REGULATORY CIRCUITS

3.11.1 Introduction

The basic concept of genetic regulation in bacteria is that the expression of a gene may be controlled by a regulator that interacts with a specific sequence or structure in DNA or mRNA at some stage prior to the synthesis of protein. This is an extremely flexible idea with many ramifications. The stage of expression that is controlled can involve transcription, when the target for regulation is DNA; or it can be at translation, when the target for regulation is RNA. When transcription is involved, the level of control can be at initiation or at termination. The regulator can be a protein or an RNA. "Controlled" can mean that the regulator turns off (represses) the target or that it turns on (activates) the target. Expression of many genes can be coordinately controlled by a single regulator gene on the principle that each target contains a copy of the sequence or structure that the regulator recognizes. Regulators may themselves be regulated, most typically in response to small molecules whose supply responds to environmental conditions. Regulators may be controlled by other regulators to make complex circuits.

Let's compare the ways that different types of regulators work.

Protein regulators work on the principle of allostery (see *Molecular Biology Supplement 32.7 Allostery*). The protein has two binding sites, one for a nucleic acid target, the other for a small molecule. Binding of the small molecule to its site changes the conformation in such a way as to alter the affinity of the other site for the nucleic acid. The way in which this happens is known in detail for the Lac repressor (see *Molecular Biology 3.10.14 DNA-binding is regulated by an allosteric change in conformation*). Protein regulators are often multimeric, with a symmetrical organization that allows two subunits to contact a palindromic target on DNA. This can generate cooperative binding effects that create a more sensitive response to regulation.

RNA regulators use changes in secondary structure as the guiding principle. An RNA regulator recognizes its target by the familiar principle of complementary base pairing. **Figure 11.1** shows that the regulator is usually a small RNA molecule with extensive secondary structure, but with a single-stranded region(s) that is complementary to a single-stranded region in its target. The formation of a double helical region between regulator and target can have two types of consequence:

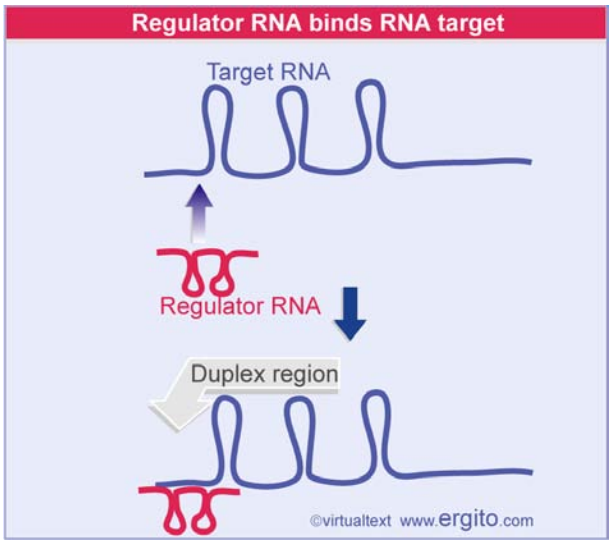

Figure 11.1 A regulator RNA is a small RNA with a single-stranded region that can pair with a single-stranded region in a target RNA.

- Formation of the double helical structure may itself be sufficient. In some cases, protein(s) can bind only to the single-stranded form of the target sequence, and are therefore prevented by acting by duplex formation. In other cases, the duplex region becomes a target for binding, for example, by nucleases that degrade the RNA and therefore prevent its expression.

- Duplex formation may be important because it sequesters a region of the target RNA that would otherwise participate in some alternative secondary structure.

Going beyond the interactions in which a protein or RNA regulates expression of a single gene, we find that bacteria have responses in which the expression of many genes is coordinated. The simplest form of regulating multiple genes occurs when they all have a copy of the same *cis*-acting regulatory elements. Genes that have a copy of the same operator sequence are coordinately repressed by a single repressor protein. Genes that have the same type of promoter are coordinately activated by the production of a sigma factor that causes RNA polymerase to use that promoter.

Regulatory networks are created when one regulator is required for the production of another. This happens in situations in which an ordered temporal expression of genes is required, for example, during phage infection (see *Molecular Biology 3.12.3 Lytic development is controlled by a cascade*) or during the development of a new cell type, such as a spore (see *Molecular Biology 3.9.18 Sigma factors may be organized into cascades*). These use one of the simplest relationships between regulators, in which one regulator is necessary for expression of the next regulator in a series, creating a cascade.

Another type of relationship occurs when one regulator directly regulates the activity of another regulator. For example, a protein that represses or activates expression of a gene may itself by inhibited by an "anti-regulator" that responds to some other signal. **Figure 11.2** shows an example. A series of such relationships can be

extended indefinitely. Some circuits are controlled by a series of regulators each of which antagonizes another. Such circuits allow the cell to control the target set of genes in response to multiple stimuli, since each stimulus can feed into the regulatory circuit at a different point.

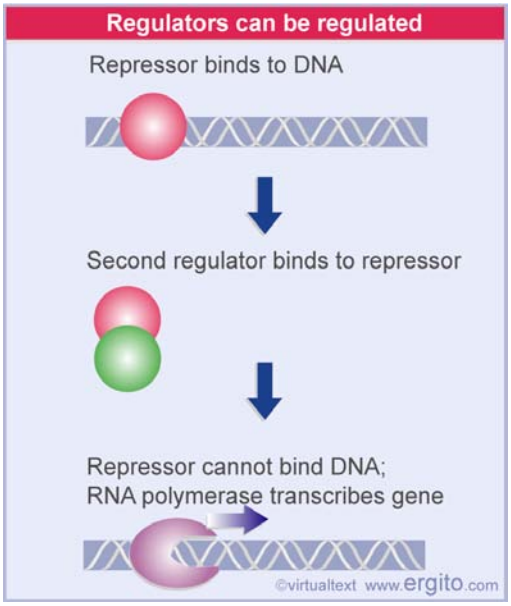

Figure 11.2 A protein that antagonizes a repressor can turn a gene on.

A special type of circuit is created when a protein regulates expression of the gene that codes for it. This is called autogenous control. It allows a protein to regulate its own level of expression without reference to any other circuit. It is often used for regulating levels of proteins that are assembled into macromolecular complexes.

Last updated on 11-15-2002

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.11.1

REGULATORY CIRCUITS

3.11.2 Distinguishing positive and negative control

Key Terms

An **activator** is a protein that stimulates the expression of a gene, typically by acting at a promoter to stimulate RNA polymerase. In eukaryotes, the sequence to which it binds in the promoter is called a response element.

An **inducible** operon is expressed only in the presence of a specific small molecule (the inducer).

A **repressible** operon is expressed unless the small molecule co-repressor is present.

The **derepressed** state describes a gene that is turned on because a small molecule corepressor is absent. It has the same effect as the induced state that is produced by a small molecule inducer for a gene that is regulated by induction. In describing the effect of a mutation, derepressed and constitutive have the same meaning.

Super-repressed is a mutant condition in which a repressible operon cannot be de-repressed, so it is always turned off.

The level of response from a system in the absence of a stimulus is its **basal level.** (The basal level of transcription of a gene is the level that occurs in the absence of any specific activation.)

Key Concepts

- Induction can be achieved by inactivating a repressor or activating an activator.
- Repression can be achieved by activating a repressor or inactivating an activator.

Positive and negative control systems are defined by the response of the operon when no regulator protein is present. The characteristics of the two types of control system are mirror images.

Genes under negative control are expressed unless they are switched off by a repressor protein (see **Figure 10.2**). Any action that interferes with gene expression can provide a negative control. Typically a repressor protein either binds to DNA to prevent RNA polymerase from initiating transcription, or binds to mRNA to prevent a ribosome from initiating translation.

Negative control provides a fail-safe mechanism: if the regulator protein is inactivated, the system functions and so the cell is not deprived of these enzymes. It is easy to see how this might evolve. Originally a system functions constitutively, but then cells able to interfere specifically with its expression acquire a selective advantage by virtue of their increased efficiency.

