators are affected, and the target is RNAP. In this review, I focus only on attenuation control and refer the reader to two excellent reviews to learn about antitermination (12, 13). The goals of this review are to summarize attenuation control mechanisms from a historical perspective; to define general mechanistic classes; and to describe representative examples in a physiological context, highlight major advances, and discuss expectations of future discovery.

TRANSCRIPTION ATTENUATION

The term transcription attenuation arose from studies by Takashi Kasai on histidine (his) operon expression in Salmonella enterica serovar Typhimurium (referred to here as Salmonella) in which he proposed that an “attenuator” region in the DNA downstream of the promoter blocked transcription into the genes encoding the histidine-biosynthetic enzymes (14). It was subsequently shown that this attenuator was in fact a transcription terminator, and the term attenuator is now generally used to define a terminator that functions conditionally in gene regulation. Today, there are many examples of gene regulation by transcription attenuation, but they can be divided into five classes based on common themes in the regulatory mechanisms.

This division is somewhat arbitrary, however, due to overlap in mechanistic elements. The first four classes involve regulation of an intrinsic transcription terminator.

CLASS I: RIBOSOME-MEDIATED TRANSCRIPTION ATTENUATION

Tryptophan Operon of E. coli

Pioneering studies by Charles Yanofsky and colleagues in the 1970s led to the discovery of transcription attenuation control of the trpEDCBA (or simply trp) operon of Escherichia coli, which encodes the enzymes required for synthesis of L-tryptophan (Trp) (15, 16). The key elements required for attenuation control of trp operon expression are found in a 162-bp leader region, which is defined as the region between the transcription start site and the first gene of the operon (Fig. 1A). These elements include four segments in the leader transcript (i.e., segments 1, 2, 3, and 4) capable of forming two alternative transcript conformations. The first conformation includes an upstream 1:2 hairpin, which can induce transcription pausing, and a downstream 3:4 hairpin, which is the G+C-rich hairpin of an intrinsic transcription terminator (the trp attenuator). The alternative conformation contains only a 2:3 hairpin, the formation of which precludes formation of the 3:4 terminator hairpin, thereby permitting readthrough transcription into the genes of the trp operon. The final regulatory element is a 14-codon open reading frame (ORF) that extends through the end of leader transcript segment 1. The leader ORF also contains two Trp “control” codons at positions 10 and 11.

According to the regulatory model (17), RNAP initiates transcription at the trp promoter and moves rapidly through the leader region that specifies transcript segments 1 and 2, which then form the 1:2 pause hairpin. RNAP stalls at this site to permit a ribosome to begin translation of the leader ORF. Early in translation, the ribosome
physically disrupts the 1:2 hairpin to release the stalled RNAP, and the ribosome then proceeds to the Trp control codons. When the intracellular level of Trp is limiting, causing a low level of Trp-tRNATrp, the ribosome pauses at the tandem Trp codons (Fig. 1B). During this time, the reengaged RNAP synthesizes transcript segment 3, permitting formation of the 2:3 antiterminator hairpin. As transcription continues, the nascent transcript is extended through leader segment 4 (without formation of the 3:4 terminator hairpin) and eventually through the entire operon. Translation of the full-length trp mRNA produces the biosynthetic enzymes needed to overcome Trp limitation. Conversely, when there is ample intracellular Trp and Trp-tRNATrp, the ribosome translating the leader transcript does not pause at the tandem Trp codons but proceeds to, and effectively pauses at, the stop codon at the end of the leader ORF (Fig. 1C) (18).

At this position, the ribosome occupies leader transcript segments 1 and 2 (19, 20). Continuing transcription synthesizes leader transcript segments 3 and 4 and an adjacent 8-residue poly(U) tract, which allows formation of the 3:4 terminator hairpin and transcription termination at the end of the U tract. Consequently, additional trp-biosynthetic enzymes are not synthesized when there is sufficient Trp to support optimal cell growth. Attenuation control regulates trp operon expression over an approximately 8-fold range. In E. coli, inhibition of transcription initiation by the Trp repressor independently controls trp expression over a 70- to 80-fold range. Together, attenuation and repression control mechanisms regulate the expression of the trp

FIG 1 Transcription attenuation control of trp operon expression in E. coli. (A) Key regulatory elements in the trp leader region. The leader region contains four segments (1, 2, 3, and 4) capable of specifying hairpins 1:2, 2:3, and 3:4 in the leader transcript. The leader region also contains a 14-codon ORF with Trp codons at positions 10 and 11 that extends through segment 1, and the trp attenuator in which segments 3 and 4 specify the terminator hairpin. Formation of the 1:2 hairpin induces transcription pausing, which is released by disruption of the hairpin by a ribosome translating the leader ORF. (B) When Trp (and Trp-tRNATrp) levels are low, the ribosome translating the leader ORF pauses at the tandem Trp codons in segment 1, while transcription continues downstream. With segment 1 sequestered, the 2:3 antiterminator hairpin forms immediately, which precludes formation of the 3:4 terminator hairpin and allows readthrough transcription. (C) When Trp (and Trp-tRNATrp) levels are high, translation proceeds through the entire leader ORF, which allows the ribosome to physically cover both segments 1 and 2 in the transcript. With segment 2 sequestered, further downstream, transcription allows the 3:4 terminator hairpin to form, which leads to transcription termination at the end of the adjacent U tract.
operon over an approximately 600-fold range in response to intracellular levels of Trp (21).

Although attenuation control of trp operon expression in E. coli is generally considered the first example of this type of regulation, concurrent studies by Bruce Ames, John Roth, and their colleagues led to the discovery of an analogous mechanism responsible for attenuation control of his operon expression in Salmonella (22, 23). The major difference between the trp and his mechanisms is that, in the case of the his operon, the control codons are 7 histidine codons in a row (24). Within the decade after the elucidation of the trp and his attenuation control mechanisms, similar mechanisms were discovered for several other amino acid-biosynthetic operons and an aminoacyl-tRNA synthetase operon of enteric bacteria (17). In each case, ribosome stalling at cognate control codons regulated the formation of alternative secondary structures in the leader transcript, one permitting and the other precluding intrinsic transcription termination.

