INTRODUCTION

As many of the features of gene expression and its regulation have been revealed, it has become apparent that every organism, from the simplest to the most complex, utilizes an appreciable fraction of its resources in regulatory processes that control its functional genes and interrelate their various activities. The development of regulatory mechanisms optimal for each organism, in its environmental niche, is thus a major theme in evolution. Macromolecular interactions are the principal events of gene regulation, with small and large molecules serving as signals for these regulatory activities.

Most prokaryotic operons are regulated transcriptionally at or near a short promoter region that binds relatively few regulatory proteins. These interactions may activate or inhibit transcription initiation. Regulatory mechanisms also target molecular events that occur subsequent to transcription initiation. These mechanisms include processes that influence transcript elongation and transcription termination. These are classified under the general heading of transcription attenuation, the subject of this article. In most eukaryotes, regulation of gene expression is more complex, with each gene's promoter/regulatory region replete with regulatory sites, or elements, that allow recognition of, and response to, many different regulatory proteins. The state of these proteins reflects their identity, the cell's environment, the stage of development, and the metabolic and other activities that are proceeding in the cells in which these regulatory proteins reside. Transcription attenuation also occurs in eukaryotes, but features of such mechanisms are only beginning to emerge.

It was once thought that repression of transcription initiation was sufficiently adaptable as a regulatory mechanism to be principally responsible for most gene regulation in most organisms. It is now apparent that each organism uses a variety of regulatory strategies in modulating gene expression. Often a single gene or operon is regulated by multiple independent mechanisms. In view of the extensive use of gene regulation, an important consideration for each organism during the course of evolution must have been how much genetic information to devote to regulatory processes. In some organisms, notably bacteria, the need for effective gene regulation must have conflicted with the desire to maintain a relatively small genome. A small genome would limit the number of regulatory genes and events that could be devoted to control and probably would favor regulatory mechanisms that require minimal genetic information. These evolutionary pressures notwithstanding, it is evident that even organisms with small genomes, such as viruses, make extensive use of gene regulation.

Once organisms were committed to differential gene regulation, it would have been advantageous to sense all external and internal events relevant to each gene's expression and to couple these events in one or more regulatory circuits. The use of both gene-specific regulatory mechanisms and global regulatory mechanisms that interrelated gene activities clearly would have been beneficial. It is perhaps not surprising that so many regulatory processes exist and that cells can sense and respond in so many ways to the numerous events that influence their behavior.

In this chapter, we focus on transcription attenuation in *Escherichia coli* and *Salmonella typhimurium* (official designation, *Salmonella enterica* serovar Typhimurium). We shall review the
features of the best-understood attenuation mechanisms that are used by these organisms. Transcription attenuation was reviewed in the previous edition of this volume (91) and, most recently, in the Cold Spring Harbor Laboratory monograph Transcriptional Regulation (89).

Early History

During the 1950s, there was considerable progress in our understanding of transcriptional regulation in bacteria, largely as a result of studies by Jacob and Monod on negative control (inhibition of transcription initiation) of lac operon expression in E. coli (65). Subsequent studies with the arabinose operon of E. coli established that positive control (activation of transcription initiation) also was used to regulate gene expression (41). During this period, several investigations with mutant tRNA synthetases suggested that these enzymes, or some consequence of their activity, affected expression of amino acid biosynthetic operons (39, 144). Extensive mutational studies on the mechanism of transcriptional regulation of the histidine (his) biosynthetic operon of S. typhimurium implicated histidinyl-tRNA synthetase, the charging and function of tRNAHis, and/or the utilization of histidine in translation of the early segment of the his operon transcript as the molecules and events that modulated his operon expression (9). These studies failed to detect a repressor gene or protein, suggesting that some other regulatory mechanism was employed (9). During this period, regulatory studies with the trp operon of E. coli mostly dealt with the role of the trp repressor in negative regulation of transcription initiation in this operon. Several observations made during the course of these studies indicated that a trpophan-specific regulatory mechanism other than repression must influence this operon's expression. Thus, kinetic studies of transcription within the initial segment of the trp operon (the leader region) indicated that the addition of tryptophan to a tryptophan-starved culture caused premature termination of transcription within this initially transcribed region of the operon (63). It was also observed that mutant strains that lacked a functional trp repressor responded to tryptophan starvation by increasing trp operon transcription (12). In addition, some tryptophanyl-tRNA synthetase mutants were found to have elevated levels of trp operon enzymes when grown in the presence of excess tryptophan (116).

Subsequent studies with both the his operon of S. typhimurium and the trp operon of E. coli demonstrated that internal deletions that removed portions of the leader region but left the promoter intact caused a large increase in synthesis of operon mRNA and proteins without affecting mRNA stability (17, 64). These findings suggested that a natural barrier to transcription elongation, such as a transcription termination site, was located in the leader regions of these operons. This barrier was presumably removed in the deletion mutants (17, 64, 78). Kasai (78) introduced the use of the term "attenuation" to describe modulation of his operon expression at such a barrier. The barrier is now known to be a transcription termination site, referred to as an attenuator (78). Transcription attenuation mechanisms can be separated into distinct classes, many of which will be considered here. The term "translation attenuation" has been used to describe processes with somewhat similar features in which gene expression is regulated by events that influence the ability of a ribosome to initiate translation. In this chapter we shall describe only transcription attenuation mechanisms.

Objectives of Transcription Attenuation as a Regulatory Mechanism

Transcription attenuation allows an organism to regulate gene expression by exploiting RNA sequences and structures, as opposed to information contained in DNA. It also allows cells to sense availability of the precursors needed for RNA and protein synthesis. An RNA signal can direct a transcribing RNA polymerase molecule to pause during transcript elongation, to terminate transcription prematurely, or to transcribe through a potential termination sequence. RNA sequences also can provide binding sites for regulatory factors that can influence the aforementioned events. Mutually exclusive RNA structures can allow determination of whether transcription will or will not be terminated at a particular site. By exploiting RNA sequences and structures in regulatory decisions, nature has engaged RNAs in ways that are not appropriate to regulation of transcription initiation. Thus, we see that translation is often used to mediate attenuation decisions. Similarly, a transcribing or paused RNA polymerase molecule itself, its associated transcription factors, or the Rho termination factor can serve as a regulatory target, thereby influencing attenuation. Transcription attenuation mechanisms therefore increase the repertoire of molecular events that an organism can exploit in optimizing gene expression.

A second attractive feature of some transcription attenuation mechanisms, as we shall see from examples described in this chapter, is that relatively little genetic information, often less than 100 nucleotides (nt), is needed for specific regulation. Also, in many instances transcription attenuation proceeds without the participation of a specific regulatory protein. It seems likely, therefore, that if a primordial RNA world preceded the DNA world we know today, transcription attenuation mechanisms probably would have been among the earliest regulatory mechanisms to have evolved.

ATTENUATION MECHANISMS IN ENTERIC BACTERIA

The hallmark of transcription attenuation is control over the continuation of transcript elongation at sites that are encountered by RNA polymerase prior to a particular gene. A wide variety of attenuation mechanisms have been discovered in enteric bacteria, and many examples fall into classes that share certain features (89). The four principal classes of attenuation mechanisms (Table 1) are (i) a class in which the location of a ribosome controls formation of alternative secondary structures in the nascent transcript (ribosome stalling, alternative RNA structure-dependent attenuation); (ii) a class in which coupling between ribosome and RNA polymerase movement can directly preclude formation of a terminator RNA hairpin (ribosome coupling-dependent attenuation); (iii) a class in which a transacting factor governs formation of a terminator structure by interacting with the nascent transcript (regulatory factor-dependent attenuation); and (iv) a class in which the action of the Rho termination protein in the RNA segment preceding a structural gene is regulated (Rho-dependent attenuation). We will consider these classes individually in the following sections.

Class I: Ribosome Stalling, Alternative RNA Structure Formation

As described above, transcription attenuation was discovered in regulatory studies with the his and trp amino acid biosynthetic
operons in E. coli and S. typhimurium. Subsequent investigations revealed that closely related mechanisms control termination in the leader regions of the _leu_ (48, 81, 169), _thr_ (44, 107), _pheA_ (62, 186), _ilvGMEDA_ (93, 117), and _ilvBN_ (43, 56) amino acid biosynthetic operons, as well as in the leader region of the _pheST_ operon, which encodes the small (_pheS_) and large (_pheT_) subunits of phenylalanyl-tRNA synthetase (153, 154; chapter 91).

The evidence underlying current knowledge of these attenuation mechanisms has been the subject of two recent reviews, which include a case-by-case account of the published information about each (91), the complete secondary structures of the alternative RNA structures that can form from each operon’s leader transcript (91), and an up-to-date and detailed account of the different intermediates and events involved in _trp_ operon attenuation (89). Therefore, in what follows we will offer a summary description of the key features of the class I attenuation mechanisms for amino acid biosynthetic operons and restrict our attention primarily to recent findings and to the key questions that remain unanswered. The reader should refer to recent reviews (55, 89, 91) for additional details.

**Common Features of Class I Attenuation Mechanisms.** The structural features of the leader regions of amino acid biosynthetic operons regulated by class I attenuation mechanisms are surprisingly similar. These similar features include sites of transcription pausing and Rho-independent termination, a coding region for a short leader peptide containing codons for the regulating amino acid(s), and transcript segments that specify alternative RNA secondary structures (Fig. 1). Transcription pausing occurs in the initial segment of the leader region and is caused, in part, by formation of an RNA secondary structure in the nascent RNA chain, termed the pause RNA hairpin. The temporary halt to transcription induced by the pause signal allows time for a ribosome to begin synthesis of the leader peptide before the Rho-independent termination site is transcribed. Resumption of transcription when the ribosome encounters the paused polymerase is largely responsible for the synchronization of transcription and translation that is essential to this class of attenuation mechanisms.

The terminator RNA hairpin (3:4) is transcribed from a segment of DNA that is downstream from the pause signal and is termed the attenuator. An alternative RNA secondary structure, termed the antiterminator, can form from the downstream segment of the 1:2 (pause) structure and the upstream segment of the 3:4 terminator structure. Thus, the antiterminator is designated structure 2:3. Formation of the 2:3 structure during transcription of the leader region precludes formation of the terminator, thereby permitting transcription readthrough. The peptide coding region overlaps the promoter-proximal segment that is part of the 1:2 RNA structure. Translation of the peptide coding region versus ribosome stalling at key sites within it selects between terminator and antiterminator formation, causing either transcription termination or readthrough into the structural genes of the respective operon.