For genes under positive control, expression is possible only when an active regulator protein is present. The mechanism for controlling an individual operon is

an exact counterpart of negative control, but instead of *interfering* with initiation, the regulator protein is *essential* for it. It interacts with DNA and with RNA polymerase to *assist the initiation event* (see **Figure 10.3***)*. The use of sigma factors to regulate transcription formally is an example of positive control. A positive regulator protein that responds to a small molecule is usually called an **activator**.

It is less obvious how positive control evolved, since the cell must have had the ability to express the regulated genes even before any control existed. Presumably some component of the control system must have changed its role. Perhaps originally it was used as a regular part of the apparatus for gene expression; then later it became restricted to act only in a particular system or systems.

Operons are defined as **inducible** or **repressible** by the nature of their response to the small molecule that regulates their expression. Inducible operons function only in the *presence* of the small-molecule inducer. Repressible operons function only in the *absence* of the small-molecule corepressor (so called to distinguish it from the repressor protein).

The terminology used for repressible systems describes the active state of the operon as **derepressed**; this has the same meaning as *induced.* The condition in which a (mutant) operon cannot be derepressed is sometimes called **super-repressed**; this is the exact counterpart of *uninducible.*

Either positive or negative control could be used to achieve either induction or repression by utilizing appropriate interactions between the regulator protein and the small-molecule inducer or corepressor. **Figure 11.3** summarizes four simple types of control circuit. Induction is achieved when an inducer inactivates a repressor protein or activates an activator protein. Repression is accomplished when a corepressor activates a repressor protein or inactivates an activator protein.

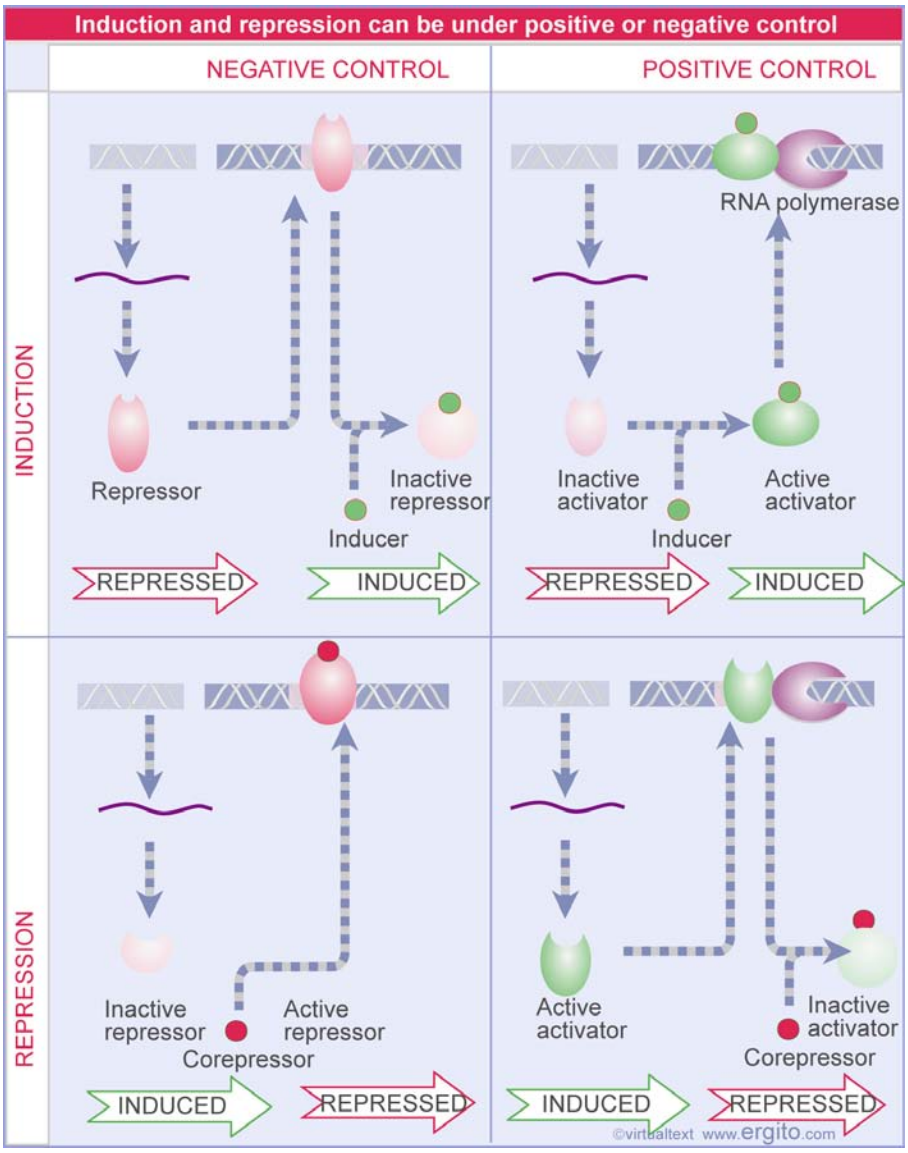

Figure 11.3 Control circuits are versatile and can be designed to allow positive or negative control of induction or repression.

The *trp* operon is a repressible system. Tryptophan is the end product of the reactions catalyzed by a series of biosynthetic enzymes. Both the activity and the synthesis of the tryptophan enzymes are controlled by the level of tryptophan in the cell.

Tryptophan functions as a corepressor that activates a repressor protein. This is the classic mechanism for repression, as seen in **Figure 11.3** (lower left). In conditions when the supply of tryptophan is plentiful, the operon is repressed because the repressor protein·corepressor complex is bound at the operator. When tryptophan is in short supply, the corepressor is inactive, therefore has reduced specificity for the operator, and is stored elsewhere on DNA.

Deprivation of repressor causes ~70-fold increase in the frequency of initiation events at the *trp* promoter. Even under repressing conditions, the structural genes

continue to be expressed at a low **basal level** (sometimes also called the repressed level). The efficiency of repression at the operator is much lower than in the *lac* operon (where the basal level is only ~1/1000 of the induced level).

We have treated both induction and repression as phenomena that rely upon allosteric changes induced in regulator proteins by small molecules. Other types of interactions also can be used to control the activities of regulator proteins. One example is OxyR, a transcriptional activator of genes induced by hydrogen peroxide. The OxyR protein is directly activated by oxidation, so it provides a sensitive measure of oxidative stress. Another common type of signal is phosphorylation of a regulator protein.

REGULATORY CIRCUITS

3.11.3 Glucose repression controls use of carbon sources

Key Terms

Glucose repression (Catabolite repression) describes the decreased expression of many bacterial operons that results from addition of glucose.

Inducer exclusion describes the inhibition of uptake of other carbon sources into the cell that is caused by uptake of glucose.

Key Concepts

- *E. coli* uses glucose in preference to other carbon sources when it has a choice.
- Glucose prevents uptake of alternative carbon sources from the medium.
- Exclusion of the alternative carbon sources from the cell prevents expression of the operons coding for the enzymes that metabolize them.

Bacteria have distinct preferences among potential carbon sources when they are offered a choice. When glucose is available as an energy source, it is used in preference to other sugars. So when *E. coli* finds (for example) both glucose and lactose in the medium, it metabolizes the glucose and represses the use of lactose. The phenomenon of **glucose repression** follows from the ability of glucose to prevent the use of alternative carbon sources. It describes the general repression of transcription of the operons (such as *lac*, *gal*, and *ara)* coding for the enzymes required to metabolize the alternative carbon sources. The same effect is found in many bacteria, although different molecular mechanisms may be responsible in different classes of bacteria (for review see 2563).

The basis for glucose repression in *E. coli* appears to be largely indirect. It involves two mechanisms:

- **Inducer exclusion** describes the ability of glucose to prevent alternative carbon sources from being taken up from the medium. These include the inducers of the operons coding for the alternative metabolic systems. As a result of the exclusion of the small molecule inducers from the cell, the operons cannot be transcribed.
- A general system for activating many operons is provided by the activator protein CRP, which is activated by the small molecule cyclic AMP (see *Molecular Biology 3.11.5 CRP functions in different ways in different target operons*). It was thought for many years that this system is inactivated by glucose, with the result that its target operons are not expressed under conditions of glucose repression. However, this view has been challenged recently, and the exact way in which it is controlled by glucose is not clear.