The discovery of these first examples of attenuation control was extremely exciting; however, the similarities between the control mechanisms raised the possibility that this new type of gene regulation was limited to amino acid-related operons and to a single mechanism for regulating transcription termination. This idea was soon dispelled by the discovery of a mechanistically different type of attenuation control involved in regulating pyrBI expression in E. coli.

**pyrBI Operon of E. coli**

The pyrBI operon encodes the catalytic (pyrB) and regulatory (pyrI) subunits of the pyrimidine nucleotide-biosynthetic enzyme aspartate transcarbamoylase. This enzyme catalyzes the first committed step in pyrimidine nucleotide biosynthesis, and its activity is allosterically regulated by pyrimidine nucleotide inhibitors and purine nucleotide activators (25). Expression of the pyrBI operon is negatively regulated over an approximately 350-fold range by the intracellular level of UTP, with most of this regulation (i.e., 50-fold) occurring through a transcription attenuation control mechanism (26). This mechanism relies on three regulatory elements located in a 156-bp leader region preceding the pyrB gene: a 44-codon ORF, an intrinsic transcription terminator (the pyrB attenuator) located within and near the end of the ORF, and UTP-sensitive transcription pause sites (i.e., sites for U addition, often as U tracts, in the nascent transcript) within the ORF and preceding the attenuator (Fig. 2A). Transcription pausing at U residues, but not at non-U residues, is due to the unique ability of the intracellular UTP pool to fall (e.g., during pyrimidine starvation) to levels that significantly slow transcription elongation (27).

According to the model of attenuation control, transcription initiates constitutively at the pyrBI promoter and continues into the leader region (28, 29). When the level of UTP in the cell is low, RNAP stalls at the UTP-sensitive pause sites, which provides time for a ribosome to begin translation of the leader ORF and proceed up to the stalled RNAP, establishing tightly coupled transcription and translation. When RNAP eventually escapes the pause sites and transcribes the attenuator, formation of the attenuator-specified terminator RNA hairpin is physically blocked by the adjacent translating ribosome (Fig. 2B). In this case, transcription termination at the attenuator is precluded, and RNAP continues transcription into the pyrBI genes. The result is an increase in the capacity of the cell to produce more UTP under conditions of UTP limitation. In contrast, when the level of UTP in the cell is high, RNAP transcribes the leader region without stalling at the UTP-sensitive pause sites. Consequently, there is insufficient time for a ribosome to translate the leader ORF and establish tight coupling with RNAP before the formation of the terminator hairpin and transcription termination at the attenuator (Fig. 2C). The result is that additional aspartate transcarbamoylase is not produced when there is no need for an increase in intracellular UTP production.

Shortly after the mechanism for attenuation control of pyrBI expression was elucidated, an analogous UTP-sensitive attenuation mechanism that controls expression of the pyrE gene of E. coli was discovered (30). The pyrE gene encodes the pyrimidine
nucleotide-biosynthetic enzyme orotate phosphoribosyltransferase, and its expression is regulated over an approximately 30-fold range by transcription attenuation. The major difference between the pyrBI and pyrE control mechanisms is that the pyrE “leader ORF” is actually the 238-codon rph gene, which encodes the tRNA-processing exoribonuclease RNase PH (31). Thus, the pyrE gene is the second gene of an rph-pyrE operon, and UTP-sensitive transcription pausing and tight coupling of transcription and translation within the rph ORF are used to prevent intrinsic transcription termination at an attenuator preceding the pyrE gene. It is noteworthy that the rph and pyrE genes appear to be metabolically unrelated and that the pyrE attenuator can double as a terminator for a monocistronic rph transcript. Another interesting difference between the pyrBI and pyrE mechanisms is that the pyrE attenuator is not in the rph ORF but is located 10 bp downstream. Nevertheless, based on the size of the ribosome footprint (~12 to 15 nt on either side of the ribosomal P site) (19, 32), translation to the end of the rph cistron would disrupt formation of the terminator hairpin and allow read-through transcription.

Historically, with the elucidation of the pyrBI and pyrE attenuation control mechanisms, there were now two distinct mechanisms for rendering transcription termination conditional, namely, formation of an alternative leader transcript secondary structure that excludes the terminator hairpin and tightly coupled transcription and translation that allow a translating ribosome to physically interfere with terminator hairpin formation. However, in both cases, attenuation control was mediated by a translating ribosome, again suggesting a limitation in the scope of attenuation control mechanisms. However, the reality was that a treasure chest of attenuation control mechanisms was just beginning to open. The next examples of attenuation control did not rely on a translating ribosome but employed RNA-binding proteins to control the formation of the terminator hairpin.
CLASS II: RNA-BINDING-PROTEIN-MEDIATED TRANSCRIPTION ATTENUATION

bgl Operon of *E. coli*

The first example of transcription attenuation control mediated by an RNA-binding protein was provided by the mechanism for β-glucoside-sensitive regulation of *bgl* operon expression in *E. coli* (33–35). The *bgl* operon contains four genes, *bglG, bglF, bglB,* and *bglH,* which encode proteins that enable cells to use certain aromatic β-glucosides as carbon sources (Fig. 3A). Two of the proteins, BglG and BglF, are involved in regulation. As a dimer, BglG is an RNA-binding protein that positively regulates operon expression. BglF is a membrane-bound, β-glucoside-specific enzyme II of the phosphoenolpyruvate-dependent phosphotransferase system (PTS) that catalyzes transport and concomitant phosphorylation of β-glucosidic sugars (Fig. 3B) (36). BglF also catalyzes reversible phosphorylation of dimeric BglG to produce monomeric and membrane-bound BglG−P that is unable to bind the nascent *bgl* transcript and disrupt terminator hairpin formation. Consequently, most transcripts initiated at the *bgl* promoter are terminated at either attenuator 1 or 2, and only minimal levels of BglG, BglF, BglB, and BglH are synthesized.