**A Detailed Model of Attenuation.** To illustrate the role of and relationship among these different segments of the leader transcript, we will describe our current understanding of events that occur during transcription of the _trp_ operon leader region. Here, a 141-nt leader RNA transcript forms the alternative terminator (3:4) or antiterminator (2:3) secondary structures (Fig. 2). The upper portion of the 1:2 structure forms the hairpin portion of the signal that causes polymerase to pause prior to the 2:3 versus 3:4 decision. In some situations, the entire 1:2 structure forms after polymerase escapes the pause (see below). The leader RNA also encodes a 14-amino-acid peptide that contains two tryptophan residues, at positions 10 and 11 (Fig. 2). Once RNA polymerase initiates synthesis of an RNA chain and clears the promoter region, it moves rapidly to a pause site located just after RNA segment 2 (Fig. 2 and 3). This pause is very short-lived in vivo, probably less than 1 s, but still long relative to the average dwell time of RNA polymerase at a single DNA position. The pause signal is complex, consisting of multiple components, one of which is a nascent transcript RNA hairpin that corresponds to the upper portion of the 1:2 secondary structure (discussed below). Halting RNA polymerase at the pause site allows time for initiation of leader peptide synthesis before RNA polymerase moves into the attenuator region where the regulatory decision
FIGURE 1 Structure and relevant features of the leader transcripts and leader regions of amino acid biosynthetic operons regulated by attenuation. The antiterminator and pause structures are depicted with unpaired bulges to emphasize that they do not contain perfectly paired stems. Note that the pause RNA hairpin corresponds to the upper portion of the 1:2 structure and is the only portion of the structure thought to form when RNA polymerase resides at the pause site. The term “terminator” describes an RNA or DNA segment that is responsible for, or contributes to, a transcription termination event. Within the context of attenuation, “terminator” refers to the terminator RNA hairpin (here 3:4) plus the run of U's that can form when formation of a mutually exclusive antiterminator structure (here 2:3) is prevented. (This figure and Fig. 2, 3, 7, 8, and 9 are adapted from reference 89 with permission of the publisher.)

FIGURE 2 Alternative RNA secondary structures of the trp operon leader transcript. (A) Termination conformation; (B) antitermination conformation. The positions in the leader peptide corresponding to the initiation, control, and termination codons are shown in boldface.
INITIAL STAGES OF TRANSCRIPTION

LEADER TRANSCRIPT

DNA

RNA POLYMERASE

Ribosome Binds to Transcript

Ribosome Movement Releases the Paused Transcription Complex

Ribosome Fails to Bind Transcript

Super-Attenuation

TRP STARVATION
Ribosome Stalls on Trp Codons, Allowing Formation of 2:3

ADERATE TRP
Ribosome Moves to Stop Codon

3:4 Formation

2:3 Forms; Terminator (3:4) Blocked

2:3 Formation Prevents 3:4 Formation

1:2 Forms; Terminator (3:4) Forms

READTHROUGH

BASE LEVEL READTHROUGH

TERMINATION

FIGURE 3 A detailed model of transcription attenuation in the trp operon. The nontranscribed DNA strand is depicted as being extruded to the outside of RNA polymerase and reannealing at the upstream edge of the transcription complex to be consistent with chemical and nuclease protection data (see discussion in reference 94). (See text for further description.)

will be made. Once a ribosome reaches the paused polymerase in the course of translation, the paused polymerase is released and resumes RNA synthesis (88).

This synchronization of transcription with ribosome movement over the leader peptide control codons is a critical feature of attenuation. Since extreme depletion of the cognate charged aminoacyl-tRNA leads to essentially complete readthrough at the attenuator, synchronization of ribosome and RNA polymerase action must be efficient. However, if translation of the leader peptide coding region fails, for instance when the leader peptide initiation codon is mutated, RNA polymerase will eventually escape from the pause site spontaneously and terminate at the attenuator with 95% efficiency (178). When amino acids are not growth limiting and translation proceeds to the stop codon normally, termination is only 85% efficient (178).

Once RNA polymerase leaves the pause site, five outcomes are possible (Fig. 3). Which of these actually is achieved is determined by the relative rates of transcription of leader DNA by RNA polymerase and ribosome movement on the newly synthesized leader RNA chain. If a ribosome does not load onto the transcript before polymerase spontaneously escapes the pause site (the situation just described), the transcript subsequently
folds into the 1:2–3:4 termination conformation efficiently, and transcription of the operon stops at the attenuator (outcome 1, superattenuation; Fig. 3). If the ribosome releases the paused transcription complex (the normal situation), the translating ribosome typically will halt synthesis at one of two positions, depending on the availability of charged tryptophanyl-tRNA: the control Trp codons when the cell is starved for Trp (position 1) or the UGA stop codon when the supply of Trp is adequate (position 2) (Fig. 2 and 3). If the supply of charged tryptophanyl-tRNA is inadequate (Trp starvation; left side of Fig. 3), the ribosome halted at the Trp control codons should block about 13 nt on either side (129). In this situation, formation of the 1:2 secondary structure is not possible, and the RNA folds into the 2:3 antitermination conformation. This precludes subsequent formation of the 3:4 terminator structure and leads to transcription through the attenuator region into the operon’s structural genes (outcome 2, readthrough; Fig. 3). If the supply of charged tryptophanyl-tRNA is adequate, the translating ribosome will reach the leader peptide stop codon (adequate Trp; right side of Fig. 3); either it will remain there until after polymerase transcribes the attenuator or it will release before polymerase reaches the attenuator. If the ribosome remains at the stop codon while RNA segment 4 is synthesized, it will block formation of the 2:3 structure. This, in turn, will promote formation of the 3:4 terminator hairpin, leading to termination of transcription at the attenuator (outcome 3, termination; Fig. 3). If instead the translating ribosome releases at the stop codon before synthesis of RNA segment 4 is completed, two outcomes are possible, depending on whether the nascent transcript spontaneously folds into the 1:2 or 2:3 conformation. If 2:3 forms, readthrough into the operon’s structural genes will occur as described above for Trp starvation (outcome 4, basal-level readthrough; Fig. 3). If 1:2 forms, termination will ensue (outcome 5; Fig. 3). We estimate that formation of 1:2 or 2:3 occurs with approximately equal probability (see below).

In the following sections, we will describe our current understanding of the major features of this model, note several potential variations on the above-described mechanism that may occur in different amino acid biosynthetic operons, and identify key questions that require further study.

Alternative RNA Structures Modulate Termination. Although all eight operons that are regulated by this class of attenuation mechanism encode potentially alternative RNA structures in their leader RNAs that are consistent with the model described above, there are significant variations for each in the exact folding patterns of these structures (see reference 91). The his operon leader, for instance, folds into a termination conformation with three significant stem-loop structures, termed A:B, C:D, and E:F (the terminator structure), as shown by the workers who first studied the his attenuator (Fig. 4). The antitermination conformation of this leader contains two significant stem-loop structures, B:C and D:E (Fig. 4). The his leader RNA encodes seven tandem histidines overlapping segment A, in contrast to the two Trp codons found in the trp leader. This probably reflects the greater sensitivity of the his attenuator to amino acid starvation (see below). Despite these differences, however, the alternative pattern of folding still allows a ribosome stalled on RNA segment A to prevent formation of the terminator hairpin (E:F) because elimination of the A:B structure allows the B:C structure to form. B:C prevents formation of the C:D structure, which in turn allows formation of the D:E structure, which then precludes formation of E:F.

Another variation occurs in the ilvGMEDA operon, which encodes enzymes for synthesis of leucine, isoleucine, and valine. The ilvGMEDA leader RNA structure contains a large bifurcation in the 1:2 RNA structure (Fig. 5). This effectively divides segment 1 into two portions, termed 1A and 1B by Hatfield and coworkers (see reference 55). Here a significant question has arisen as to whether a ribosome stalled at the tandem isoleucine codons in the loop of the 1:2 structure could prevent base pairing between segments 1A and 2, since it would be more than 13 nt from the RNA segments that pair (Fig. 5). Thus, if 1A:2 re-formed when a ribosome halted in response to starvation for isoleucine, but not starvation for valine or leucine, 2:3 formation could be inhibited and termination might still result. This scenario led Hatfield to propose a “double-ribosome” model of attenuation in the ilvGMEDA leader (55). In this model, a second ribosome loads onto the transcript before the attenuation decision is made and could stop immediately behind one stalled at the isoleucine codons. This second ribosome would prevent pairing between segments 1A:2.

Despite these variations, the importance of competition between the formation of mutually exclusive RNA secondary structures in controlling attenuation is established beyond reasonable doubt. Four independent lines of evidence support this conclusion. First, for the E. coli trp, S. typhimurium his, and E. coli thr leader regions, many leader RNA base substitutions have been observed to alter termination (see reference 91 for a review). Outside the leader peptide coding regions, the effects of all of these substitutions can be explained by their effects on the relative stabilities of the different RNA structures. For example, replacement of G-93 with A in the trp leader (Fig. 2) or of U-86 or of C-118 with A in the his leader (Fig. 4) increases termination at the corresponding attenuator because these substitutions weaken base pairing in the antiterminator conformation of the leader RNA. Further, the effects of the C-118→A substitution in the his leader are reversed by a compensatory G-157→U substitution on the opposite side of the D:E RNA secondary structure because base pairing is restored by the second substitution.

Second, sequential deletions of RNA segments 1, 2, 3, and 4 have been examined for both the E. coli (92) and Serratia marcesens (160–162) trp operons. Again, their effects are consistent with the model of attenuation outlined above. Deletion of segment 1 greatly increases readthrough at the attenuator, presumably because the 2:3 structure forms unhindered. Deletions that extend into segment 2 restore termination to high levels, and further deletion into segment 3 or 4 eliminates termination.

Third, nucleotide sensitivity studies with the isolated trp leader RNAs from S. marcesens (86) and E. coli (128) reveal patterns of cleavage and protection by single- and double-strand-specific nucleases that are consistent with the structures depicted in Fig. 2. Moreover, a recent nuclear magnetic resonance study of a synthetic RNA corresponding to residues +114 to +134 of the trp leader RNA is consistent with the assigned structure (134).

Finally, the role of the 3:4 terminator hairpin has been studied in particular detail. Substitutions of bases in the paired stems of the terminator hairpins from the his, trp, leu, and thr leader regions have been shown to reduce termination (see references 89 and 91 and references therein). Experiments with both the trp (143) and thr (174) leader regions have shown that these substitutions affect
transcription only when present in the transcribed DNA strand, and thus the RNA transcript, and have no effect if present only in the nontranscribed strand of a heteroduplex template.

Role of the Leader Peptide Coding Region. In the attenuation mechanisms that control expression of amino acid biosynthetic operons, regulatory information is transduced through synthesis of the leader peptide. In essence, the rate of ribosome movement over the control codons senses the availability of cognate charged tRNAs for protein synthesis and communicates this information to the transcription complex via its influence on nascent RNA structure. Thus, the leader peptide coding region plays a central role in these mechanisms. Many aspects of this role have been covered in detail in earlier reviews (89,91). We emphasize four central points here.

First, translational capacity is sensed through the availability of charged tRNAs for efficient incorporation into the leader peptide, not by the actual pools of amino acids. This is shown most conclusively by the following experiments (92).

FIGURE 4 Alternative RNA secondary structures of the his operon leader transcript. (A) Termination conformation; (B) antitermination conformation.