Figure 11.4 outlines the interactions involved in inducer exclusion. The key molecular component in inducer exclusion is the PTS (phosphoenolpyruvate:glycose phosphotransferase system), a complex of proteins in the bacterial membrane, which simultaneously phosphorylates and transports sugars into the cell (for review see 2561). One of the proteins of this complex (IIA^{Glc}, which is coded by the *crr* gene) becomes dephosphorylated as a result of glucose transport. It then binds to the Lac permease and prevents it from importing lactose into the cell.

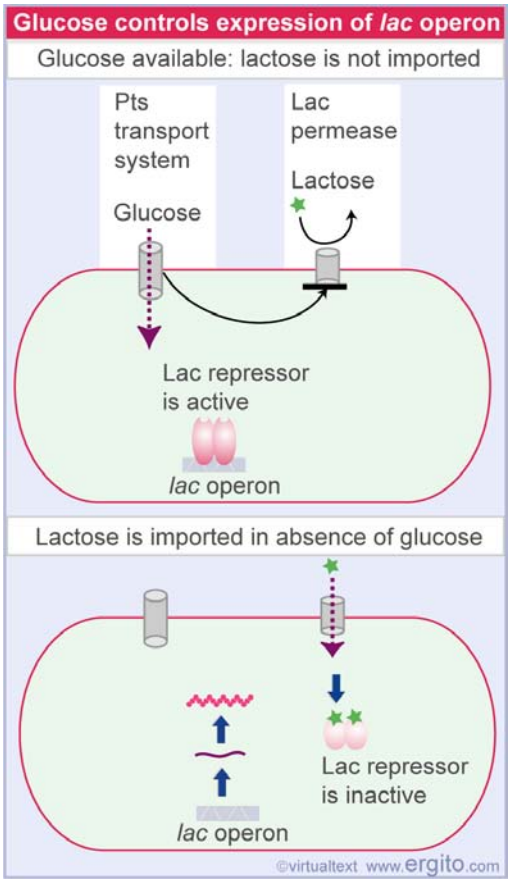

Figure 11.4 Glucose is imported through the Pts system. This causes the Lac permease to be inactivated. The absence of lactose means that the *lac* repressor switches off the operon. When glucose is absent, lactose can be imported, and inactivates the *lac* repressor, so the operon is switched on.

Last updated on 7-29-2002

Reviews

2561. Meadow, N. D., Fox, D. K., and Roseman, S. (1990). *The bacterial phosphoenolpyruvate: glycose phosphotransferase system.* Annu. Rev. Biochem. 59, 497-542.

2563. Stalke, J. and Hillen, W. (2000). *Regulation of carbon catabolism in Bacillus species.* Annu. Rev. Microbiol. 54, 849-880.

REGULATORY CIRCUITS

3.11.4 Cyclic AMP is an inducer that activates CRP to act at many operons

Key Terms

CRP activator (CAP activator) is a positive regulator protein activated by cyclic AMP. It is needed for RNA polymerase to initiate transcription of many operons of *E. coli*.

Adenylate cyclase is an enzyme that uses ATP as a substrate to generate cyclic AMP, in which 5′ and 3′ positions of the sugar ring are connected via a phosphate group.

Key Concepts

- CRP is an activator protein that binds to a target sequence at a promoter.
- A dimer of CRP is activated by a single molecule of cyclic AMP.

So far we have dealt with the promoter as a DNA sequence that is competent to bind RNA polymerase, which then initiates transcription. But there are some promoters at which RNA polymerase cannot initiate transcription without assistance from an ancillary protein. Such proteins are positive regulators, because their presence is necessary to switch on the transcription unit. Typically the activator overcomes a deficiency in the promoter, for example, a poor consensus sequence at –35 or –10.

One of the most widely acting activators is a protein called **CRP activator** that controls the activity of a large set of operons in *E. coli*. The protein is a positive control factor whose presence is necessary to initiate transcription at dependent promoters. CRP is active *only in the presence of cyclic AMP*, which behaves as the classic small-molecule inducer (see **Figure 11.3**; upper right).

Cyclic AMP is synthesized by the enzyme **adenylate cyclase**. The reaction uses ATP as substrate and introduces a 3′–5′ link via phosphodiester bonds, generating the structure drawn in **Figure 11.5**. Mutations in the gene coding for adenylate cyclase (cya^-) do not respond to changes in glucose levels.

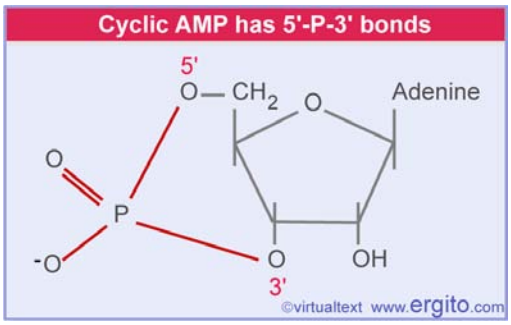

Figure 11.5 Cyclic AMP has a single phosphate group connected to both the 3 ′ and 5 ′ positions of the sugar ring.

The level of cyclic AMP is inversely related to the level of glucose. The basis for this effect lies with the same component of the Pts system that is responsible for controlling lactose uptake. The phosphorylated form of protein IIA^{Glc} stimulates adenylate cyclase. When glucose is imported, the dephosphorylation of IIA^{Glc} leads to a fall in adenylate cyclase activity.

Figure 11.6 shows that reducing the level of cyclic AMP renders the (wild-type) protein unable to bind to the control region, which in turn prevents RNA polymerase from initiating transcription. So the effect of glucose in reducing cyclic AMP levels is to deprive the relevant operons of a control factor necessary for their expression.

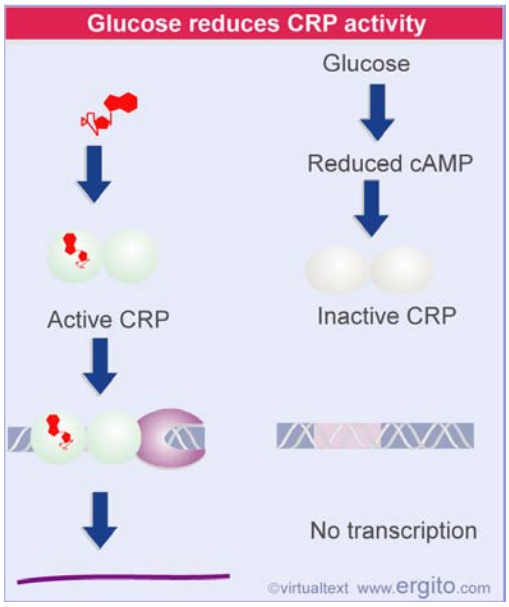

Figure 11.6 By reducing the level of cyclic AMP, glucose inhibits the transcription of operons that require CRP activity.

Last updated on 7-25-2002

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.11.4

REGULATORY CIRCUITS

3.11.5 CRP functions in different ways in different target operons

Key Concepts

- CRP-binding sites lie at highly variable locations relative to the promoter.
- CRP interacts with RNA polymerase, but the details of the interaction depend on the relative locations of the CRP-binding site and the promoter.

The CRP factor binds to DNA, and complexes of cyclic AMP·CRP·DNA can be isolated at each promoter at which it functions. The factor is a dimer of two identical subunits of 22.5 kD, which can be activated by a single molecule of cyclic AMP. A CRP monomer contains a DNA-binding region and a transcription-activating region (for review see 99).