**FIG 3** Basic mechanism of BglG-mediated attenuation control of *bgl* operon expression in *E. coli.* (A) Regulatory elements of the *bgl* operon including the attenuators and partially overlapping RAT sites preceding the *bglG* and *bglF* genes. (B) In the presence of a β-glucoside substrate, BglF-catalyzed dephosphorylation of monomeric BglG−P allows formation of active (i.e., RNA-binding-proficient) BglG dimers. These dimers bind to and stabilize RAT secondary structures in the nascent leader transcript that prevent terminator hairpin formation at attenuators 1 and 2. The result is readthrough transcription at each site. (C) In the absence of β-glucosides, BglF-catalyzed phosphorylation of dimeric BglG produces monomeric and membrane-bound BglG−P that is unable to bind the nascent *bgl* transcript and disrupt terminator hairpin formation. Consequently, most transcripts initiated at the *bgl* promoter are terminated at either attenuator 1 or 2, and only minimal levels of BglG, BglF, BglB, and BglH are synthesized.
BglH) are synthesized. On the other hand, in the presence of β-glucosides, BglF acquires phosphate residues for sugar phosphorylation from BglG~P, which releases unphosphorylated BglG into the cytoplasm, where it dimerizes (40). Dimeric BglG then binds independently to specific transcript sites called RAT (ribonucleic antiterminator) sites that precede and partially overlap each attenuator-specified terminator hairpin (Fig. 3A), thereby preventing termination at attenuators 1 and 2 and allowing expression of the operon and β-glucose catabolism (Fig. 3B). The basic model became a bit more complicated with the discovery that BglG is phosphorylated at a second site and that this phosphorylation is required for dimeric BglG binding to RNA. Phosphorylation at the second site is catalyzed by the PTS enzyme HPr, which is not specific for a particular sugar, and its paralog, FruB. Accordingly, when other PTS sugars become available in addition to β-glucosides, second-site phosphorylation of BglG is restricted and BglG activity is suppressed (43).

Overall, the critical feature in attenuation control of bgl expression is the binding of dimeric BglG to RAT sites in the nascent transcript, which physically blocks formation of attenuator-encoded terminator hairpins. RAT site sequences form an alternative RNA hairpin that is stabilized by the binding of BglG (Fig. 3B) (41, 44, 45). The RNA-binding activity of BglG is regulated through antagonistic phosphorylation events: one ensures that high levels of the enzymes required for β-glucose catabolism are produced only when β-glucose substrates are available, and the other couples bgl expression to the general carbohydrate supply. Subsequent to the discovery of attenuation control of bgl expression in E. coli, several analogous regulatory mechanisms for operons involved in PTS-transported sugar catabolism and employing a BglG-like antiterminator protein were found in other bacteria, especially Gram-positive bacteria (46). Additionally, a steadily growing number of other attenuation control mechanisms that rely on unique RNA-binding proteins to control transcription termination have been discovered. Two of the best studied are TRAP and PyrR of Bacillus subtilis.

**TRAP and PyrR of B. subtilis**

Expression of the B. subtilis trpEDCFBA (or trp) operon, which contains six of the seven genes needed for Trp biosynthesis, is regulated over an approximately 100-fold range by Trp-sensitive transcription attenuation (47). A 203-bp leader region can specify two alternative RNA hairpins, an antiterminator (A:B) hairpin and a downstream (C:D) terminator hairpin specified by an intrinsic transcription terminator (the trp attenuator) (Fig. 4). The two hairpins are mutually exclusive due to a 4-base overlap of leader transcript segments B and C (48). An RNA-binding protein called TRAP (trp RNA-binding attenuation protein) controls formation of the alternative hairpins. TRAP is composed of 11 identical 75-amino-acid subunits, each encoded by the mtrB gene (49), which assemble into a β-wheel with a large central hole (50). TRAP is activated to bind RNA by positive cooperative binding of up to 11 molecules of Trp (51). The TRAP binding site in the leader transcript consists of 11 repeats of GAG or UAG separated by 2- or 3-nt nonconserved spacers (52). The triplet repeats are located between leader residues 36 and 91, which overlap segment A and the 5′ portion of segment B of the antiterminator hairpin (Fig. 4A). When bound to TRAP, the leader transcript wraps around the outside of the protein ring (50, 52).

According to the standard model for attenuation control (44), transcription initiation at the B. subtilis trp promoter occurs constitutively, and transcript elongation proceeds into the leader region. Under conditions of limiting Trp, TRAP is not activated and does not bind the nascent leader transcript, which permits the antiterminator hairpin to form as soon as segments A and B are synthesized. As transcription proceeds downstream, formation of the terminator hairpin is precluded, allowing transcription into and expression of the structural genes of the operon (Fig. 4A). Alternatively, when cells are growing in the presence of excess Trp, TRAP is activated and binds the (G/U)AG repeat region of the nascent transcript, inhibiting formation of the A:B antiterminator hairpin. This inhibition permits formation of the alternative C:D terminator hairpin, which leads...
to premature leader transcript termination at the trp attenuator and suppression of trp operon expression (Fig. 4B).

Over the years, the standard model has been expanded to include regulatory roles for leader features, such as a 5′ stem-loop (5′ SL) and NusA- and NusG-stimulated transcription pausing (54, 55), which enhance TRAP binding, and a role for an anti-TRAP protein, the production of which is induced by uncharged tRNATrp (56, 57). More recent studies even suggest a second role for TRAP in attenuation control. The trp leader terminator hairpin is only 40% G+C rich, which should make the trp attenuator a weak transcription terminator. However, transcription termination at the trp attenuator is highly efficient in vivo. To account for this apparent contradiction, it has been proposed that TRAP, in addition to its role in controlling leader hairpin formation, interacts with RNAP and perhaps other cellular factors to directly stimulate transcription termination at the trp attenuator (58, 59). Although not directly related to attenuation control, TRAP also represses translation initiation of trpE and several other genes involved in tryptophan metabolism (60).