FIGURE 5 Alternative RNA secondary structures of the ilvGMEDA operon leader transcript. (A) Termination conformation; (B) antitermination conformation.
sively from the findings that defects in the tRNA that reduce its translational efficiency (e.g., failure to isopentenylate tRNATrp in a midA strain (179) or to pseudouridylate tRNAhis in a hisT strain (72, 99)) or defects in the aminocayl-tRNA synthetases responsible for charging tRNA^{trp} or tRNA^{his} (99, 179) lead to elevated readthrough at the corresponding attenuators.

Second, it is the act of translation and not the leader peptide itself that regulates attenuation. Synthesis of the complete thr leader peptide was observed in coupled transcription-translation experiments (108), and the trp leader peptide has been detected both in vitro (32) and in vivo (35). However, all tests for trans action have been negative. Further, the leader peptides themselves are unstable, having half-lives of less than 5 min in cell extracts (32, 108). At least for the trp operon, sequences present downstream from the attenuator site fold back and pair with the leader peptide ribosome-binding site, effectively precluding more than a single round of leader peptide synthesis (32). Finally, where a single amino acid is encoded by both rare and frequently used codons, such as in the leu operon, the presence of rare codons in the leader peptide coding region confers a strong regulatory response, whereas the presence of common codons does not (14; reviewed in reference 89).

Third, the number and positions of control codons within the leader peptide coding regions vary among the different operons regulated by attenuation (Fig. 6). Some leader peptide coding regions (trp, his, phe, and leu) confer a regulatory response primarily to a single cognate amino acid, whereas others (thr, ilvGMEDA, and ilvBN) confer a multivalent response. In at least one case (the ilvGMEDA operon in S. marcescens), a single rare codon is sufficient to confer regulation (61), whereas the phe and his operons each contain seven codons for a single regulatory amino acid. Presumably this reflects both the relative ease of translating a particular codon and the degree of sensitivity to amino acid starvation desired for a particular operon. Although all of the leader peptide coding regions begin between nt 22 and 45 of the different leader transcripts, some extend only to the loop region of the 1:2 hairpin, whereas others continue over segment 2 to near or past the pause site. Interestingly, the aromatic amino acid biosynthetic operons all fall into the first class and the branched-chain amino acid biosynthetic operons fall into the second. Presumably this reflects an early divergence in the evolutionary paths of amino acid biosynthetic operons. It also may have regulatory significance. In the double-ribosome model of ilvGMEDA operon attenuation, pausing conceivably could be induced by base pairing between segments 1A and 2 even after a ribosome has passed segment 1A (55). Continuation of the leader peptide coding region up to or past the pause site would allow the ribosome to release such a paused RNA polymerase.

Fourth, the range of regulation conferred by the different leader peptide coding regions varies dramatically among the different operons. The presence or absence of Trp in the growth medium does not normally affect readthrough of the trp attenuator; the 10-fold change in readthrough occurs only upon extreme Trp starvation in mutant bacteria or transiently upon transfer of bacteria from a Trp-containing to a Trp-free medium (177). Thus, the trp attenuator may represent adaptation of the mechanism to an operon also regulated by repression (see discussion in reference 89). In other cases, significant changes in

FIGURE 6 Leader peptide coding regions and pause signals of eight operons regulated by a ribosome-stalling, alternative RNA structure mechanism of transcription attenuation. The RNA sequences of the leader regions for E. coli trp, S. typhimurium his, E. coli phe, E. coli thr, S. typhimurium leu, E. coli ilvGMEDA, and E. coli ilvBN were obtained from the GenBank DNA sequence database. The locations of RNA secondary-structure elements (indicated by numbered horizontal arrows) are shown as illustrated in reference 91. The locations of pause sites (P adjacent to vertical line or arrow) are shown for the trp (171), his (25), thr (45), leu (14), ilvGMEDA (57), and ilvBN (57) operons.
Attenuator readthrough are observed when the cognate amino acid is present or withdrawn from the medium. For instance, such manipulation changes pheA operon expression 17-fold (46).

Transcription-Translation Coupling. In principle, the coupling of polymerase movement over the leader region to ribosome positioning in the leader peptide coding region could be accomplished in two ways, stochastically or by pausing and release of RNA polymerase. Synchronization could be achieved stochastically if the time between exposure of the leader peptide initiation site and extrusion from RNA polymerase of the portion of RNA segment 2 whose pairing with segment 1 or 3 dictates the attenuation decision were commensurate with the time required for ribosome binding and initiation of leader peptide synthesis. The distance between the leader peptide start codon and this critical portion of segment 2 varies from −50 to −75 nt (Fig. 6). At its average rate over long transcriptional units (35 to 50 nts) (21, 22, 167), RNA polymerase would require 1 to 2 s to complete synthesis of this portion of leader RNA. This appears to be sufficient time for stochastic synchronization with ribosome movement (but see below). However, the available data favor the pausing-release model for synchronization. Every leader region examined to date contains a pause site at a location appropriate to stop RNA polymerase before the critical portion of RNA segment 2 is exposed (Fig. 6). Further, experiments performed with a coupled in vitro transcription-translation system showed that the paused complex decays much more slowly when leader peptide synthesis is prevented by addition of inhibitors (88). Nonetheless, direct demonstration of a requirement for pausing to obtain synchronization has not yet been obtained because of the complications that alternative RNA folding imposes on experimental design.

Basal-Level Expression. One interesting feature of attenuation that has received some attention in recent years is the role of basal-level readthrough of the attenuator even in the presence of an adequate supply of cognate amino acid. The defining observation of basal-level readthrough is the fivefold reduction in trp operon expression caused by the trpL29 mutation, which changes the leader peptide initiation codon from AUG to AUA (83, 187). In S. marcescens, deleting the trp leader peptide start codon causes a 10-fold reduction in operon expression (161). Similar reductions in operon expression also have been observed for leader peptide initiation codon mutations in the his (73) and phe (46) operons. Although occasional slow movement of ribosomes through the leader peptide coding region even under optimal conditions could explain this effect (91), at least for the trp operon the major cause of basal-level readthrough appears to be the release of ribosomes at the leader peptide stop codon (Fig. 2) (138, 139). From the effects of several stop codon and ribosome release factor mutants, Roesser and Yanofsky (139) suggest that in tryptophan excess, 75% of the ribosomes remain attached to the leader RNA at the stop codon while RNA polymerase moves over the attenuator, thereby directing terminator formation. The remaining 25% of the ribosomes release, with the RNA folding into terminator and antiterminator conformations with equal probability. This produces 15% basal-level readthrough when combined with the 5% readthrough inherent in the trp attenuator. This mechanism allows significant adjustment of basal-level readthrough in different organisms by changes in the relative stabilities of the leader RNA termination and antitermination conformations (176).

Polymerase Recognition of Pause and Termination Signals. Until recently, explanations for transcription pausing and Rho-independent termination have focused on the role of the nascent transcript RNA hairpin. Findings obtained since the first edition of this book was published suggest that these mechanisms, and indeed the entire process of RNA chain elongation, may be considerably more complex than had been appreciated (24, 27, 31; chapter 55). Because the complexity of this subject is beyond the scope of this review and because it is covered in another chapter, we will offer only a summary of the most important points regarding RNA chain elongation, pausing, and termination and then consider their implications for the mechanisms of attenuation. The reader is referred to chapter 55 and the other recent reviews cited above for more detail.

RNA chain elongation. Early studies of in vitro transcription revealed that RNA chain elongation is discontinuous and punctuated by long pauses at certain sites and rapid elongation through other regions. It now appears that this variability in the elongation reaction may reflect significant heterogeneity in the structure of transcription complexes located at different template positions. Studies of isolated transcription complexes halted at single template positions reveal significant differences in the size of the DNA footprint on DNA (varying between −25 and 40 nt) (84, 90, 175), the position of the transcript relative to the downstream edge of the complex (24, 90, 94, 175), the size of the single-stranded region of DNA within the complex (24, 94), and the stability of individual complexes (85).

The pattern with which these properties change in two different promoter-proximal regions led Chamberlin (24) to propose an "inchworm" model for discontinuous translation of RNA polymerase (see also references 19, 27, and 31). In this model, RNA polymerase is envisioned to contain two separate binding sites for DNA. These sites alternately lock and slide to allow addition of up to 10 nt to the growing RNA chain while the upstream contact moves forward and the downstream contact remains fixed. When sufficient strain accumulates in the complex, the downstream contact unlocks and undergoes a discontinuous forward step of up to 10 bp. The RNA transcript also is envisioned to be positioned on RNA polymerase in two distinct sites. Approximately 8 nt at the 3' end of the RNA chain may be paired to or positioned near the DNA template in a "loose" RNA-binding site, such that the 3' end is near the leading edge of the transcription bubble. The next 10 nt appear to remain tightly bound to RNA polymerase, perhaps in a distinct transcript exit channel. Filling of the loose site with RNA allows chain extension during the time the downstream contact is locked.

This model immediately raises several fundamental questions (27): (i) Is a constant inchworm cycle maintained as polymerase moves along the DNA? (ii) Are discontinuous changes in RNA polymerase contacts to DNA or RNA somehow linked to recognition of pause or termination signals? (iii) Do different RNA polymerase molecules transcribe with precisely the same discontinuous movements (i.e., remain in phase with each other)? (iv) What is the effect of an upstream polymerase on movement of a downstream polymerase? The answer to the first question probably is no. Not only should this produce periodic effects on transcript elongation patterns that have not been observed, but also it now appears that RNA polymerase behavior is
more monotonic over at least certain sequences located further from the promoter (125). We address the second question below, but answers to the last two await further study (see also reference 27).

**Pause signals.** Extensive studies of pausing in the *his* and *trp* leader regions suggest that pause site recognition is complex and relies on multiple interactions between RNA polymerase and different segments of RNA or DNA. Rather than depending simply on formation of the 1:2 (or A:B) secondary structure to stop chain elongation, the *his* and *trp* pause sites are multipartite and consist of at least four components (26, 27, 95): (i) a 5- or 6-bp stem-loop RNA secondary structure (the pause hairpin) 10 or 11 nt upstream of the transcript 3' end, (ii) a 3'-proximal region of transcript or DNA template, (iii) the apparently unfavorable geometry of U or C at the 3' end of the transcript reacting with an incoming GTP, and (iv) up to 14 bp of DNA downstream of the pause site. Recognition of these pause sites appears to be coupled to a discontinuous rearrangement of the transcription complex at the pause site: the pause hairpin can be detected by RNase VI digestion only at or 1 nt before RNA polymerase reaches the pause site and occurs coincident with a 10-bp advance in the downstream edge of the RNA polymerase-DNA contact. However, the downstream edge of the complex remains fixed as RNA polymerase approaches the pause site, apparently causing the complex to compress as it passes over the 10 bp upstream from the pause site (168). This compressed complex, which may form either by movements of parts of the polymerase or by loopings of the RNA and DNA chains, isomerizes into a paused configuration by a 10-bp jump in the downstream DNA contact and simultaneous formation of the pause hairpin. The downstream sequence, and perhaps sequences just upstream from the pause site, appear to set up isomerization into the paused configuration by controlling the inchworm movement of RNA polymerase. An altered downstream sequence causes more than 75% of the complexes to pass the site in vitro without entering the paused state (168), probably because the requisite discontinuous movements in the complex have been compromised. The pause hairpin, on the other hand, appears to stabilize the paused configuration by an ionic interaction between the phosphate backbone and a positively charged region on RNA polymerase (26, 27). Whether this interaction slows elongation through an allosteric effect on RNA polymerase or by altering the positioning of the transcript in the active site remains to be determined.