A CRP dimer binds to a site of ~22 bp at a responsive promoter. The binding sites include variations of the consensus sequence given in **Figure 11.7**. Mutations preventing CRP action usually are located within the well conserved pentamer TGTGA / ACACT, which appears to be the essential element in recognition. CRP binds most strongly to sites that contain two (inverted) versions of the pentamer, because this enables both subunits of the dimer to bind to the DNA. Many binding sites lack the second pentamer, however, and in these the second subunit must bind a different sequence (if it binds to DNA). The hierarchy of binding affinities for CRP helps to explain why different genes are activated by different levels of cyclic AMP *in vivo.*

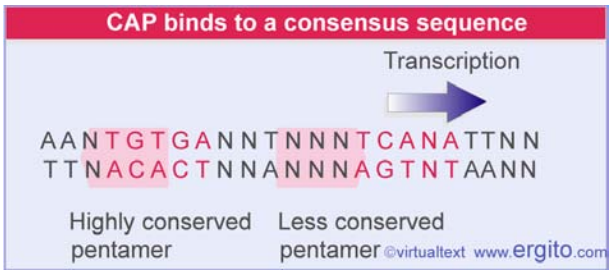

Figure 11.7 The consensus sequence for CRP contains the well conserved pentamer TGTGA and (sometimes) an inversion of this sequence (TCANA).

The action of CRP has the curious feature that its binding sites lie at different locations relative to the startpoint in the various operons that it regulates. And the TGTGA pentamer may lie in either orientation. The three examples summarized in **Figure 11.8** encompass the range of locations:

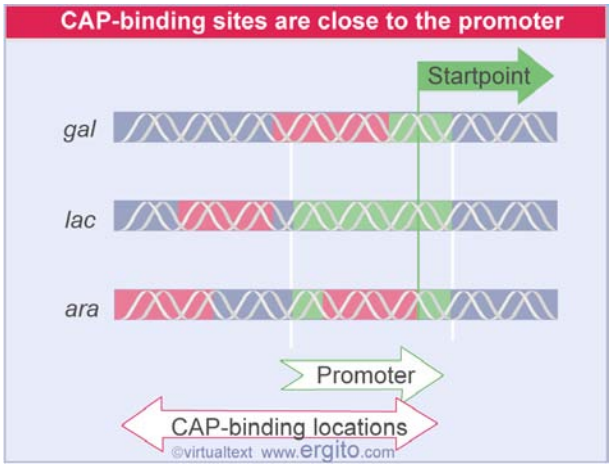

Figure 11.8 The CRP protein can bind at different sites relative to RNA polymerase.

- The CRP-binding site is adjacent to the promoter, as in the *lac* operon, in which the region of DNA protected by CRP is centered on –61. It is possible that two dimers of CRP are bound. The binding pattern is consistent with the presence of CRP largely on one face of DNA, the same face that is bound by RNA polymerase. This location would place the two proteins just about in reach of each other.

- Sometimes the CRP-binding site lies within the promoter, as in the *gal* locus, where the CRP-binding site is centered on –41. It is likely that only a single CRP dimer is bound, probably in quite intimate contact with RNA polymerase, since the CRP-binding site extends well into the region generally protected by the RNA polymerase.

- In other operons, the CRP-binding site lies well upstream of the promoter. In the *ara* region, the binding site for a single CRP is the farthest from the startpoint, centered at –92.

Dependence on CRP is related to the intrinsic efficiency of the promoter. No CRP-dependent promoter has a good –35 sequence and some also lack good –10 sequences. In fact, we might argue that effective control by CRP would be difficult if the promoter had effective –35 and –10 regions that interacted independently with RNA polymerase.

There are in principle two ways in which CRP might activate transcription: it could interact directly with RNA polymerase; or it could act upon DNA to change its structure in some way that assists RNA polymerase to bind. In fact, CRP has effects upon both RNA polymerase and DNA (for review see 101).

Binding sites for CRP at most promoters resemble either *lac* (centered at –61) or *gal* (centered at –41 bp). The basic difference between them is that in the first type (called class I) the CRP-binding site is entirely upstream of the promoter, whereas in the second type (called class II) the CRP-binding site overlaps the binding site for RNA polymerase. (The interactions at the *ara* promoter may be different.)

In both types of promoter, the CRP binding site is centered an integral number of turns of the double helix from the startpoint. This suggests that CRP is bound to the same face of DNA as RNA polymerase. However, the nature of the interaction between CRP and RNA polymerase is different at the two types of promoter.

When the α subunit of RNA polymerase has a deletion in the C-terminal end, transcription appears normal except for the loss of ability to be activated by CRP. CRP has an "activating region", which consists of a small exposed loop of ~ 10 amino acids, that is required for activating both types of its promoters. The activating region is a small patch of amino acids that interacts directly with the α subunit of RNA polymerase to stimulate the enzyme (2565). At class I promoters, this interaction is sufficient. At class II promoters, a second interaction is also required, involving another region of CRP and the N-terminal region of the RNA polymerase α subunit (2564).

Experiments using CRP dimers in which only one of the subunits has a functional transcription-activating region shows that, when CRP is bound at the *lac* promoter, only the activating region of the subunit nearer the startpoint is required, presumably because it touches RNA polymerase (2566). This offers an explanation for the lack of dependence on the orientation of the binding site: the dimeric structure of CRP ensures that one of the subunits is available to contact RNA polymerase, no matter which subunit binds to DNA and in which orientation.

The effect upon RNA polymerase binding depends on the relative locations of the two proteins. At class I promoters, where CRP binds adjacent to the promoter, it increases the rate of initial binding to form a closed complex. At class II promoters, where CRP binds within the promoter, it increases the rate of transition from the closed to open complex.

Last updated on 7-25-2002

Reviews

99. Botsford, J. L. and Harman, J. G. (1992). *Cyclic AMP in prokaryotes*. Microbiol. Rev. 56, 100-122.

101. Kolb, A. (1993). *Transcriptional regulation by cAMP and its receptor protein*. Annu. Rev. Biochem. 62, 749-795.

References

2564. Niu, W., Kim, Y., Tau, G., Heyduk, T., and Ebright, R. H. (1996). *Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase.* Cell 87, 1123-1134.

2565. Zhou, Y., Merkel, T. J., and Ebright, R. H. (1994). *Characterization of the activating region of E. coli catabolite gene activator protein (CAP). II. Role at Class I and class II CAP-dependent promoters.* J. Mol. Biol. 243, 603-610.

2566. Zhou, Y., Busby, S., and Ebright, R. H. (1993). *Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers.* Cell 73, 375-379.

REGULATORY CIRCUITS

3.11.6 CRP bends DNA

Key Concepts

- CRP introduces a 90° bend into DNA at its binding site.

The structure of the CRP-DNA complex is interesting: *the DNA has a bend.* Proteins may distort the double helical structure of DNA when they bind, and several regulator proteins induce a bend in the axis.

Figure 11.9 illustrates a technique that can be used to measure the extent and location of a bend. A target sequence containing the site is cut with different restriction enzymes to generate a set of fragments all of the same length, but each containing the protein-binding site at a different location.

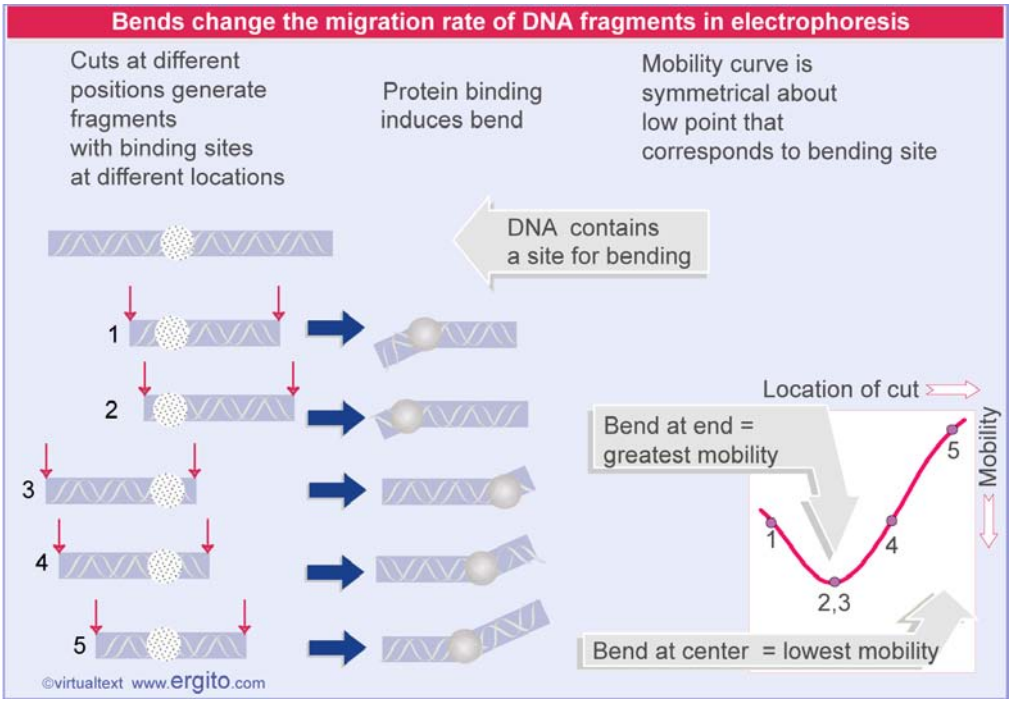

Figure 11.9 Gel electrophoresis can be used to analyze bending.