Like TRAP, the PyrR protein of B. subtilis is an RNA-binding protein that regulates transcription termination, but their mechanisms of action are quite distinct. PyrR regulates expression of the pyrRPBCD(AA)(AB)KDFE (or pyr) operon, the last eight genes of which encode the six enzymes required for de novo synthesis of UMP (29, 61). The first and second genes of the operon, pyrR and pyrP, encode PyrR and uracil permease, respectively. The operon also contains three similar-strength intrinsic transcription terminators that function as attenuators (numbered 1, 2, and 3), the first located in the leader region and the second and third in untranslated regions between pyrR and pyrP and pyrP and pyrB, respectively (Fig. 5A). The RNA specified by all three of these untranslated regions is also capable of folding into a large antiterminator hairpin that sequesters residues in the 5′ segment of each downstream terminator hairpin, thereby preventing termination at each attenuator (Fig. 5B). Furthermore, at each location, the RNA embedded in the 5′ half of each antiterminator can form a third hairpin called the
antiantiterminator, which precludes antiterminator formation and thus allows termination at the attenuator (62). Each antiantiterminator hairpin includes structures and conserved sequences that constitute a PyrR binding site (Fig. 5C) (63, 64). PyrR binding to these sites is stimulated by uridine nucleotides and phosphoribosyl pyrophosphate (PRPP) and antagonized by guanosine nucleotides, all at physiological concentrations (63, 65, 66). Binding of these effectors also affects the aggregation state of PyrR, with activator and inhibitor binding resulting in dimer and tetramer formation, respectively (67).

In the model for regulation, transcription of the pyr operon is initiated at a constitutive promoter, and expression of the pyrimidine nucleotide-biosynthetic genes is negatively regulated over a ≥200-fold range by uridine nucleotide-sensitive attenuation (68). When intracellular levels of uridine nucleotides (e.g., UMP and UTP) are low, PyrR forms tetramers (or hexamers) that do not bind RNA (69). In this case, antiterminator hairpin formation is favored in the nascent transcript, which prevents transcription termination and allows readthrough transcription into the downstream gene(s). (C) At high levels of uridine nucleotides, UMP/UTP-bound dimeric PyrR binds to and stabilizes the antiantiterminator hairpin (including segments 1 and 2), which promotes formation of the attenuator-encoded terminator hairpin (including segments 3 and 4) and immediate transcription termination. See the text for additional details.

It was also observed that guanosine (e.g., GMP and GTP) and uridine nucleotides compete for the same PyrR binding site (65, 67). Thus, a high ratio of guanine to uridine nucleotides results in guanine nucleotide-bound PyrR, which forms non-RNA-binding tetramers. In this case, expression of the pyr operon is increased. This effect provides a
means of coordinating the rate of pyrimidine nucleotide biosynthesis with the size of intracellular guanine nucleotide pools. Additionally, PRPP stimulates PyrR dimerization and binding to antiantiterminator hairpins, which leads to reduced pyr operon expression. This response makes physiological sense, because an increase in the intracellular PRPP level is a consequence of guanosine nucleotide shortage and hence a diminished demand for pyrimidine nucleotide synthesis. One cautionary note about the PyrR regulatory models described above is that correlation between the aggregation state of PyrR and RNA binding is just the current hypothesis; it has not been rigorously demonstrated.

Another interesting aspect of PyrR is that it possesses uracil phosphoribosyltransferase activity; however, the activity is weak and apparently plays a negligible role in uracil salvage in vivo (70). A close structural resemblance of GMP-bound PyrR to an evolutionarily distant hypoxanthine-guanine phosphoribosyltransferase suggests that the two enzymes share ancestry (65). Such a relationship might explain the ability of PyrR to bind guanine nucleotides. It seems clear, though, that PyrR was derived from a phosphoribosyltransferase that later acquired the ability to bind RNA. Finally, homologs of PyrR have been identified in several hundred distinct bacterial species. In many cases, these homologs appear to be involved in transcription attenuation (29).

Although the discovery of attenuation control based on RNA-binding proteins was a major advance in the field, the next class of control mechanisms revolutionized the study of gene regulation. These mechanisms relied on a new type of RNA called a riboswitch.

**CLASS III: RIBOSWITCH-MEDIATED TRANSCRIPTION ATTENUATION**

**Definition of a Riboswitch**

A riboswitch is broadly defined as a cis-acting regulatory segment of an mRNA that directly senses a physiological signal, which causes a change in RNA structure that impacts downstream gene expression (71). In many instances, the change in RNA structure controls the formation of an attenuator-specified terminator hairpin (72). When the term riboswitch was first coined by Ron Breaker in 2002, it was defined more narrowly to mean a regulatory segment of mRNA that specifically binds a small-molecule cellular metabolite (73). However, the first example of a riboswitch that meets the broad definition of the term was discovered a decade earlier, and in this case, the mRNA-binding ligand was a tRNA (74–76).

**tRNA-Binding Riboswitches**

In 1992, analysis of the *B. subtilis* tyrS operon, which encodes tyrosyl-tRNA synthetase, indicated that tyrS expression was regulated by tyrosine-sensitive transcription attenuation, with conditional termination occurring at an intrinsic terminator located near the end of an ~300-bp leader region. Transcription past the attenuator was dependent on a 14-base transcript sequence, designated the T box, which was located immediately upstream of the terminator hairpin (Fig. 6A) (75). It was quickly realized that T-box sequences and downstream intrinsic terminators were present in long leader regions of numerous aminoacyl-tRNA synthetase and amino acid-biosynthetic operons in different *Bacillus* species, suggesting a common attenuation control mechanism (74). Additional analysis revealed that the leader transcripts of each of these operons were capable of forming three conserved stem-loop structures (I, II, and III) that preceded the T box. Each stem I contains a bulge region that includes a triplet sequence called the specifier sequence, which is actually a codon that specifies the amino acid associated with the operon being regulated (Fig. 6A). This result suggested that the specifier sequence is part of a binding site for a cognate tRNA, which was confirmed by further experimentation (77). Inspection of sequences downstream of stem III showed that the region could form an antiterminator hairpin by base pairing a segment of the T-box sequence with residues in the upstream portion of the terminator hairpin (Fig. 6B). In each antiterminator hairpin, the central 7 bases of the T-box sequence formed a bulge that included the sequence UGGN, which is complementary to the NCCA sequence.
present at the 3’ acceptor ends of cognate tRNAs (Fig. 6B). Interaction of these 4-base sequences was subsequently shown to provide a second leader binding site for an uncharged— but not a charged— cognate tRNA (77).