**Termination signals.** Like pause sites, Rho-independent termination sites are multipartite (136). In addition to the terminator RNA hairpin and following uridine-rich segment, the traditional hallmarks of a Rho-independent terminator, both the DNA sequence downstream from the release site and the RNA or DNA sequence upstream from the terminator hairpin also contribute to the efficiency of termination for at least some Rho-independent terminators with imperfect uridine-rich regions (e.g., UUU-CUGCCG [136, 163]). These different interactions may explain the terminator-specific effects of salt and nucleoside triphosphate (NTP) concentration on termination efficiency (137). As we suggest above for the pause site, precise and discontinuous movements in the transcription complex also appear to contribute to the mechanism of termination (126, 168).

**Outstanding Questions about Attenuation in Amino Acid Biosynthetic Operons.** Although work on the mechanisms of attenuation in the amino acid biosynthetic operons of enteric bacteria has slowed in recent years, it is not for lack of fundamentally important questions. Rather, these questions have become increasingly difficult to address experimentally. We outline here several among them that concern fundamental macromolecular events that could fruitfully occupy talented investigators in the years ahead.

**Structures and dynamics of nascent transcript RNA folding.** First, what are the true stabilities in vivo of the alternative RNA secondary structures that participate in attenuation control, and even more important, how fast do they form once the relevant segments of nascent transcript emerge from the transcription complex? Do as yet unidentified proteins interact generally with RNA and influence stability? Attenuation control relies on the relative stability of RNA secondary structures, yet we know surprisingly little about the precise conformations of these structures, their true stabilities, their interactions, or the kinetics of their formation. Although available enzymatic digestion data support the presumed structures of the *trp* leader transcript (Fig. 2; 86, 90, 128), many leader transcript structures have been assigned simply on the basis of predictions of a computer algorithm. These algorithms do not accommodate the known complexity of RNA structures, which includes many different types of non-Watson-Crick base pairs and tertiary interactions (173). Thus, it is altogether possible that structures presumed for some leader transcripts, particularly the more complicated ones like those for ilvGMDA (Fig. 5), may be missing important details. Moreover, we are even less knowledgeable about the kinetics of nascent transcript folding or of hairpin melting. For instance, how fast after a ribosome releases from the leader peptide stop codon does folding into the 1:2 or 2:3 structures occur? With the increasing power and sophistication available to study structure with techniques like nuclear magnetic resonance spectroscopy, these questions should be approachable.

**Role of transcription pausing.** A second fundamental question concerns the role of transcription pausing in attenuation. In the case of pyrimidine biosynthetic operons, the essential role of transcription pausing to the mechanism appears indisputable (see below). However, it is currently possible to argue that transcription and translation in the attenuator regions of amino acid biosynthetic operons are synchronized either stochastically or by the formation of a paused transcription complex and its subsequent release by the ribosome translating the leader peptide coding region. Although several arguments favor the latter possibility (see above), it has not been confirmed by experimental observation. If pausing is required, elimination of the pause signal or of the ability of RNA polymerase to recognize it should prevent increased readthrough of the attenuator in response to undercharged tRNAs or defective aminoacyl-tRNAs. This is not simple to test experimentally, however, because most base changes that would weaken the pause signal would also alter the stability of the alternative RNA secondary structures, and because the pause signal is multiparte (see above). Elimination of any single component of the pause signal has at most a five- to sixfold effect on the pause half-life (26). Nonetheless, if we have greater understanding of the pause signal now makes it possible to design appropriate experiments. One could eliminate pausing almost completely by combining selected base substitutions in several components of the pause signal (particularly the DNA sequence downstream from the pause site, which appears to be critical for efficient recognition of the signal). These substitu-
tions could be inserted in attenuation control regions in other­wise isogenic strains with or without a mutation, such as miaA or hisT, that causes elevated readthrough at the attenuator. If transcription pausing is required to ensure ribosome participation within the critical window of time, then elimination of pausing should also eliminate the increased readthrough caused by miaA or hisT.

Relative timing of RNA polymerase movements. The essence of attenuation control is the kinetic relationships among RNA polymerase movement on DNA, ribosome movement on nascent RNA, and RNA folding. It therefore follows that we will not understand the mechanism of attenuation completely until we understand the kinetics of leader transcript formation and leader peptide synthesis. Although the effects on attenuation of different ribosome and RNA polymerase locations now are relatively clear (Fig. 3), the actual timing of events during the movement of RNA polymerase and the ribosome over the leader DNA and RNA are at present simply best guesses. Assumptions about the time required for RNA polymerase to reach the pause signal or for a ribosome to reach the control codons, for instance, are based on average elongation rates over long transcriptional or translational units (e.g., see kinetic arguments in references 55 and 91). As the foregoing discussion of RNA chain elongation, pausing, and termination should make abundantly clear, the movements of RNA polymerase are complex (see above). The average rate of transcript elongation, in fact, reflects the rate-limiting influence of pause sites within a transcriptional unit (30, 79, 97). The rate of transcription between these pause sites is uncertain but could be up to 10 times faster than the average. Hence the apparent reasonableness of stochastic coupling between transcription and translation in a leader region (discussed above) is the kinetic argument for the necessity of a double-ribosome model for attenuation (discussed above) may be illusory. There is no reason to assume that ribosome movements will be any less complex.

A true understanding of attenuation will require that we explain the behavior of each of these complex molecular machines in detail. Precisely how RNA polymerase moves through the leader region matters a great deal to models of attenuation. One intriguing possibility, for instance, arises from the possible requirement for a particular coupling between discontinuous translation by RNA polymerase and recognition of the terminator. If RNA polymerase is capable of transcribing the terminator in different phases of the discontinuous cycle, it is conceivable that the transcription pause also plays a key role in attenuation by establishing the configuration of the transcribing complex so that it enters the attenuator region properly phased for terminator recognition.

Evolution of attenuation in amino acid biosynthetic operons. One intriguing question is why attenuation control in E. coli and presumably S. typhimurium is used to regulate expression of particular amino acid biosynthetic operons (trp, his, phe, leu, thr, ilvGMEDA, and ilvBN) and an aminoacyl-tRNA synthetase operon (pheST) and not others that at least superficially appear equally suitable (e.g., cysDNC, argCBH, or metBFL). Interestingly, the attenuation-controlled, amino acid biosynthetic operons encode enzymes that catalyze assembly of related amino acids (Trp, Phe, and His [aromatic] and Leu, Thr, Ile, and Val [branched chain]). Perhaps the bacterium requires a particularly rapid response to starvation for these particular amino acids when its growth medium changes quickly from one rich in these particular amino acids (such as the human gut) to one in which they are absent. It is also possible that the energetic cost of synthesizing the aromatic and branched-chain amino acids makes it desirable to maintain particularly efficient control over expression of their biosynthetic enzymes. Finally, one should not disregard the simple possibility of historical accident leading to retention or loss of attenuation control during divergent evolution of related operons.

Class II: Ribosome-RNA Polymerase Coupling

The first several examples of attenuation control discovered in E. coli and S. typhimurium involved mechanisms that regulated the expression of amino acid biosynthetic operons. Each example employed ribosome stalling at control codons as a regulatory signal and an alternative transcript secondary structure as a means of preventing terminator hairpin formation. These similarities raised the possibility that attenuation control was limited to amino acid biosynthetic operons and to a single mechanism for regulating transcription termination. The first clear indication to the contrary was the discovery of attenuation control of pyrB1 operon expression in E. coli (118, 142, 165). The pyrB1 operon encodes the pyrimidine biosynthetic enzyme aspartate transcarbamylase, which is not involved directly in amino acid metabolism. Elucidation of the pyrB1 attenuation control mechanism revealed that transcription termination at the attenuator, a Rho-independent transcription terminator that precedes the pyrB1 structural genes, was regulated in a way fundamentally different from that described for the amino acid biosynthetic operons (reference 89 and references therein); namely, transcription termination was controlled by the extent of coupling (i.e., movement and location) of a translating ribosome and a transcribing RNA polymerase within the pyrB1 leader region. In this case, the ribosome directly controls the formation of the terminator RNA hairpin by steric hindrance.

Shortly after the discovery of attenuation control of pyrB1 expression, a similar mechanism was described for the pyrE gene of E. coli (18, 130, 132), which also encodes a pyrimidine biosynthetic enzyme. Attenuation control mechanisms equivalent to their E. coli counterparts also appear to regulate pyrB1 (114) and pyrE (120) expression in S. typhimurium. At present, the pyrB1 operon and pyrE gene provide the only examples of this second class of attenuation control, which we designate the ribosome-RNA polymerase coupling mechanism; however, attenuation control of E. coli ampC expression could be considered a special case (see below). In this section, we will describe the general features of this class of attenuation control, provide a detailed description of the well-studied mechanism of the pyrB1 operon of E. coli, and comment on unique features of pyrE regulation. The reader also is referred to an earlier review (89) and to the chapter in this book on pyrimidine metabolism (chapter 35) for additional information.

Common Features and General Model. There are three regulatory elements required for the known examples of attenuation control by ribosome-RNA polymerase coupling. These elements are present in the leader region located between the promoter and the first regulated structural gene (Fig. 7). These elements include (i) a polypeptide-encoding open reading frame, which is referred to as the leader open reading frame; (ii) a Rho-independent transcription terminator (attenuator) near the downstream end of the leader open reading frame; and (iii) UTP-sensitive transcription pause sites (runs of uridines in the leader transcript) that precede the attenuator in the leader open reading frame. These elements
permit UTP-mediated regulation of gene expression according to the following general model (Fig. 7). Transcription is initiated at a promoter preceding the leader open reading frame and proceeds into this region. When the intracellular level of UTP is low, RNA polymerase stalls at UTP-sensitive pause sites, which provides time for a ribosome to initiate translation of the leader transcript and translate up to the stalled polymerase. When polymerase eventually escapes the pause region and transcribes the attenuator, formation of the terminator RNA hairpin is blocked by the presence of the adjacent translating ribosome. In this case, polymerase continues transcription into the downstream gene(s). In contrast, when the level of UTP is high, polymerase transcribes the leader region without stalling at UTP-sensitive pause sites. In this situation, there is insufficient time for a ribosome to establish tight coupling with polymerase before the formation of the terminator hairpin. The result is transcription termination before the regulated structural gene(s).