The fragments move at different speeds in an electrophoretic gel, depending on the position of the bend. (If there is no bend, all fragments move at the same rate.) The greatest impediment to motion, causing the lowest mobility, happens when the bend is in the center of the DNA fragment. The least impediment to motion, allowing the greatest mobility, happens when the bend is at one end.

The results are analyzed by plotting mobility against the site of restriction cutting. The low point on the curve identifies the situation in which the restriction enzyme

has cut the sequence immediately adjacent to the site of bending.

For the interaction of CRP with the *lac* promoter, this point lies at the center of dyad symmetry. The bend is quite severe, >90°, as illustrated in the model of **Figure 11.10**. There is therefore a dramatic change in the organization of the DNA double helix when CRP protein binds. The mechanism of bending is to introduce a sharp kink within the TGTGA consensus sequence. When there are inverted repeats of the consensus, the two kinks in each copy present in a palindrome cause the overall 90° bend. It is possible that the bend has some direct effect upon transcription, but it could be the case that it is needed simply to allow CRP to contact RNA polymerase at the promoter (518).

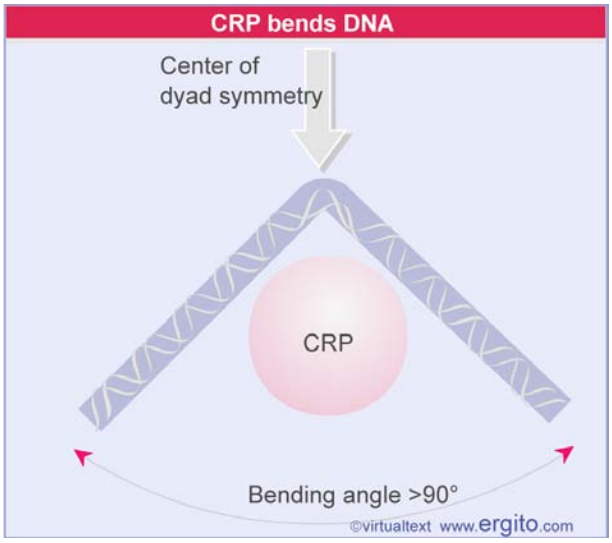

Figure 11.10 CRP bends DNA >90° around the center of symmetry.

Whatever the exact means by which CRP activates transcription at various promoters, it accomplishes the same general purpose: to turn off alternative metabolic pathways when they become unnecessary because the cell has an adequate supply of glucose. Again, this makes the point that coordinate control, of either negative or positive type, can extend over dispersed loci by repetition of binding sites for the regulator protein.

References

518. Gaston, K. A. et al. (1990). *Stringent spacing requirements for transcription activation by CRP*. Cell 62, 733-743.

REGULATORY CIRCUITS

3.11.7 The stringent response produces (p)ppGpp

Key Terms

Stringent response refers to the ability of a bacterium to shut down synthesis of tRNA and ribosomes in a poor-growth medium.

An **alarmone** is a small molecule in bacteria that is produced as a result of stress and which acts to alter the state of gene expression. The unusual nucleotides ppGpp and pppGpp are examples.

ppGpp is guanosine tetraphosphate. Diphosphate groups are attached to both the 5′ and 3′ positions.

pppGpp is a guanosine pentaphosphate, with a triphosphate attached to the 5′ position and a diphosphate attached to the 3′ position.

Key Concepts

- Poor growth conditions cause bacteria to produce the small molecule regulators ppGpp and pppGpp.

When bacteria find themselves in such poor growth conditions that they lack a sufficient supply of amino acids to sustain protein synthesis, they shut down a wide range of activities. This is called the **stringent response**. We can view it as a mechanism for surviving hard times: the bacterium husbands its resources by engaging in only the minimum of activities until nutrient conditions improve, when it reverses the response and again engages its full range of metabolic activities.

The stringent response causes a massive (10-20×) reduction in the synthesis of rRNA and tRNA. This alone is sufficient to reduce the total amount of RNA synthesis to ~5-10% of its previous level. The synthesis of certain mRNAs is reduced, leading to an overall reduction of ~3× in mRNA synthesis. The rate of protein degradation is increased. Many metabolic adjustments occur, as seen in reduced synthesis of nucleotides, carbohydrates, lipids, etc.

The stringent response causes the accumulation of two unusual nucleotides (sometimes called **alarmones**). **ppGpp** is guanosine tetraphosphate, with diphosphates attached to both 5′ and 3′ positions. **pppGpp** is guanosine pentaphosphate, with a 5′ triphosphate group and a 3′ diphosphate. These nucleotides are typical small-molecule effectors that function by binding to target proteins to alter their activities. Sometimes they are known collectively as (p)ppGpp (520; for review see 94).

(p)ppGpp functions to regulate coordinately a large number of cellular activities. Its production is controlled in two ways. A drastic increase in (p)ppGpp is triggered by the stringent response. And there is also a general inverse correlation between (p)ppGpp levels and the bacterial growth rate, which is controlled by some unknown means.

Reviews

94. Cashel, M. and Rudd, K. E. (1987). *The stringent response In E. coli and S. typhimurium*. E. coli and S. typhimurium, 1410-1429.

References

520. Cashel, M. and Gallant, J. (1969). *Two compounds implicated in the function of the RC gene of E. coli*. Nature 221, 838-841.

REGULATORY CIRCUITS

3.11.8 (p)ppGpp is produced by the ribosome

Key Terms

The **idling reaction** results in the production of pppGpp and ppGpp by ribosomes when an uncharged tRNA is present in the A site; this triggers the stringent response.

Relaxed mutants of *E. coli* do not display the stringent response to starvation for amino acids (or other nutritional deprivation).

The **stringent factor** is the protein RelA, which is associated with ribosomes. It synthesizes ppGpp and pppGpp when uncharged aminoacyl-tRNA enters the A site.

Key Concepts

- The stringent factor RelA is a (p)ppGpp synthetase that is associated with ~5% of ribosomes.
- RelA is activated when the A site is occupied by an uncharged tRNA.
- One (p)ppGpp is produced every time an uncharged tRNA enters the A site.

Deprivation of any one amino acid, or mutation to inactivate any aminoacyl-tRNA synthetase, is sufficient to initiate the stringent response. The trigger that sets the entire series of events in train is *the presence of uncharged tRNA in the A site of the ribosome*. Under normal conditions, of course, only aminoacyl-tRNA is placed in the A site by EF-Tu (see *Molecular Biology 2.6.10 Elongation factor Tu loads aminoacyl-tRNA into the A site*). But when there is no aminoacyl-tRNA available to respond to a particular codon, the uncharged tRNA becomes able to gain entry. Of course, this blocks any further progress by the ribosome; and it triggers an **idling reaction**.

The components involved in producing (p)pGpp via the idling reaction have been identified through the existence of **relaxed** (*rel*) mutants. *rel* mutations abolish the stringent response, so that starvation for amino acids does not cause any reduction in stable RNA synthesis or alter any of the other reactions that are usually seen.

The most common site of relaxed mutation lies in the gene *relA*, which codes for a protein called the **stringent factor**. This factor is associated with the ribosomes, although the amount is rather low – say, <1 molecule for every 200 ribosomes. So perhaps only a minority of the ribosomes are able to produce the stringent response.