Based on the features described above, a novel mechanism to regulate expression of amino acid-related operons in Bacillus and closely related species was proposed (74, 77). According to the model, when the cell has ample levels of a particular amino acid, its cognate tRNA is efficiently aminoacylated. This charged tRNA can bind to the specifier sequence but not to the T-box bulge, which prevents stabilization of the antiterminator hairpin. Consequently, the terminator hairpin forms, and the leader transcript is terminated at the attenuator (Fig. 6A). Conversely, when the cell is starved for the amino acid, cognate uncharged tRNA accumulates and interacts with both tRNA-binding sites of the leader region. This binding stabilizes formation of the antiterminator hairpin (78), which allows transcription past the attenuator and into the structural genes of the operon (Fig. 6B). In this model, the regulatory mechanism monitors the ratio of charged to uncharged tRNA and, in response, adjusts the level of associated operon expression to meet the physiological needs of the cell (71). This mechanism has now been shown to be widely used by members of the phyla Firmicutes and Actinobacteria, with the number of examples per genome ranging from 1 to 40.
More recent studies have focused on identifying additional tRNA-leader transcript contacts and contacts that are tRNA specific (Fig. 6) (81–84).

**Small-Molecule-Binding Riboswitches**

Most examples of riboswitches bind molecules much smaller than a tRNA, with the first of these identified about 20 years ago (85–87). To date, more than 40 different classes of riboswitches that bind a variety of metabolites, coenzymes, signaling molecules, and inorganic ions have been discovered (88–90). These riboswitches control the formation of alternative leader transcript secondary structures that often affect intrinsic transcription termination, although they can also control translation initiation and RNA stability and splicing (91). When riboswitches control transcription termination, they dictate the formation of an RNA secondary structure that either includes or excludes the attenuator-encoded terminator hairpin. Leader regions that include riboswitches are atypically long (usually >200 nt) and contain aptamers that bind a target molecule with high selectivity and specificity.

One of the earliest examples of a small-molecule-binding riboswitch was the mechanism controlling expression of the riboflavin-biosynthetic operon ribDEAHT in *B. subtilis* (92). Riboflavin is a precursor of flavin mononucleotide (FMN), and *ribDEAHT* expression is controlled by the binding of FMN to an aptamer (called an *rfn* box) formed by the *ribD* leader transcript. When FMN levels are low, it fails to bind its aptamer, and the *ribD* leader transcript adopts a secondary structure that includes an antiterminator stem (Fig. 7A). The downstream segment of this stem is shared with a potential attenuator-encoded terminator hairpin located further downstream. Therefore, formation of the antiterminator stem precludes formation of the terminator hairpin, allowing readthrough transcription into the structural genes of the operon. At high levels, FMN binds its aptamer, which promotes the formation of an alternative leader transcript secondary structure that includes the terminator hairpin (Fig. 7B). As a result, transcrip-
tion is terminated upstream of the structural genes of the operon under conditions of ample riboflavin and FMN. FMN riboswitches like the one described above are now known to control expression of genes responsible for the biosynthesis and transport of riboflavin and FMN in many bacteria (88). In some cases, the structure of FMN bound by its aptamer has been determined (Fig. 7C).

Pioneered by the Breaker laboratory, a productive strategy for discovering small-molecule-binding riboswitches has been the use of computer algorithms to compare sequences of noncoding regions of bacterial genomes and the RNA structures that they encode (88, 93–98). This method uncovers variants of known riboswitches that can be subsequently verified by genetic, biochemical, and biophysical studies. The limitation of this method is that it fails to detect rare classes of riboswitches, and it is possible that many undiscovered riboswitches fall into this category (99).

CLASS IV: TRANSCRIPTION ATTENUATION WITHOUT AN RNA-BINDING ELEMENT
A Class of One, So Far, and Reiterative Transcription

All the examples of attenuation control described above employ an RNA-binding element; however, there is one known example where this is not the case. In this example, attenuation control relies on conditional (but permanent) modification of the leader transcript primary sequence by a noncanonical form of transcription called reiterative transcription (100). This reaction is defined by repetitive addition of nucleotides to the 3′ end of a nascent transcript due to one or more rounds of a 1-base upstream shift of the transcript without movement of the DNA template. Accordingly, the same nucleotide in the template specifies multiple (or extra) residues in the transcript (101). Reiterative transcription occurs primarily within a 3-base homopolymeric tract in the DNA template and can involve the repetitive addition of any nucleotide substrate of RNAP.

pyrG Operon of B. subtilis

Expression of the pyrG operon of B. subtilis, which encodes the pyrimidine-biosynthetic enzyme CTP synthetase, is regulated over a 20-fold range by a CTP-sensitive attenuation control mechanism that targets an intrinsic transcription terminator near the downstream end of a 189-bp leader region. The mechanism that renders the pyrG attenuator conditional relies on reiterative transcription that involves repetitive addition of G residues during transcription initiation (102). This reaction occurs during transcription of the pyrG initially transcribed region (Fig. 8A), which specifies the sequence 5′-GGGCU at the start of the pyrG transcript. Another key control element is an attenuator-encoded terminator hairpin in which most of the upstream segment of the stem consists of the 9-residue pyrimidine tract 5′-CUCCCUUUC (Fig. 8B), all of which can base pair with a poly(G) tract of RNA (Fig. 8C) (103).

According to the model for regulation (102, 104), when the intracellular level of CTP is high (reflecting ample pyrimidine availability), transcription of the pyrG leader region is rapid and produces transcripts that are copied base for base from the DNA template. However, these transcripts are efficiently terminated at the pyrG attenuator (Fig. 8B). Thus, when CTP is abundant, transcription of the pyrG gene is repressed. On the other hand, when the intracellular level of CTP is low (reflecting pyrimidine limitation), transcription of the pyrG leader region pauses at position +4 C (i.e., after the synthesis of the 5′-GGG transcript) because of insufficient substrate. The time provided by this pause allows the nascent transcript to slip upstream relative to the DNA template (i.e., into the 3′-CCC tract), which directs the addition of an extra G residue to the transcript. This process can be repeated up to at least 10 times, until eventually a C residue is added to the 3′ end of the transcript by canonical transcription. Transcript elongation then continues normally until RNAP transcribes the attenuator sequence specifying the upstream segment of the terminator hairpin, which includes the long pyrimidine tract. This tract immediately base pairs with the poly(G) tract at the 5′ end of the transcript to form an antiterminator hairpin (Fig. 8C). Formation of a highly effective antiterminator hairpin requires a minimum of three extra G residues in the poly(G) tract (103).
As RNAP continues to elongate the pyrG transcript, the antiterminator hairpin precludes terminator hairpin formation, which allows the synthesis of full-length transcripts. Translation of these transcripts produces more CTP synthetase, which is needed to overcome the CTP deficiency. Although the model is described in binary terms of high and low intracellular CTP concentrations, regulation can occur continuously over a range of CTP concentrations that affect the extent of pausing at position +4 (104).