**Attenuation Control of pyrBl Expression in E. coli.** Most of what we know about class II attenuation control comes from the study of the pyrBl operon of E. coli. This operon encodes the catalytic (pyrB) and regulatory (pyrI) subunits of the allosteric enzyme aspartate transcarbamylase, which catalyzes the first committed step in the de novo synthesis of pyrimidine nucleotides. Expression of the pyrBl operon is negatively regulated over an approximately 300-fold range by pyrimidine availability, specifically by the intracellular concentration of UTP (103, 104, 148, 164, 165). The first indication that attenuation was involved in this regulation came from the sequence determination of the pyrBl promoter-leader region, which revealed a potential Rho-independent transcription terminator (attenuator) located 23 bp before the pyrB structural gene (118, 142, 165). Transcription from the pyrBl promoter (102) was shown to be efficiently (98 to 99%) terminated at this attenuator in vitro (165) and in vivo under conditions of pyrimidine excess (96). The sequence of the 158-bp pyrBl leader region indicated that the leader transcript could not adopt alternative stem-loop structures to regulate terminator hairpin formation, which implied that attenuation control of pyrBl expression must be mechanistically different from that described for the amino acid biosynthetic operons (165).

Elucidation of the pyrBl attenuation control mechanism relied on the discovery of two additional regulatory elements. The first...
was a translatable 44-codon open reading frame that extends through the leader region and ends 6 nt before the pyrB gene (Fig. 8) (141). The second was strong UTP-sensitive transcription pause sites in the pyrBl leader region upstream of the attenuator (36, 165) (Fig. 8). These features suggested a regulatory model (as drawn in Fig. 7) in which low UTP levels cause RNA polymerase to transcribe slowly through the first half of the leader region, which results in tight coupling between polymerase and a ribosome that is translating the 44-codon open reading frame. When polymerase eventually transcribes the attenuator, the terminator hairpin is precluded from forming or is disrupted by the adjacent translating ribosome. The result is transcription into the pyrBl structural genes and synthesis of aspartate transcarbamylase when it is needed by the cell.

To describe the regulatory elements more clearly and to illustrate the interactions required for class II attenuation control, we will discuss in detail the key features of the E. coli pyrBl control mechanism.

Attenuation versus attenuation-independent control of pyrBl expression. An early question about the pyrBl attenuation control mechanism was whether it could account for all pyrimidine (i.e., UTP)-mediated regulation of operon expression. This regulation occurs over a much wider range than the typical 10- to 40-fold range for attenuation control of amino acid biosynthetic operon expression in E. coli and S. typhimurium (91). To answer this question, pyrBl expression was measured in a mutant E. coli strain which carries a 9-bp chromosomal deletion that removes the run of eight T·A base pairs at the end of the pyrBl attenuator plus one additional base pair to maintain the reading frame of the leader polypeptide (104). All Rho-independent transcription termination is abolished at this mutant attenuator. When the mutant strain was grown under conditions of pyrimidine excess, pyrBl expression was approximately 50-fold higher than that in an isogenic pyrBl" strain. When growth of the mutant was limited for pyrimidines, operon expression increased an additional sevenfold. Growth of the pyrBl" strain under the same pyrimidine-limiting conditions resulted in a 300- to 350-fold increase in operon expression (103, 104). These results indicate that attenuation control is responsible for most (i.e., 50-fold), but not all, of the pyrimidine-mediated regulation. Recent studies indicate that the residual sevenfold pyrimidine-mediated regulation in the mutant strain is due to an additional pyrBl control mechanism. In this mechanism, high UTP levels induce reiterated transcription within a run of three T·A base pairs in the initially transcribed region, which inhibits promoter clearance (103).

Translation of the pyrBl leader transcript. In the model for pyrBl attenuation control, translation of the 44-codon open reading frame of the leader transcript plays the critical role of disrupting the formation of the terminator hairpin. It was necessary, therefore, to demonstrate that such translation occurs physiologically. This task was accomplished by fusing the pyrBl promoter region and leader open reading frame to the lacZ gene and detecting the predicted

![Pause Hairpin](image)

![ Terminator Hairpin](image)

FIGURE 8 Nucleotide sequence and secondary structures of the pyrBl leader transcript. Transcription initiation occurs at the first two A residues in vitro and almost exclusively at the second A in vivo. Nucleotides 21 through 152 encode the 44-amino-acid leader polypeptide, and the sequence ends with the AUG initiation codon of the pyrB cistron. The Shine-Dalgarno sequences for the leader polypeptide and pyrB are underlined. Two RNA secondary structures are shown; the first (pause hairpin) is flanked by uridine-rich sequences, within which strong UTP-sensitive transcription pausing occurs, and the second (terminator hairpin) is specified by the pyrBl attenuator. The sites of transcription termination at the pyrBl attenuator correspond to positions 133 to 135.
leader polypeptide-β-galactosidase fusion protein in cells (141). However, it is not the 44-amino-acid leader polypeptide that is important for regulation but the act of translation of the leader transcript. This fact was demonstrated by showing that near-normal attenuation control was observed in a mutant bearing a +1 frameshift at codon 6 of the 44-codon open reading frame, which still allows translation of the entire leader region (28). This situation is comparable to that described above for the leader peptides of amino acid biosynthetic operons.

To show that regulation requires translation of the 44-codon leader open reading frame, the effects of mutations that either strongly inhibit the initiation of translation of the leader transcript or introduce stop codons early in the open reading frame, well before the attenuator, were measured (28, 140, 141). Each mutation reduced operon expression to ≤6% of the wild-type level under conditions of pyrimidine limitation and to ≤30% of the wild-type level under conditions of pyrimidine excess; the latter effect presumably reflects a low level of tight coupling of transcription and translation within the wild-type leader region even in cells grown with an ample pyrimidine source. The net effect of the mutations was to significantly reduce the range of pyrimidine-mediated regulation. Further support for the regulatory role of translation of the leader transcript comes from a study of pyrBl (and pyrE) expression in cells in which ribosomes translate at either one-third or two-thirds of their normal rate because of a mutation in the rpsL gene encoding the ribosomal protein S12 (69). When mutant cells were grown with excess pyrimidines, they exhibited significantly reduced pyrBl expression compared to the wild-type level. Apparently, the slower rate of translation reduced coupling of transcription and translation in the leader region, resulting in increased transcription termination at the attenuator. In contrast, the slow ribosomes did not significantly affect pyrBl expression under conditions of mild or severe pyrimidine limitation, indicating that in these situations, transcription pausing within the leader region is sufficient to permit extensive coupling of transcription and translation even with slow ribosomes.

The proposed attenuation control model employs a translating ribosome to physically block the formation of the terminator RNA hairpin under conditions of pyrimidine limitation. To test this aspect of the model, the distance that a ribosome must translate within the leader region to suppress transcription termination at the attenuator was determined by introducing stop codons at various sites in the leader open reading frame from codons 6 to 33 (140). On the basis of the estimated size of the ribosome-binding site (77, 156), translation would have to proceed to within approximately 15 nt of the terminator hairpin to permit the ribosome to interact directly with this sequence. The results show that translation termination at or before codon 24, which is 16 nt upstream of the terminator hairpin (Fig. 8), reduces operon expression to approximately 5% of the wild-type level under pyrimidine-limiting conditions. In contrast, when translation termination occurs at codon 25, which should be the first stop codon at which ribosome-binding overlaps the sequence of the terminator hairpin, expression is 64% of the wild-type level. In general, the level of operon expression increases as the stop codon is moved further downstream of codon 25, with the highest level of expression (i.e., 91% of the wild-type level) occurring with translation termination at codon 33, which is within the loop of the terminator hairpin. In S. typhimurium, the leader open reading frame normally stops at codon 34 as a result of a sequence difference between the two bacteria (114). These results provide strong support for the proposed regulatory role of the ribosome.

According to the model for regulation, a single round of translation of the leader transcript is sufficient to elicit readthrough transcription. Consistent with this limited requirement for translation, the ribosome-binding site preceding the leader open reading frame is relatively weak, only 7% as efficient as the pyrB ribosome-binding site (K. L. Roland, C. Liu, and C. L. Turnbough, Jr., unpublished data). Interestingly, mutations in the leader ribosome binding site that increase leader translation by as much as 10-fold cause less than a 2-fold increase in operon expression under conditions of pyrimidine excess, and they have no significant effect on operon expression under conditions of pyrimidine limitation (C. Liu and C. L. Turnbough, Jr., unpublished data). Thus, the pyrBl attenuation control mechanism appears fine-tuned to produce strong regulation without unnecessary leader translation. Additional control of translation in the leader region is suggested by the presence of sequences between nt 75 and 143 of the leader transcript (Fig. 8) that are complementary to the leader ribosome-binding site (118, 141). Formation of a secondary structure by these sequences could block multiple rounds of translation of readthrough transcripts and perhaps all translation of attenuated transcripts. It is also possible that this secondary structure plays a more active role in attenuation control. For example, in cells grown under conditions of pyrimidine excess, rapid transcription up to the attenuator may be quickly followed by the formation of a secondary structure which precludes ribosome binding to the leader transcript and thereby ensures transcription termination at the attenuator. Such a role remains to be established experimentally.

UTP-sensitive transcription pausing in the pyrBl leader region. The discovery of UTP-sensitive transcription pausing in the pyrBl leader region was a key finding in the development of the attenuation control model. This pausing provided the regulatory sensor, equivalent to the control codons in the case of the amino acid biosynthetic operons, that could respond to different levels of UTP in a way that controlled transcription termination at the attenuator. UTP-sensitive pause sites are defined as positions in RNA (or DNA) at which RNA polymerase stalls while awaiting the addition of a uridine residue to the 3' end of a growing RNA chain. Generally, these pause sites are detectable only at low UTP concentrations (e.g., below 200 μM). In E. coli and S. typhimurium, the UTP concentration varies from approximately 50 μM in cells grown under conditions of pyrimidine limitation to 1 mM or slightly above in cells grown under conditions of pyrimidine excess (6, 119, 164).

The first in vitro experiments to detect UTP-sensitive transcription pausing in the pyrBl leader region revealed only a small cluster of pause sites that correspond to a uridine-rich region located 20 nt before the terminator hairpin in the leader transcript (Fig. 8) (165). Subsequent in vitro studies employing more accurate and sensitive procedures have provided a significantly different view of pausing in the leader region preceding the attenuator (36). Instead of one cluster of pause sites, there are a large number of sites throughout the leader transcript at which RNA polymerase pauses when the UTP concentration is low (i.e., 20 μM with other NTPs at 400 μM). Nearly all of these sites correspond to positions where uridines are added to the
leader transcript. Pausing at these sites decreases with increasing UTP concentrations and is no longer detectable at a concentration of 400 μM. Although some degree of pausing apparently can occur before the addition of every uridine residue in the leader transcript at 20 μΜ UTP, the strength of individual pause sites is variable. This variability presumably reflects the effects of DNA sequence and RNA secondary structure (25, 95). In this regard, strong pausing occurs before the addition of uridines at leader transcript positions 81 to 85, which are located 7 nt downstream of an RNA hairpin (designated the pause hairpin in Fig. 8). These sites correspond to the originally identified cluster of UTP-sensitive transcription pause sites. Recent studies of hairpin-induced transcription pausing suggest that the pyrBI pause hairpin would affect specifically the rate of transcript elongation through positions 81 to 85 in the leader region (26, 27). A mutation that destabilizes the pause hairpin has in fact been shown to reduce pausing in this region (K. Mixter-Mayne and C. L. Turnbough, Jr., unpublished data).