Ribosomes obtained from stringent bacteria can synthesize ppGpp and pppGpp *in vitro*, provided that the A site is occupied by an uncharged tRNA *specifically responding to the codon.* Ribosomes extracted from relaxed mutants cannot perform this reaction; but they are able to do so if the stringent factor is added.

Figure 11.11 shows the pathways for synthesis of (p)ppGpp. The stringent factor (RelA) is an enzyme that catalyzes the synthetic reaction in which ATP is used to donate a pyrophosphate group to the 3′ position of either GTP or GDP. The formal name for this activity is (p)ppGpp synthetase.

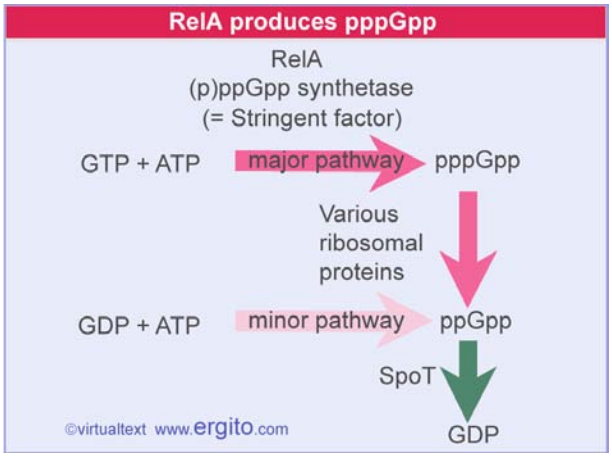

Figure 11.11 Stringent factor catalyzes the synthesis of pppGpp and ppGpp; ribosomal proteins can dephosphorylate pppGpp to ppGpp.

How is ppGpp removed when conditions return to normal? A gene called *spoT* codes for an enzyme that provides the major catalyst for ppGpp degradation. The activity of this enzyme causes ppGpp to be rapidly degraded, with a half-life of ~20 sec; so the stringent response is reversed rapidly when synthesis of (p)ppGpp ceases. *spoT* mutants have elevated levels of ppGpp, and grow more slowly as a result.

The RelA enzyme uses GTP as substrate more frequently, so that pppGpp is the predominant product. However, pppGpp is converted to ppGpp by several enzymes; among those able to perform this dephosphorylation are the translation factors EF-Tu and EF-G. The production of ppGpp via pppGpp is the most common route, *and ppGpp is the usual effector of the stringent response.*

The response of the ribosome to entry of uncharged tRNA is compared with normal protein synthesis in **Figure 11.12**. When EF-Tu places aminoacyl-tRNA in the A site, peptide bond synthesis is followed by ribosomal movement. But when uncharged tRNA is paired with the codon in the A site, the ribosome remains stationary and engages in the idling reaction (521).

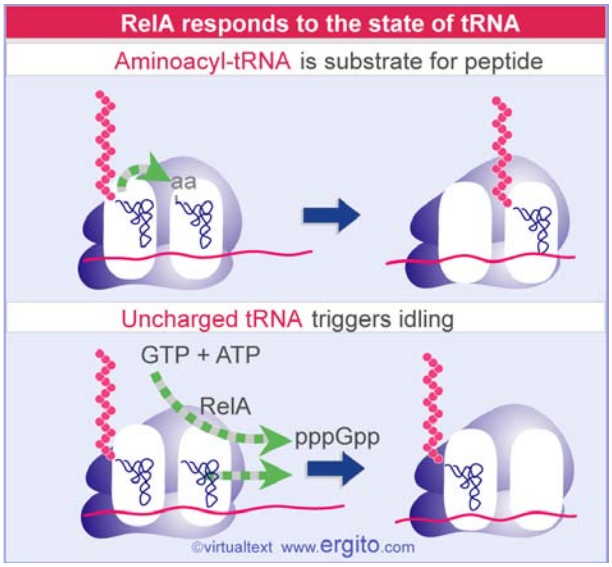

Figure 11.12 In normal protein synthesis, the presence of aminoacyl-tRNA in the A site is a signal for peptidyl transferase to transfer the polypeptide chain, followed by movement catalyzed by EF-G; but under stringent conditions, the presence of uncharged tRNA causes RelA protein to synthesize (p)ppGpp and to expel the tRNA.

How does the state of the ribosome control the activity of RelA enzyme? An indication of the nature of the interaction is revealed by relaxed mutations in another locus, originally called *relC*, which turns out to be the same as *rplK*, which codes for the 50S subunit protein L11. This protein is located in the vicinity of the A and P sites, in a position to respond to the presence of a properly paired but uncharged tRNA in the A site. A conformational change in this protein or some other component could activate the RelA enzyme, so that the idling reaction occurs instead of polypeptide transfer from the peptidyl-tRNA.

One round of (p)ppGpp synthesis is associated with release of the uncharged tRNA from the A site, so that synthesis of (p)ppGpp is a continuing response to the level of uncharged tRNA. So under limiting conditions, a ribosome stalls when no aminoacyl-tRNA is available to respond to the codon in the A site. Entry of uncharged tRNA triggers the synthesis of a (p)ppGpp molecule, and the resulting expulsion of the uncharged tRNA allows the situation to be reassessed. Depending upon the availability of aminoacyl-tRNA, the ribosome resumes polypeptide synthesis or undertakes another idling reaction.

References

521. Haseltine, W. A. and Block, R. (1973). *Synthesis of guanosine tetra and pentaphosphate requires the presence of a codon specific uncharged tRNA in the acceptor site of ribosomes*. Proc. Natl. Acad. Sci. USA 70, 1564-1568.

REGULATORY CIRCUITS

3.11.9 pGpp has many effects

Key Concepts

- ppGpp inhibits transcription of rRNA.

What does ppGpp do? It is an effector for controlling several reactions, including the inhibition of transcription. Many effects have been reported, among which two stand out:

- *Initiation of transcription is specifically inhibited at the promoters of operons coding for rRNA.* Mutations of stringently regulated promoters can abolish stringent control, which suggests that the effect requires an interaction with specific promoter sequences.
- *The elongation phase of transcription of many or most templates is reduced by ppGpp.* The cause is increased pausing by RNA polymerase. This effect is responsible for the general reduction in transcription efficiency when ppGpp is added *in vitro.*

The use of ppGpp is just one aspect of a more general regulatory network that relates production of ribosomes to the growth rate. The level of protein synthesis increases in proportion with the growth rate. This is accomplished by increasing the production of ribosomes as cells grow more rapidly. The cell therefore needs some general indicator of growth rate that can be used to control the synthesis of ribosomes. The indicator appears to be NTP levels, and the target for their action is the control of transcription of rRNA.

Figure 11.13 summarizes the systems that are used to control rRNA transcription in response to growth rate. Under conditions of starvation, ppGpp is produced, and (among its various actions) inhibits initiation at the promoters of the *rrn* loci that code for rRNA. As growth rate increases, the levels of ATP and GTP increase. These increase the rate of initiation at the *rrn* promoters (for review see 103).

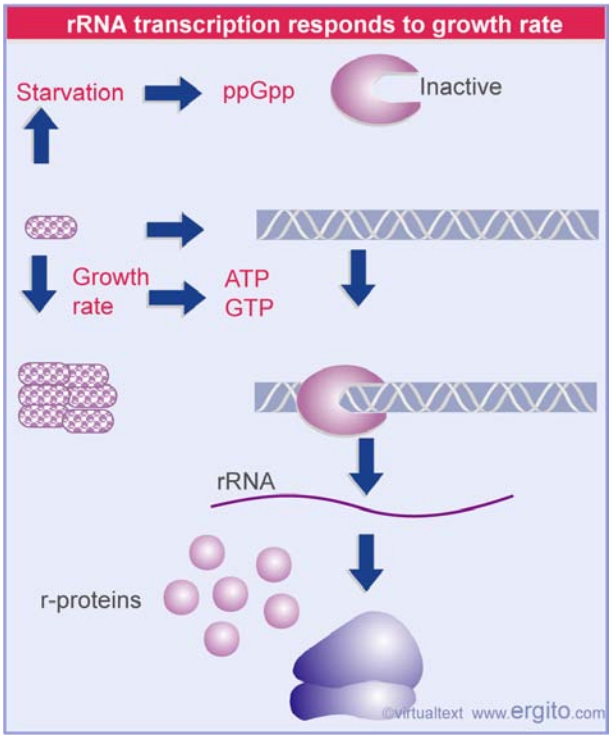

Figure 11.13 Nucleotide levels control initiation of rRNA transcription.