Although attenuation control of pyrG expression in *B. subtilis* is the only example in its class to date, it is reasonable to expect that analogous mechanisms will be discovered. One difficulty in this process is that there are no computer algorithms that identify sites of reiterative transcription. Reiterative transcription that might be involved in attenuation control could be detected by the presence of extra nucleotides in the leader regions of sequenced RNA transcripts. However, the discovery of extra nucleotides is typically considered a sequencing error and dismissed as unimportant. There is also the possibility that attenuation control without an RNA-binding element could rely on something other than reiterative transcription, although that “something” remains to be discovered.

**CLASS V: TRANSCRIPTION ATTENUATION AT RHO-DEPENDENT TERMINATORS**

Finding Rho-Dependent Terminators, Their Expanding Role in Gene Regulation, and Familiar Regulators

Unlike intrinsic terminators, Rho-dependent terminators have been difficult to find because they lack a consensus sequence (6). For example, the only requirements for a rut site are that it is long enough to accommodate all six Rho subunits; that it is rich in cytidines, which preferentially bind Rho; and that it is free of ribosomes and extensive secondary structure that can preclude Rho binding (6, 106). Consequently, the discovery of a role for Rho in gene regulation was initially a slow process. However, with the advent in the last decade of high-throughput sequencing techniques capable of identifying the 3’ ends of RNA transcripts, the pace of discovery has quickened dramatically (107). Today, we recognize a variety of regulatory roles for Rho, including control of individual gene expression (106), as a housekeeper of genome-wide transcription (108, 109) and genome integrity (110), and even in the determination of cell fate (111).

Perhaps not surprisingly, the regulatory factors that mediate transcription attenuation at Rho-dependent terminators are often the same as those involved in making...
intrinsic terminators function conditionally. The difference is the target: the terminator hairpin of intrinsic attenuators and primarily the rut site for Rho-dependent attenuation. As with the first examples of attenuation involving an intrinsic terminator, the first example of Rho-dependent transcription attenuation employed ribosome positioning on the leader transcript to control termination, but with interesting twists.

Ribosome-Mediated Rho-Dependent Attenuation

The first well-studied example of Rho-dependent attenuation was the mechanism controlling Trp-induced (and also cAMP/CRP-dependent) expression of the Trp-degradative tnaCAB (or tna) operon of *E. coli* (112). The 320-bp leader region of the operon includes three features required for regulation: a 24-codon ORF designated *tnaC*, a *rut* site, and transcription pause sites, all of which precede the structural genes encoding tryptophanase (*tnaA*) and tryptophan permease (*tnaB*) (Fig. 9A). According to an early regulatory model, after synthesizing the *tnaC* segment of the nascent *tna* transcript, RNAP pauses in the *tnaC*-*tnaA* intergenic space to allow a ribosome to begin translating the leader ORF (113). When intracellular levels of Trp are low, translation of the leader region is completed, and the translating ribosome is released from the leader transcript, thereby exposing the *rut* site. Rho then binds to this site and translocates downstream until it contacts the stalled RNAP, and terminates transcription (Fig. 9B). On the other hand, when Trp levels are high, it binds to the nascent TnaC polypeptide while it is within the exit channel of the translating ribosome. This interaction causes the ribosome to stall at codon Pro24, which masks the *rut* site and prevents Rho-dependent transcription termination. Eventually, the stalled RNAP resumes transcription, allowing expression of the proteins capable of catabolizing excess Trp.

FIG 9 Mechanism of ribosome-mediated Rho-dependent attenuation control of *tna* expression in *E. coli.* (A) Organization of the *tna* promoter-leader region, including *tnaC*, with residues Trp12 and Pro24 marked, and the following spacer region that contains a *rut* site and downstream transcription pause sites. Following transcription through most of the leader region and RNAP stalling at a pause site, a ribosome begins translating the *tnaC* cistron. (B) When the level of Trp is low, the translating ribosome completes synthesis of the TnaC polypeptide and dissociates from the transcript. Rho factor then binds the exposed *rut* site, translocates downstream until it contacts the stalled RNAP, and terminates transcription. (C) When the level of Trp is high, it binds to the nascent TnaC polypeptide while it is within the exit channel of the translating ribosome. This interaction causes the ribosome to stall at codon Pro24, which masks the *rut* site and prevents Rho-dependent transcription termination. Eventually, the stalled RNAP resumes transcription, allowing expression of the proteins capable of catabolizing excess Trp.
This model describes how \textit{tna} expression occurs only when intracellular levels of Trp are high, but it does not explain how Trp levels control the extent of translation in the leader ORF. The answer was a huge surprise. Further studies revealed that when translation of the leader region reaches codon Pro24, particular residues of the nascent TnaC peptide (e.g., Trp12 and Ile19) and the 23S rRNA form a binding pocket for free Trp within the exit channel of the ribosome (114–116). Binding of Trp to this pocket causes the ribosome to stall, thereby masking the \textit{rut} site and preventing premature Rho-dependent termination within the leader region (Fig. 9C). The TnaC peptide was an early example of regulatory peptides called ribosome arrest peptides, which induce translational arrest to control either transcription or translation of downstream genes in the same operon. Today, there are numerous examples of ribosome arrest peptides that are involved in a variety of regulatory mechanisms (117, 118).