Although some pause sites within the leader region may be stronger than others, the large number of these sites indicates that it is the cumulative effect of pausing at multiple positions that is the key factor in controlling coupling between RNA polymerase and the ribosome translating the pyrBI leader transcript. Consistent with this view, replacing all seven uridines with adenines between positions 81 and 88 in the leader transcript (Fig. 8) causes only a twofold reduction in the range of pyrimidine-mediated regulation of pyrBI expression (Mixter-Mayne and Turnbough, unpublished data). Interestingly, this large substitution mutation causes an approximately 10-fold increase in the level of readthrough transcription at the pyrBI attenuator in vitro (from 2 to 20%). This result may indicate that strong UTP-sensitive pausing at positions 81 to 88 in the wild-type leader transcript is used to establish a transcription complex configuration that is optimally phased for subsequent attenuator recognition as discussed above.

Unlike transcription at low UTP concentrations, appreciable transcription pausing in the pyrBI leader region is not detected at 20 μΜ ATP or CTP and is detected at only one site at 20 μM GTP (36). Strong pausing occurs before the addition of a guanine residue at position 55 in the leader transcript (Fig. 8). This pausing is substantially reduced, although not completely eliminated, at 400 μΜ GTP. The reason for pausing at this site and not before the addition of other guanine residues in the leader transcript remains to be established, but presumably it is associated with the function of a secondary structure in the transcript. Although nucleotide G-55 is included in the pause hairpin of the leader transcript (Fig. 8), its position in the 5’ segment of the stem indicates that the observed pausing is unrelated to the formation of this hairpin. The physiological consequences of pausing at this unique GTP-sensitive site are not known. However, pyrBI expression is increased fourfold in cells defective in guanine nucleotide synthesis (68). Possibly, this effect is due to GTP-sensitive transcription pausing that permits coupling of transcription and translation in the pyrBI leader region and readthrough transcription past the pyrBI attenuator.

The more pronounced pausing at 20 μΜ UTP than at 20 μΜ ATP, CTP, or GTP appears to be due, at least in part, to a difference in the apparent Km values for these nucleotides during transcript elongation. The apparent Km for UTP during elongation appears to be significantly higher than the apparent Km values for the other NTPs (70, 82). A higher apparent Km for UTP and the fact that UTP (unlike CTP) pools fall to low levels in cells limited for pyrimidines (131) make UTP an ideal regulatory effector for attenuation control based on ribosome-RNA polymerase coupling.

Additional data supporting the proposed role of UTP-sensitive transcription pausing in attenuation control come from studies of a strain of S. typhimurium carrying an altered RNA polymerase that exhibits an approximately sixfold-higher apparent Km for the binding of UTP (6 mM) and ATP (4 mM) during transcript elongation. This mutant displays constitutive expression of the pyrBI operon and the pyrE gene at high intracellular levels of UTP (70), indicating that transcription pausing during the addition of uridine (or other) residues to the pyrBI leader transcript, and not the UTP level per se, is the key determinant in regulation. Furthermore, the transcription elongation factor NusA enhances UTP-sensitive pausing within the pyrBI and pyrE leader regions in vitro (6, 36) and appears to be important in determining the level of expression of these genes in vivo (6). Presumably, NusA plays a key role in establishing a rate of transcript elongation that permits tight coupling of transcription and translation in cells limited for pyrimidines. It has been shown that pyrE expression is hyper-repressed by the addition of uracil to a culture of E. coli containing a defective NusA (6). Apparently, transcript elongation is too fast in the mutant strain to permit coupling of transcription and translation in the pyrE attenuator region, particularly in cells containing high levels of UTP. These results indicate that the activity of NusA or of any factor that influences the rate of transcript elongation can affect the expression of genes regulated by class II attenuation control.

Unique Features of pyrE Attenuation Control. The pyrE gene of E. coli encodes the pyrimidine biosynthetic enzyme orotate phosphoribosyltransferase. Expression of this gene is regulated over a 30-fold range almost entirely by an attenuation control mechanism which is analogous to that described for the pyrBI operon (18, 60, 130, 131). A striking difference, however, is that the pyrE gene is the second gene of a bicistronic operon, and the cell uses transcription and translation of the first cistron for the purpose of attenuation control of pyrE expression. UTP-sensitive as well as GTP-sensitive transcription pausing has been demonstrated in the rph transcript (6). Also noteworthy is that the rph region of the transcript ends eight bases before the terminator RNA hairpin specified by the pyrE attenuator. However, on the basis of the size of a typical ribosome-binding site, translation to the end of the rph cistron would result in efficient disruption of the terminator hairpin, thereby allowing readthrough transcription.

Outstanding Questions. Perhaps the most intriguing question at present is whether this general class of attenuation control, which was the first to "broaden our attenuation horizons," is actually very limited itself. For example, is it limited to pyrimidine genes because only changes in UTP pools can affect transcription pausing appropriately? Alternatively, this class of attenuation control may be more prevalent and based, in some
cases, on changes in the activity of other factors that control coupling of transcription and translation.

Class III: Regulatory Factor Dependent

In the two classes of attenuation control described so far, regulation of transcription attenuation is achieved by coupling transcription and translation within the regulatory leader region. In the third class, called regulatory factor-dependent attenuation, the regulatory role of translation is replaced by a trans-acting RNA-binding factor. This factor can be a protein or an antisense RNA that binds a specific sequence in the leader (or intercistronic) region of the regulated transcript. Factor binding can either directly block the formation of an attenuator-specified terminator hairpin or indirectly control its formation by favoring a particular RNA secondary structure.

*bgl* Operon. In *E. coli*, the best studied example of regulatory factor-dependent attenuation is regulation of *bgl* operon expression. In this case, the regulatory factor is a specialized RNA-binding protein. The *bgl* operon contains three genes, *bglG*, *bglF*, and *bglB*, which enable cells to use certain aromatic β-glucosides such as salicin and arbutin as carbon sources (133, 147). The *bglG* gene encodes a positive regulatory protein, *BglG*, which is required for operon expression. The *bglF* gene encodes the β-glucoside-specific enzyme II of the phosphoenolpyruvate-dependent phosphotransferase system (designated enzyme II*βgl* or BglF), which is present in the cytoplasmic membrane and transports β-glucosidic sugars into the cell with concomitant phosphorylation of the sugar. The *bglB* gene encodes phospho-β-glucosidase B, which hydrolyzes phosphorylated β-glucosides to yield glucose-6-phosphate. The *bgl* operon is cryptic in wild-type cells, but a variety of spontaneous mutations activate the operon by enhancement of transcription initiation from a preexisting but silent promoter (105). Once the operon is activated, full expression requires a β-glucoside inducer and also cyclic AMP (cAMP) and the cAMP receptor protein (CRP) (135).

β-Glucoside-mediated regulation of *bgl* operon expression occurs through an attenuation control mechanism (Fig. 9). In the absence of a β-glucoside inducer, most transcripts initiated at the activated *bgl* promoter are terminated at a Rho-independent attenuator located just upstream of *bglG*, the first gene in the operon (109). A second Rho-independent attenuator is located between *bglG* and the adjacent *bglF* gene (147). The positive regulator *BglG* is required to prevent transcription termination at both attenuators (109, 145). Low levels of *BglG* and *BglF* are synthesized in the absence of β-glucosides, but under these conditions, *BglF* inactivates *BglG* by phosphorylation (2, 3, 146). Phosphorylation of *BglG* prevents the formation of dimers, which are required for RNA binding and antitermination activity (4). In the presence of β-glucosides, *BglF* dephosphorylates *BglG*, allowing it to dimerize and function as an antitermination factor (2, 3, 146). Nonphosphorylated dimers of *BglG* prevent transcription termination within the operon by binding to a sequence in the *bgl* mRNA that precedes and partially overlaps the terminator hairpin encoded by each attenuator. This binding blocks the formation of the terminator hairpins, allowing expression of the operon. Apparently, the *BglG*-binding site forms an alternative secondary structure that is recognized by *BglG* dimers (5, 60).

At present, it is not known if *S. typhimurium* carries a *bgl* operon. Attempts to clone parts of the operon by PCR have been unsuccessful. *S. typhimurium* does not grow on aromatic β-glucosides, indicating that if the genome includes a *bgl* operon, it is silent (A. Wright, personal communication).

Antisense RNA-Mediated Control of Plasmid Replication. At present, there are no clear cases of antisense RNA-dependent attenuation in *E. coli*, but this type of regulation can be illustrated by examining the mechanism that controls replication of the staphylococcal plasmid pT181 (124). An analogous mechanism controls replication of the streptococcal plasmid pIP501 (20). Replication of plasmid pT181 is controlled by the level of the initiator protein RepC. Constitutive transcription of plasmid pT181 produces an antisense RNA capable of base pairing with a target sequence near the 5′ end of the repC transcript. When this pairing occurs early during *repC* transcription, it promotes the formation of an RNA hairpin 5′ to the *repC* start codon. This hairpin causes premature Rho-independent transcription termination, which precludes RepC synthesis. In the absence of the antisense RNA (i.e., when plasmid copy number is low), an upstream sequence in the *repC* leader transcript pairs with the 5′ segment of the terminator hairpin, thereby blocking the signal for transcription termination and allowing synthesis of the full-length *repC* transcript. In this case, RepC is synthesized and can direct another round of plasmid replication.

Class IV: Rho Factor Dependent

*E. coli* and *S. typhimurium*, and perhaps most bacteria, have two distinct classes of transcription termination sites. One class, discussed previously, is responsible for termination in many operons regulated by transcription attenuation. Sites of this class appear to be recognized by RNA polymerase acting alone; they are referred to as Rho-independent or intrinsic termination sites. Sites of the second class require the action of the protein factor Rho. Rho interacts with both transcript and RNA polymerase and directs polymerase to terminate transcription (chapter 55). The first example of the use of transcription termination and its relief in gene regulation involved Rho factor and its role in regulating early gene transcription in bacteriophage lambda. The phage-specified N protein was shown to be responsible for antitermination at Rho-dependent termination sites (53).
Much is known about Rho's structure and mechanism of action (23, 47, 53; chapter 55). The characteristics of many Rho-binding sites and Rho termination sites have been determined. Rho has been shown to be an ATP-dependent RNA-DNA helicase (23); helicase activity is required for termination. Rho causes termination by crystallizing a paused RNA polymerase molecule and altering its behavior. Rho action requires accessory factors, such as NusG (chapter 55). Models that explain Rho's binding to RNA, its relative movement on RNA, and the mechanism by which it promotes termination have been proposed (23, 47). The end of an operon is often defined by a Rho-dependent termination site. Occasionally Rho termination events are regulated by accessory factors, such as NusG (chapter 55). Models that explain Rho's binding to RNA, its relative movement on RNA, and the mechanism by which it promotes termination have been proposed (23, 47).