The *rrn* promoters in *E. coli* form atypical open complexes with RNA polymerase. The open complexes are unusually unstable. The result is that the main factor governing the rate of initiation becomes the decay rate of the open complex. Increased concentration of the initiating nucleotide (which is ATP at six of the *rrn* promoters and GTP at the seventh) drives the initiation reaction forward by stabilizing the open complex.

The level of rRNA controls the production of ribosomes (by a feedback loop in which the absence of rRNA inhibits synthesis of ribosomal proteins, see **Figure 11.18**). This means that the production of ribosomes, and thus the level of protein synthesis, in turn respond to the levels of ATP and GTP, which reflect the nutritional condition of the cell. So concentrations of particular nucleotides control ribosome synthesis in response to normal changes in growth rate and the more extreme conditions of starvation.

Reviews

103. Condon, C., Squires, C., and Squires, C. L. (1995). *Control of rRNA transcription in E. coli*. Microbiol. Rev. 59, 623-645.

REGULATORY CIRCUITS

3.11.10 Translation can be regulated

Key Concepts

- A repressor protein can regulate translation by preventing a ribosome from binding to an initiation codon.
- Accessibility of initiation codons in a polycistronic mRNA can be controlled by changes in the structure of the mRNA that occur as the result of translation.

Translational control is a notable feature of operons coding for components of the protein synthetic apparatus. The operon provides an arrangement for *coordinate* regulation of a group of structural genes. But, superimposed on it, further controls, such as those at the level of translation, may create *differences* in the extent to which individual genes are expressed.

A similar type of mechanism is used to achieve translational control in several systems. *Repressor function is provided by a protein that binds to a target region on mRNA to prevent ribosomes from recognizing the initiation region.* Formally this is equivalent to a repressor protein binding to DNA to prevent RNA polymerase from utilizing a promoter. **Figure 11.14** illustrates the most common form of this interaction, in which the regulator protein binds directly to a sequence that includes the AUG initiation codon, thereby preventing the ribosome from binding.

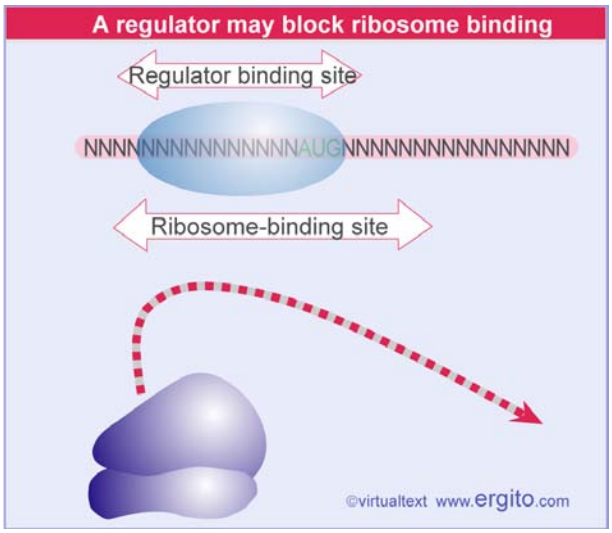

Figure 11.14 A regulator protein may block translation by binding to a site on mRNA that overlaps the ribosome-binding site at the initiation codon.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.3.11.10 to view properly.

Some examples of translational repressors and their targets are summarized in **Figure 11.15**. A classic example is the coat protein of the RNA phage R17; it binds to a hairpin that encompasses the ribosome binding site in the phage mRNA. Similarly the T4 RegA protein binds to a consensus sequence that includes the AUG initiation codon in several T4 early mRNAs; and T4 DNA polymerase binds to a sequence in its own mRNA that includes the Shine-Dalgarno element needed for ribosome binding.

Translational repressors bind to mRNA

Repressor	Target Gene	Site of Action
R17 coat protein	R17 replicase	hairpin that includes ribosome binding site
T4 RegA	early T4 mRNAs	various sequences including initiation codon
T4 DNA polymerase	T4 DNA polymerase	Shine-Dalgarno sequence
T4 p32	gene 32	single-stranded 5' leader

Figure 11.15 Proteins that bind to sequences within the initiation regions of mRNAs may function as translational repressors.

Another form of translational control occurs when translation of one cistron requires changes in secondary structure that depend on translation of a preceding cistron. This happens during translation of the RNA phages, whose cistrons always are expressed in a set order. **Figure 11.16** shows that the phage RNA takes up a secondary structure in which only one initiation sequence is accessible; the second cannot be recognized by ribosomes because it is base paired with other regions of the RNA. However, translation of the first cistron disrupts the secondary structure, allowing ribosomes to bind to the initiation site of the next cistron. In this mRNA, secondary structure controls translatability.

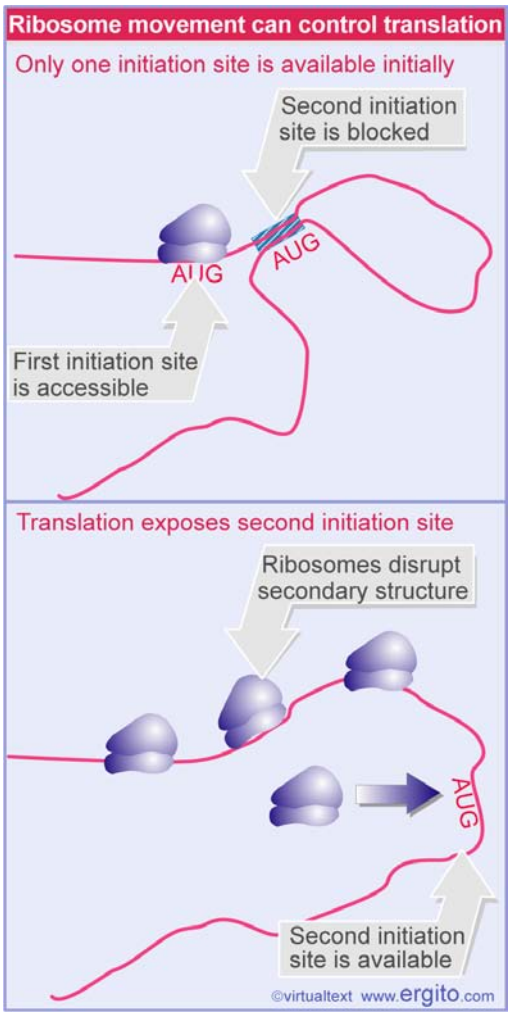

Figure 11.16 Secondary structure can control initiation. Only one initiation site is available in the RNA phage, but translation of the first cistron changes the conformation of the RNA so that other initiation site(s) become available.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.3.11.10 to view properly.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.11.10

REGULATORY CIRCUITS

3.11.11 r-protein synthesis is controlled by autogenous regulation

Key Terms

An **r-protein** is one of the proteins of the ribosome.

Autogenous control describes the action of a gene product that either inhibits (negative autogenous control) or activates (positive autogenous control) expression of the gene coding for it.

Key Concepts

- Translation of an r-protein operon can be controlled by a product of the operon that binds to a site on the polycistronic mRNA.

About 70 or so proteins constitute the apparatus for bacterial gene expression. The ribosomal proteins are the major component, together with the ancillary proteins involved in protein synthesis. The subunits of RNA polymerase and its accessory factors make up the remainder. The genes coding for ribosomal proteins, protein-synthesis factors, and RNA polymerase subunits all are intermingled and organized into a small number of operons. Most of these proteins are represented only by single genes in *E. coli.*

Coordinate controls ensure that these proteins are synthesized in amounts appropriate for the growth conditions: when bacteria grow more rapidly, they devote a greater proportion of their efforts to the production of the apparatus for gene expression. An array of mechanisms is used to control the expression of the genes coding for this apparatus and to ensure that the proteins are synthesized at comparable levels that are related to the levels of the rRNAs.