A recent and equally interesting example of ribosome-mediated Rho-dependent attenuation is the mechanism controlling expression of the \textit{mgtA} operon of \textit{Salmonella}, which encodes the Mg\textsuperscript{2+} transport protein MgtA and is regulated over a 100-fold range by the intracellular level of Mg\textsuperscript{2+} (119–121). The leader region of the operon includes a short, proline-rich ORF, \textit{mgtL}, and mutually exclusive secondary structures, only one of which has an exposed \textit{rut} site. Rapid and complete translation of \textit{mgtL} permits formation of the secondary structure with the exposed \textit{rut} site, resulting in Rho-dependent termination at a transcription pause site preceding the \textit{mgtA} gene (122). Efficient translation of \textit{mgtL}, specifically the included proline codons, is thought to require high concentrations of Mg\textsuperscript{2+}. This putative requirement apparently reflects the involvement of Mg\textsuperscript{2+}-dependent components of the translation machinery that are uniquely involved in translation of proline codons (123, 124). Thus, when Mg\textsuperscript{2+} levels are high, \textit{mgtA} expression is repressed. On the other hand, when Mg\textsuperscript{2+} levels are low, the proline codons in \textit{mgtL} presumably present an impediment to translation. The resulting slow translation directs formation of the leader secondary structure that occludes the \textit{rut} site, thereby preventing Rho-dependent termination and allowing \textit{mgtA} expression. This mechanism shares many key features with the archetypal \textit{trp} attenuation control mechanism of \textit{E. coli} (Fig. 1). In both cases, the efficiency of translating control codons in a leader ORF dictates transcript secondary structures that control the availability of an RNA element that is essential for transcription termination (i.e., a \textit{rut} site or terminator hairpin).

It should be noted that a somewhat different model for regulation of \textit{mgtA} expression proposes a different role for Mg\textsuperscript{2+} in controlling the formation of alternative transcript secondary structures. Instead of the indirect role proposed above, this alternative model proposes that it is the extent of direct Mg\textsuperscript{2+} binding to the leader transcript that dictates secondary-structure formation. In other words, the leader transcript is a multi-ion Mg\textsuperscript{2+}-sensitive riboswitch (125). Consistent with this model, changes in Mg\textsuperscript{2+} concentrations within the physiological range alter the secondary structure of the \textit{mgtA} leader transcript in ways that affect transcription elongation and termination, and the regulation by Mg\textsuperscript{2+} was recapitulated \textit{in vitro} using only DNA, nucleotides, and RNAP (i.e., no ribosomes) (120). In addition, replacing the proline codons in \textit{mgtL} with other sense codons did not eliminate the response to Mg\textsuperscript{2+} (126). Furthermore, ribosome profiling failed to show slow translation at proline codons at low Mg\textsuperscript{2+} concentrations (127). The differences in the two proposed mechanisms for attenuation control of \textit{mgtA} expression remain to be resolved, although it is possible that both mechanisms are correct.

**Protein- and Riboswitch-Mediated Rho-Dependent Attenuation**

In addition to the ribosome, RNA-binding proteins and riboswitches can target \textit{rut} site accessibility as a means of controlling Rho-dependent attenuation; however, there are presently few examples of these mechanisms. The archetype regulatory protein is CsrA (\textit{carbon storage regulator A}) of \textit{E. coli}, a key regulator of carbon and stationary-phase metabolism, biofilm formation, quorum sensing, and motility (128). One of the targets of CsrA binding is the leader transcript of the \textit{pgaABCD} operon, which encodes...
proteins responsible for the production, modification, and export of the biofilm adhesin poly-β(1-6)-N-acetyl-D-glucosamine. CsrA binding negatively regulates pgaABCD expression and associated biofilm formation (in part) by remodeling leader transcript secondary structure to unmask a rut site, which permits Rho-dependent termination before transcription of the structural genes (129) (Fig. 10). This regulation allows biofilm formation to be linked to the carbon metabolism-related signals that control the RNA-binding activity of CsrA (128).

The clearest example of riboswitch-mediated Rho-dependent attenuation appears to be the mechanism controlling expression of the ribB operon of *E. coli*, which encodes an enzyme required for riboflavin synthesis (125). The ribB leader transcript contains a highly conserved aptamer for FMN, which closely resembles the previously described rfn box formed by the ribD leader transcript of *B. subtilis* (Fig. 7). Binding of FMN to the ribB leader transcript results in a secondary structure that enhances Rho-dependent transcription termination within the leader region. Presumably, the FMN-induced secondary structure unmasks a rut site, but that important detail remains to be determined. The only other reported examples of riboswitch-mediated Rho-dependent attenuation are Mg²⁺-sensitive mechanisms that regulate expression of the mgtA, mgtCRB, and corA operons of *Salmonella*. Each operon encodes a separate Mg²⁺ uptake system. The riboswitch-mediated mechanism that controls mgtA expression was discussed above, and apparently, analogous mechanisms control mgtCRB and corA expression (120, 121, 130, 131).

**RNA-Mediated Mechanisms of Rho-Dependent Attenuation and Regulatory Targets Other than the rut Site**

Recent evidence suggests that small RNAs (sRNAs) capable of binding leader transcripts are important regulators of Rho-dependent attenuation (132). For example, such a mechanism is proposed to regulate expression of the *E. coli* rpoS gene, which encodes the stationary-phase sigma factor, σ² (132). The small RNAs DsrA, ArcZ, and

---

**FIG 10** Model for CsrA-induced Rho-dependent attenuation control of pgaABCD expression in *E. coli*. The upstream region of the pgaA leader transcript folds into an imperfect hairpin that includes two potential CsrA binding sites in the upstream segment of the stem. The downstream segment of the stem includes a large part of a rut site, and its inclusion in the secondary structure prevents Rho binding and allows transcription to read through into the structural genes of the pgaABCD operon. In the presence of active CsrA, a dimeric form of the protein binds both CsrA binding sites simultaneously and unfolds the leader hairpin to expose the entire rut site. This configuration allows Rho to bind to the rut site and cause transcription termination within the leader region.
RprA apparently bind to the 5′ untranslated region of nascent mRNA and inhibit Rho activity, thereby stimulating rpoS expression. The mechanism of action of the sRNAs has not been defined for this or any other mechanism of this type, and it could be something other than controlling rut site availability. Additionally, it has been reported that cis-acting RNAP-binding aptamers (RAPs) are present in the leader regions of many E. coli transcripts and that these RAPs promote Rho-dependent transcription termination. The model is that RAPs promote termination either by uncoupling transcription and translation to expose rut sites or by enhancing transcription pausing (133).