Rho-Dependent Transcription Attenuation in the tna Operon. The tna operons of E. coli and Proteus mirabilis are regulated by tryptophan-induced transcription attenuation (52, 75, 157-159). These operons contain two major structural genes: tnaA, encoding the hydrolytic enzyme tryptophanase, and tnaB, specifying a tryptophan permease (34, 75). The presence of these two proteins permits these organisms to use tryptophan as a carbon or nitrogen source. In both organisms, tnaA is preceded by a leader regulatory region containing a short coding region, tnaC, encoding 24- and 36-residue leader peptides in E. coli and P. mirabilis, respectively (Fig. 10) (75, 158). tnaC and tnaA are separated by a ca. 100-200-bp region that contains regulated sites of Rho-dependent transcription termination (75, 158). A third tna operon, that of Enterobacter aerogenes SM-18, is organized similarly (80). S. typhimurium does not have a tna operon.

Transcription initiation in the tna operons of E. coli and P. mirabilis, as in many degradative operons, is subject to catabolite repression. A cAMP-CRP-binding site is located just upstream of the -35 site in each tna promoter. Activated CRP presumably binds at these sites and promotes initiation. Transcription initiation is not subject to tryptophan control. Rather, the presence of exogenous tryptophan induces an attenuation mechanism that prevents Rho-dependent termination in the leader region of the operon (157, 158). Mutational studies have shown that Rho alterations that reduce its termination proficiency increase basal tna operon expression (157, 158). Similarly, deletions that remove a presumed "rut" (Rho utilization) site located immediately following tnaC, or that remove segments of the spacer segment between tnaC and tnaA, also increase basal operon expression (52; A. Kamath and C. Yanofsky, unpublished data).

Studies on the role of tnaC in tna operon induction in E. coli and Proteus vulgaris have shown that translation of tnaC is essential for tryptophan-induced attenuation. Thus, mutations that alter the tnaC start codon eliminate tryptophan induction (52, 159). tnaC of these two organisms contains a single Trp codon, which, when replaced by other codons, prevents induction by tryptophan or by the amino acid encoded by the substituted codon (52). Replacing the Trp codon by a stop codon also prevents induction. However, introducing an appropriate tRNA<sup>Trp</sup> nonsense suppressor gene, but not other tRNA suppressor genes, restores induction (52). Frameshift mutations that allow translation of tnaC to proceed beyond its stop codon result in elevated, constitutive expression of the operon (52, 158). In addition, when the tna leader transcript is highly expressed from a multicopy plasmid, induction of the resident tna operon is reduced appreciably (50), suggesting that some cell component becomes limiting for attenuation.

Recent mutational studies have provided evidence suggesting that the TnaC leader peptide itself plays a role in tryptophan-induced attenuation (K. Gish and C. Yanofsky, unpublished data). Nonconservative amino acid changes at six positions in TnaC prevent induction, whereas conservative changes at these same positions, or silent changes in codons specifying the residues at these positions, have no effect on induction (Gish and Yanofsky, unpublished data). However, with the exception of six consecutive amino acid residues of TnaC that include the single Trp residue, the TrpCs from the three bacterial species have very
different amino acid sequences. This observation suggests that some or all of these six residues may be crucial in promoting antitermination.

Among the mutational changes in tnaC of E. coli that have been examined, one important class results in high-level constitutive expression of the operon. Mutations of this class convert an out-of-frame stop codon near the middle of tnaC (Fig. 10) to any one of several sense codons. In these mutants, tna operon expression is five times higher than the tryptophan-induced level observed with the wild type (Gish and Yanofsky, unpublished data). An out-of-frame stop codon is located at a similar position in tnaC of the tna operons of the other two bacterial species.

On the basis of studies with the tna operons of E. coli and P. vulgaris (Gish and Yanofsky, unpublished data; Kamath and Yanofsky, unpublished data), a tentative model that explains many features of this example of transcription attenuation can be proposed (Fig. 10). According to this model, segments of the tna leader region have two crucial functions. One is to allow Rho to bind to the transcript and cause RNA polymerase to terminate transcription. This is the principal event that must occur in cells grown in the absence of tryptophan. This function is believed to be mediated by the out-of-frame stop codon within tnaC. Presumably, ribosome stoppage or dissociation at this stop codon allows Rho to recognize and bind to the following untranslated segment of mRNA. The second function of the leader region is to permit an unidentified tryptophan induction event to provide an activated leader peptide (i.e., TnaC) that prevents Rho molecules from attaching to the leader transcript (perhaps by inhibiting ribosome release), hence allowing expression. A boxA-like sequence has been identified at the end of tnaC in the E. coli and P. vulgaris operons (158; Kamath and Yanofsky, unpublished data), raising the possibility that some RNA-binding accessory factor binds here and influences Rho’s action. The precise binding site for Rho in the tna leader transcript is not known, nor is it known how tryptophan promotes antitermination. This example, like others in which Rho is a participant, is complex.

rRNA Operons. A special case of Rho-dependent attenuation occurs in the seven homologous rRNA operons of E. coli and S. typhimurium, which encode a 16S-23S-5S rRNA precursor from which the mature tRNAs are excised nucleolytically (see chapter 90). The elongating rRNA transcript, unlike other nascent RNAs, is not translated and thus is a potential target for Rho protein throughout its synthesis. To prevent Rho from stopping rRNA synthesis, the cell has evolved a special mechanism, termed rRNA antitermination. The degree to which this process is regulated is not known. We will describe RNA antitermination here because, although it is related to N- and Q-dependent antitermination during bacteriophage lambda growth (see chapter 55), it is the only clear-cut example of a cellular regulatory mechanism that alters the capacity of RNA polymerase to elongate an RNA chain through an entire transcriptional unit.

Transcription initiation in rRNA (rrn) operons occurs at two promoters, P1 and P2, that couple the rate of initiation to growth rate, stringent response, and translational capacity of the cell (see chapter 90). Between the start site and the first nucleotide of the mature 16S rRNA, the rRNA transcripts contain two sequences, boxA and boxB, that are important for antitermination and that also are present in the nut regions of phage lambda, where they participate in N-dependent antitermination (chapter 55). boxB encodes an RNA stem-loop structure. boxA encodes the consensus sequence 5′-UGGCUUUAACA, which appears to be a target for binding of a heterodimer of NusB and ribosomal protein S10 (NusE) (111, 122). Interestingly, the order of these sequences in the rrn leaders is boxB-boxA, opposite that found in the lambda nut sites. Mutations in boxA and NusB abrogate RNA antitermination (100, 150, 155). Although deletion of boxB does not prevent antitermination, some boxB mutations do, perhaps because boxB can bind both an inhibitor and an activator of antitermination (155). The boxB-boxA motif is highly conserved among both eubacteria and archaebacteria (15), suggesting that rRNA antitermination may be a universal mechanism except in eukaryotes, in which a special RNA polymerase has evolved for rRNA synthesis.

rRNA antitermination has been recapitulated in vitro. In addition to RNA polymerase, an rrn template, NusB, NusA, NusE (S10), and NusG, RNA antitermination in vitro requires an unidentified cellular protein that may be the rrn homolog of the phage lambda N protein (155).

Several other features of the rRNA antitermination mechanism raise interesting questions that are worthy of further study. First, rRNA antitermination blocks only Rho-dependent, not Rho-independent, termination (1). In fact, tandem Rho-independent terminators halt transcription at the end of the rrn operons. How this is accomplished and why it differs from the antitermination mechanisms involved in lambda gene expression are not known.

Second, the involvement of a ribosomal protein, S10 (NusE), in rRNA antitermination raises the as yet unexamined possibility of a feedback control circuit in which excess 16S rRNA inhibits antitermination by sequestering S10.

Third, by analogy to lambda N-dependent antitermination, the boxB-boxA region of the rrn nascent RNA is thought to be anchored to antitermination-modified RNA polymerase as it transcribes the rrn genes (112, 121). This finding raises the possibility, suggested first by Morgan (115), that the 5′ stem of the “processing stalk,” which is required for excision of 16S rRNA, is delivered to the 3′ stem by the transcribing polymerase. Such a mechanism also could account for the existence of a second boxA downstream from the 16S gene, where it could facilitate a second antitermination modification of RNA polymerase, after cotranscriptional excision of 16S rRNA, and possibly set up 23S processing by a similar delivery mechanism.

Finally, the finding that fivefold-faster transcription of rrn genes by T7 RNA polymerase produces inactive ribosomes (98) raises the possibility that a precise elongation rate is necessary for proper folding of rRNA (155). Thus, by determining the time available for 5′-proximal segments of the rRNA to fold before potentially competing 3′-proximal segments are synthesized, rRNA antitermination, by itself doubling the normal elongation rate, may control the sequence of events by which rRNA reaches its proper three dimensional conformation.

Other Interesting Examples

Ribosomal Protein L4 Mediates Transcription Attenuation in the Leader Region of the S10 Operon of E. coli. The S10 operon of E. coli contains the structural genes for 11 ribosomal proteins. Four of these proteins are constituents of the 30S ribosomal subunit, while the remaining seven are constituents of the 50S ribosomal
subunit (71; chapter 90). The S10 operon, like other operons encoding ribosomal proteins, is autogenously regulated by one of its specified proteins. This regulatory protein, L4, is encoded by the third gene of the operon (71; chapter 90). L4 is one of several primary rRNA-binding proteins that recognize specific segments of 23S rRNA during the initial stages of assembly of the 50S ribosomal subunit (152).

The L4 protein is believed to regulate S10 operon expression by two distinct mechanisms, translational feedback inhibition and transcription attenuation (for a recent review, see reference 185). Feedback inhibition occurs whenever synthesis of L4 provides a molar excess of this protein over 23S rRNA. The free L4 molecules are believed to bind to the S10 transcript and inhibit translation of the S10 coding region (101, 180). Inhibition of translation of the S10 coding region also reduces translation of the downstream coding regions in the S10 transcript. Thus, S10 synthesis and L4 synthesis are translationally coupled (70, 101).

Translational feedback inhibition is a common regulatory feature of ribosomal protein operons of E. coli, i.e., ribosomal proteins specified by genes of other ribosomal protein operons also autogenously regulate their own synthesis. These proteins presumably also act by binding to their homologous transcripts (70). Binding of L4 to the S10 transcript during the initial stages of the transcript's synthesis also appears to reduce transcription of the structural genes of the operon, via transcription attenuation (42, 181). It has been shown that some mutations in the S10 leader region influence translational control, others alter transcriptional control, while some mutations affect both processes (42). These findings suggest that the regulatory sites essential for the two processes have both unique and common features.