Direct evidence for regulatory targets other than the rut site comes from studies of Rho-dependent attenuation control of mgtA, mgtCBR, and corA expression in Salmonella. In each case, the leader transcript can adopt two mutually exclusive conformations that control rut site accessibility. In the case of mgtA, the secondary structure that exposes the rut site under conditions of high Mg2+ levels also contains elements that create a long-lived pause site that corresponds to the only Rho-dependent termination site in the leader region (Fig. 11A). Furthermore, this strong pause site is required for Mg2+-sensitive regulation of mgtA expression (122). In the case of mgtBCR, the leader transcript contains a Rho-antagonizing RNA element (RARE) that precedes the rut site. The RARE does not block Rho binding to the rut site but, when single stranded, traps Rho in an inactive complex. Under conditions of low Mg2+, inefficient translation of a short leader ORF promotes formation of a transcript secondary structure that exposes a single-stranded RARE, which can inactivate Rho and permit transcription through the mgtBCR structural genes (Fig. 11B) (134). In the case of corA, the leader transcript conformation that exposes the rut site also promotes pausing at a site >100 bases downstream at which Rho-dependent termination can occur. In addition, a stem-loop structure that forms downstream of the alternative secondary structures controls the location at which Rho-dependent termination occurs (135). The stem-loop blocks Rho’s access to RNAP when paused at position 192 but not at a subsequent pause site at position 240 (Fig. 11C). These results indicate that multiple steps in the process of Rho-dependent termination can be targets for gene regulation.

Finally, another type of Rho-dependent attenuation broadly referred to as Rho-dependent polarity is worth mentioning. In this mechanism, uncoupled transcription and translation within an early gene of an operon exposes latent ribosome-free rut sites that permit Rho-dependent termination before the transcription of downstream genes. Any regulatory scheme that prevents translation of an early cistron can be employed, although typically translation initiation is the target, using regulators such as riboswitches and sRNAs (136, 137). Rho-dependent polarity differs fundamentally from the other attenuation control mechanisms described above in that translation is the primary regulatory target and transcription termination occurs within the structural genes of an operon.

CONCLUSIONS

This review describes nearly a half century of seminal research on the regulation of bacterial gene expression based on conditional transcription termination. Perhaps the three things that best characterize these studies are as follows: there are a lot of these mechanisms, and they are everywhere; the mechanisms that render terminators conditional are incredibly diverse; and the discovery of each class of regulatory mechanisms was a huge surprise. For example, the discovery of the trp mechanism in E. coli certainly made it easier to discover the pyrBl mechanism in the same bacterium, because both mechanisms employ the position of a translating ribosome as the central regulatory feature. However, these studies did little to aid the discovery of attenuation mechanisms that rely on RNA-binding proteins. The same can be said about attenuation mechanisms based on riboswitches and reiterative transcription. Finally, could anyone have guessed that the central element of Rho-dependent attenuation control of the E. coli tna operon was binding of the amino acid Trp to a nascent peptide while it was still in the exit channel of the ribosome?

Although the mechanisms discussed in this review represent only a small sample of
many interesting examples of attenuation control, they were chosen because they highlight both major and subtle elements of this type of gene regulation. Certainly, they represent each class of known attenuation control mechanisms. They also demonstrate the economy, versatility, dynamic range, and physiological importance of attenuation control mechanisms. Many of the mechanisms rely only on the basic machinery of transcription or transcription and translation, and none rely on DNA binding regulatory proteins, suggesting that attenuation control has an ancient origin (138). The mechanisms also take advantage of subtle effects. For example, many rely on the fact that RNA secondary structures form quickly, so that when there are competing structures, formation of the more upstream one is strongly favored. Additionally, the mechanism controlling *tna* expression in *E. coli* and the mechanisms controlling expression of the Mg$^{2+}$ transporters in *Salmonella* appear to use uniquely hard-to-
translate Pro codons as speed bumps to appropriately position a translating ribosome. Finally, how much more ingenious can riboswitches and reiterative transcription be?

The discovery of new transcription attenuation mechanisms will be easier in the future. Computational analyses can find virtually every intrinsic transcription terminator, common riboswitch, short ORF, and long leader region in any sequenced bacterial genome. Powerful new sequencing techniques and perhaps a reliable algorithm will make it much easier to find Rho-dependent terminators (107, 139). Even sites of regulatory elements will be unlike anything seen before. The elucidation of these new mechanisms, as with so many of the old ones, will require thorough examination of individual operons and researchers who are challenged by unexpected and puzzling results.

ACKNOWLEDGMENTS

I thank Paul Babitzke, Ron Breaker, Keith Corbino, Mike Gray, Eduardo Groisman, Tina Henkin, Patrick Lane, Bob Landick, Tony Romeo, and Bob Switzer for their help with the manuscript.

Work on bacterial gene regulation in my laboratory was supported by National Institutes of Health grants GM29466 and GM94466. I declare no conflict of interest.

REFERENCES

Turnbough Microbiology and Molecular Biology Reviews


57. Hobl B, Mack M. 2007. The regulator protein PyrR of Bacillus subtilis specifically interacts in vivo with three untranslated regions within pyr mRNA.


Charles L. Turnbough, Jr., received a Ph.D. in biochemistry from the University of Illinois, Champaign-Urbana, and studied bacterial gene expression as a postdoctoral fellow at the Berkeley and Davis campuses of the University of California. In 1980, he joined the faculty of the Department of Microbiology at the University of Alabama at Birmingham, where he has spent his entire academic career. His primary research interest during this time has been the regulation of gene expression in bacteria, focusing on the genes involved in pyrimidine nucleotide biosynthesis and salvage. These studies have resulted in the discovery of a number of new regulatory mechanisms, most of which employ only the basic machinery of transcription and translation. Several of these are described in this review. He is now a Professor Emeritus and actively continues his research.