The L4 protein regulates S10 operon transcription by promoting transcription termination (attenuation) in the S10 leader region (42, 181). The extent of attenuation control in vivo by L4 is about fourfold. Here, as in translational control, when L4 is present in excess over 23S rRNA, it functions as a regulator. The binding site for the L4 protein in 23S rRNA has been localized to a 110-base segment designated domain I (184). The presumed L4-binding site(s) in the leader segment of the S10 transcript that is responsible for attenuation control has not been determined (149). Transcript sequences and structures that are somewhat similar to sequences and structures in 23S rRNA have been identified in the vicinity of the ribosome-binding site used for S10 synthesis (180). The L4-S10 example is currently the only one in which transcription attenuation regulation expresses a ribosomal protein operon. As mentioned earlier, rRNA synthesis in E. coli also is subject to transcription attenuation (155).

L4-mediated transcription termination has been studied in vivo and in vitro. Under both conditions, L4 promotes transcription termination at a typical Rho-independent termination site located 140 bp from the transcription start site of the S10 operon (182). Translation of leader RNA is not involved in operon regulation. Rather, it appears that L4 recognizes some feature of the transcript or transcription complex during transcription of the leader region of the operon (183). The NusA protein is required for transcription termination at the S10 attenuator in vitro (183). NusA is believed to act by increasing the stability of a paused transcription complex that forms at the S10 leader termination site. The NusA-facilitated paused complex is thought to be stabilized further by association with the L4 protein (149, 183); this complex then presumably promotes transcription termination at the attenuator. Although attempts to demonstrate direct binding of L4 to the S10 leader transcript have been unsuccessful (149, 184), mutational studies have identified two segments of the leader transcript that form hairpin structures as being essential for L4 action (149, 183). The L4 protein promotes transcription termination in vitro when added to a NusA-RNA polymerase complex that has reached a pause site (183). This observation suggests that L4 recognizes the NusA-mediated paused complex. However, although ability to form the attenuator hairpin is sufficient for NusA action, an RNA segment upstream of this hairpin is necessary for L4 stabilization of the paused complex (183). This finding suggests that additional sites or factors are necessary for L4 action. When these sites and factors are discovered, and their mechanisms of action are elucidated, it may be possible to describe the precise events involved in L4-mediated transcription attenuation.

Transcription Attenuation in the ampC Operon of E. coli. The gene ampC of E. coli encodes a β-lactamase that is produced in small amounts and is secreted into the periplasm, where it can hydrolyze penicillin and related β-lactam antibiotics (123). Expression of ampC is subject to an interesting form of growth rate control (67). As the growth rate increases, the rate of production of the ampC-encoded β-lactamase also increases (54). Transcription of the ampC operon is initiated 41 bp preceding ampC, in the frd gene, which resides in a separate operon (16). The leader segment of the ampC transcript contains a typical Rho-independent transcription terminator (66). This transcript segment also has a ribosome-binding site just upstream of the terminator hairpin. The AUG start codon of this ribosome-binding site is immediately followed by an ochre stop codon (66). A ribosome bound at this ribosome-binding site would be expected to disrupt the terminator, thereby allowing transcription to proceed into the ampC coding region (66). On the basis of what is known about ampC operon regulation, the following model has been proposed to explain growth rate control. As the number of ribosomes per cell increases or decreases, coincident with changes in growth rate, it becomes more or less likely that a ribosome will bind at the leader ribosome-binding site and disrupt the terminator. Upon disruption of the terminator, synthesis of the ampC transcript and its specified periplasmic β-lactamase would be increased. Thus, the ribosome content per cell would determine the rate of transcription of ampC. Much evidence that is consistent with this model has been gathered. Transcription studies performed in vitro have demonstrated formation of the expected terminated transcript. Mutational studies have shown that ribosome binding and formation of the terminator hairpin are both essential for growth rate regulation (54, 66).

TRANSCRIPTION ATTENUATION MECHANISMS IN OTHER ORGANISMS

Although this review deals with attenuation mechanisms in E. coli and S. typhimurium, we think it essential that the reader be aware of several recent discoveries of novel attenuation mechanisms in other organisms, including both gram-negative (11, 37, 40) and gram-positive (13, 29, 33, 49) bacteria. Similar mechanisms could well be found in future studies of gene regulation in E. coli and S. typhimurium. Interestingly, these examples are principally class III mechanisms (i.e., regulatory factor-de-
pended attenuation), and they involve some operons whose homologs in *E. coli* and *S. typhimurium* are regulated by class I and class II mechanisms.

**B. subtilis trp Operon**

Attenuation control of the *trp* operon in *Bacillus subtilis*, as in *E. coli*, involves formation in the leader transcript of alternative secondary structures that either include or exclude a terminator hairpin (87, 151). However, in *B. subtilis*, formation of the transcript secondary structures is controlled not by the position of a translating ribosome but by the binding of a regulatory protein, TRAP (trp RNA-binding attenuation protein), which is encoded by the *mtrB* gene (10, 51). Under conditions of abundant tryptophan, TRAP binds to a segment of the leader transcript upstream of the terminator hairpin. This binding blocks formation of an antiterminator secondary structure equivalent to the *E. coli* 23 hairpin, thereby promoting formation of the terminator hairpin and causing transcription termination. Recent studies indicate that TRAP consists of 11 identical 8-kDa subunits (7, 8) and that each tryptophan-activated subunit binds one closely spaced G/UAG repeat in the leader transcript (10).

**B. subtilis pyr Operon**

A similar mechanism also appears to regulate the *pyr* operon in *B. subtilis*. This operon encodes the six pyrimidine nucleotide biosynthetic enzymes and includes two additional, 5'-proximal genes: *pyrR*, which encodes a 20-kDa regulatory RNA-binding protein that also can catalyze the conversion of uracil and phosphoribosylpyrophosphate to UMP, and *pyrP*, which encodes an integral membrane uracil permease (166). Distinct Rhö-independent terminators (attenuators) occur in the leader region, the *pyrR-pyrP* intercistronic region, and the *pyrP-pyrB* intercistronic region. Recent studies suggest that when pyrimidines are abundant, the PyrR protein binds to target sequences in the nascent transcript and disrupts formation of an antiterminator secondary structures that, when the pyrimidine pool is depleted, prevent formation of the terminator hairpins (166). An intriguing feature of this mechanism is the use of an enzyme involved in pyrimidine metabolism as the RNA-binding protein, perhaps allowing the binding site for UMP to be used both in the enzymatic reaction and as an allosteric site to regulate the RNA-binding activity of PyrR (R. Switzer, personal communication).

**tRNA-Directed Transcription Antitermination**

An exciting recent discovery is the finding that transcription attenuation can be directly controlled by binding of a cognate uncharged tRNA to the nascent leader transcripts of at least 12 aminoacyl-tRNA synthetase and 6 amino acid biosynthetic genes and operons in a variety of gram-positive bacteria (reference 59 and references therein). These leader RNAs each contain three conserved stem–loop structures (I, II, and III), followed by a highly conserved 14-base element (T box) and then the attenuator-specified terminator hairpin. An unpaired triplet sequence (called the specifier sequence), which corresponds to a codon for the cognate amino acid, occurs in a bulge region of each stem I. For all genes, an antiterminator hairpin can be formed by base pairing of a segment of the T box with a conserved sequence in the upstream portion of the terminator hairpin. In the antiterminator hairpin, the central seven bases of the T box form a bulge. When uncharged cognate tRNA is abundant, it presum-ably binds to the specifier sequence in stem I through codon-anticodon interactions and, near its 3' end, to the T-box bulge in the antiterminator hairpin. This stabilizes the antiterminator hairpin and permits readthrough transcription into the downstream gene(s). When the cognate tRNA is charged, it either does not bind stably or is unavailable for interaction because it is complexed with aminoacyl-tRNA synthetase or elongation factors. This allows the terminator hairpin to form and cause transcription termination. Thus, limitation of a particular amino acid or aminoacyl-tRNA synthetase increases the level of uncharged cognate tRNA(s), which in turn increases transcription of the genes encoding the enzymes that synthesize this amino acid or that charge it to tRNA.

**Possibility of Transcription Attenuation in Eukaryotes**

Although our review focuses on bacterial attenuation mechanisms, it is worth mentioning the possibility that transcription attenuation occurs in eukaryotes. Although similarities of features in eukaryotic transcriptional units to those found for bacterial attenuators have led to several suggestions of transcription attenuation in eukaryotes (38, 58, 110), to date there is no documented example of a eukaryotic gene whose expression is regulated at a discrete transcription termination site. However, in at least some cases (the human and murine c-myc proto-oncogenes, for instance), the susceptibility of RNA polymerase II to multiple, weak termination sites within or upstream from a gene appears to be controlled by events at or near the promoter (recently reviewed in references 76 and 172). These mechanisms are most formally analogous to bacterial antitermination mechanisms, such as those found in rRNA operons (see above and chapter 90) or during growth of phage lambda (chapter 55). For example, the lambda N and Q proteins play roles somewhat similar to that of human immunodeficiency virus type 1 Tat protein, whose binding to a promoter-proximal RNA structure (the transactivation response region) is required for efficient synthesis of full-length mRNA (recently reviewed in reference 74).

Although other types of attenuation mechanisms have not been found in eukaryotes, we caution against concluding they do not exist. An important lesson from studies of bacterial attenuation is that regulatory mechanisms evolve special features that uniquely couple transcription of a particular gene to the metabolic need for its gene product. Control of RNA chain elongation in eukaryotes is poorly understood, and these linkages may still be unrecognized. Both class III mechanisms and, as we have noted elsewhere (89), coupling of transcription by RNA polymerase II to assembly of the mRNA splicing machinery (in analogy to coupling to ribosome movement in bacteria) could plausibly occur in eukaryotes.

**EVOlUTIONARY ASPECTS, CONCLUSIONS, AND EXPECTATIONS**

We have described a variety of transcription attenuation mechanisms that are used by *E. coli* and/or *S. typhimurium* in regulating expression of specific operons. The diversity of these mechanisms and the variety of molecules and events involved lead us to suspect that there are many as yet undiscovered mechanisms of transcription attenuation used by bacteria. Potential targets for regulatory events that could modulate transcription termination are RNA polymerase, polymerase accessory proteins, RNA-binding proteins, ribosomal components, polypeptide and ribosome
release factors, terminator factors, and the transcripts themselves. Thus, as mentioned in the introduction, the many potential targets for regulation by transcription attenuation probably offer cells options that are not available for regulation of transcription initiation. A second consideration is the impact of evolution on adoption of successful regulatory strategies. If RNA did serve as the sole genetic material during some early period, as some suggest, and if this RNA was translated by a primitive translational process, then some of the transcription attenuation mechanisms in use today could have evolved from regulatory mechanisms that were operating prior to the appearance of DNA. It seems likely that successful strategies would have been retained.

As we have mentioned, there are only a few well-documented examples of transcription attenuation in eukaryotes. The patterns of gene expression in higher organisms suggest that a common regulatory objective is to provide each genetic unit with a variety of cis regulatory sites or elements that allow an appropriate response to numerous regulatory proteins that are produced or modified during different stages of growth and development. Given the need for regulatory diversity, it seems likely that some forms of transcription attenuation will be found to be common in eukaryotes.

LITERATURE CITED