The organization of six operons is summarized in **Figure 11.17**. About half of the genes for ribosomal proteins (**r-proteins**) map in four operons that lie close together (named *str, spc, S10,* and α simply for the first one of the functions to have been identified in each case). The *rif* and *L11* operons lie together at another location.

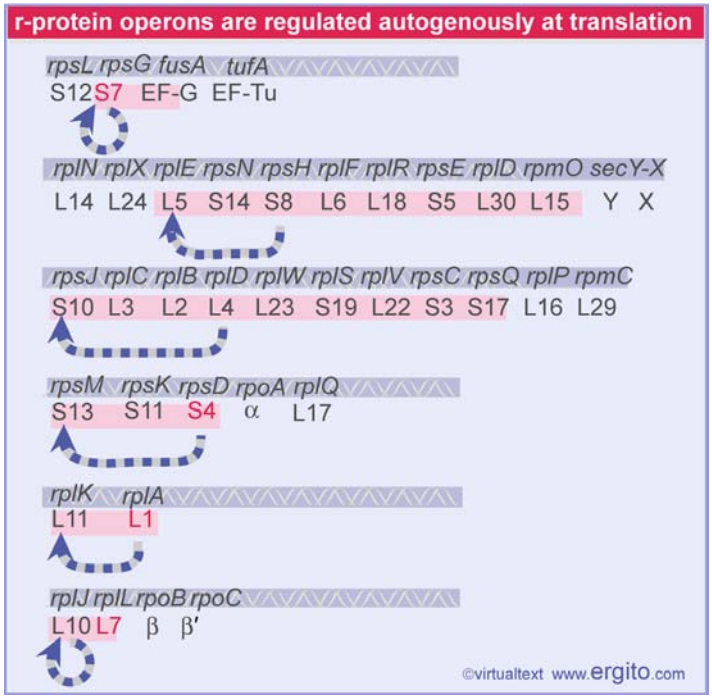

Figure 11.17 Genes for ribosomal proteins, protein synthesis factors, and RNA polymerase subunits are interspersed in a small number of operons that are autonomously regulated. The regulator is named in red; the proteins that are regulated are shaded in pink.

Each operon codes for a variety of functions. The *str* operon has genes for small subunit ribosomal proteins as well as for EF-Tu and EF-G. The *spc* and *S10* operons have genes interspersed for both small and large ribosomal subunit proteins. The α operon has genes for proteins of both ribosomal subunits as well as for the α subunit of RNA polymerase. The *rif* locus has genes for large subunit ribosomal proteins and for the β and β′ subunits of RNA polymerase.

All except one of the ribosomal proteins are needed in equimolar amounts, which must be coordinated with the level of rRNA. The dispersion of genes whose products must be equimolar, and their intermingling with genes whose products are needed in different amounts, pose some interesting problems for coordinate regulation.

A feature common to all of the operons described in **Figure 11.17** is regulation of some of the genes by one of the products. In each case, the gene coding for the regulatory product is itself one of the targets for regulation. **Autogenous** regulation occurs whenever a protein (or RNA) regulates its own production. In the case of the r-protein operons, the regulatory protein inhibits expression of a contiguous set of genes within the operon, so this is an example of negative autogenous regulation.

In each case, *accumulation of the protein inhibits further synthesis of itself and of some other gene products.* The effect often is exercised at the level of translation of the polycistronic mRNA. Each of the regulators is a ribosomal protein that binds directly to rRNA. *Its effect on translation is a result of its ability also to bind to its own mRNA.* The sites on mRNA at which these proteins bind either overlap the sequence where translation is initiated or lie nearby and probably influence the

accessibility of the initiation site by inducing conformational changes. For example, in the S10 operon, protein L4 acts at the very start of the mRNA to inhibit translation of S10 and the subsequent genes. The inhibition may result from a simple block to ribosome access, as illustrated previously in **Figure 11.15**, or it may prevent a subsequent stage of translation. In two cases (including S4 in the α operon), the regulatory protein stabilizes a particular secondary structure in the mRNA that prevents the initiation reaction from continuing after the 30S subunit has bound (517).

The use of r-proteins that bind rRNA to establish autogenous regulation immediately suggests that this provides a mechanism to link r-protein synthesis to rRNA synthesis. A generalized model is depicted in **Figure 11.18**. Suppose that the binding sites for the autogenous regulator r-proteins on rRNA are much stronger than those on the mRNAs. Then so long as any free rRNA is available, the newly synthesized r-proteins will associate with it to start ribosome assembly. There will be no free r-protein available to bind to the mRNA, so its translation will continue. But as soon as the synthesis of rRNA slows or stops, free r-proteins begin to accumulate. Then they are available to bind their mRNAs, repressing further translation. This circuit ensures that each r-protein operon responds in the same way to the level of rRNA: as soon as there is an excess of r-protein relative to rRNA, synthesis of the protein is repressed (for review see 92).

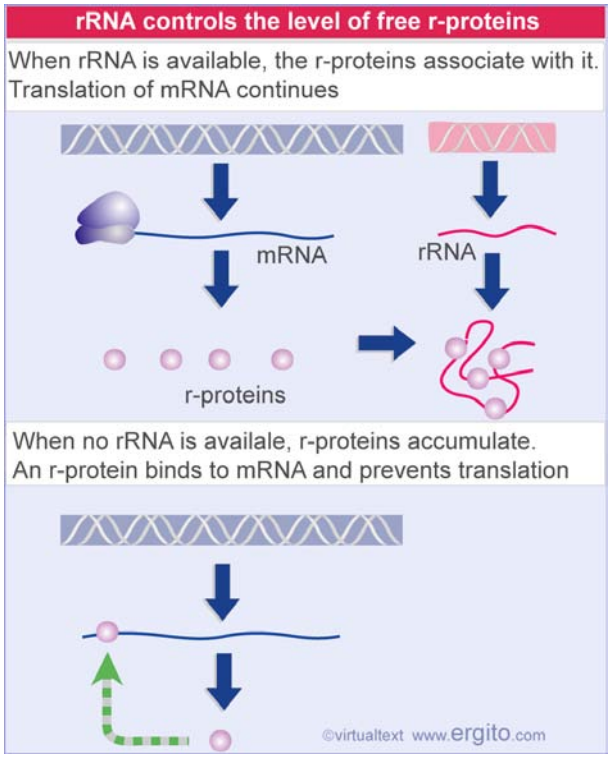

Figure 11.18 Translation of the r-protein operons is autogenously controlled and responds to the level of rRNA.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.3.11.11 to view properly.

Reviews

92. Nomura, M. et al. (1984). *Regulation of the synthesis of ribosomes and ribosomal components.* Annu. Rev. Biochem. 53, 75-117.

References

517. Baughman, G. and Nomura, M. (1983). *Localization of the target site for translational regulation of the L11 operon and direct evidence for translational coupling in E. coli*. Cell 34, 979-988.

REGULATORY CIRCUITS

3.11.12 Phage T4 p32 is controlled by an autogenous circuit

Key Concepts

- p32 binds to its own mRNA to prevent initiation of translation.

Autogenous regulation has been placed on a quantitative basis for gene *32* of phage T4. The protein (p32) plays a central role in genetic recombination, DNA repair, and replication, in which its function is exercised by virtue of its ability to bind to single-stranded DNA. Nonsense mutations cause the inactive protein to be overproduced. *So when the function of the protein is prevented, more of it is made.* This effect occurs at the level of translation; the gene *32* mRNA is stable, and remains so irrespective of the behavior of the protein product.

Figure 11.19 presents a model for the gene *32* control circuit. When single-stranded DNA is present in the phage-infected cell, it sequesters p32. However, in the absence of single-stranded DNA, or at least in conditions in which there is a surplus of p32, the protein prevents translation of its own mRNA. The effect is mediated directly by p32 binding to mRNA to prevent initiation of translation. Probably this occurs at an A·T-rich region that surrounds the ribosome binding site.

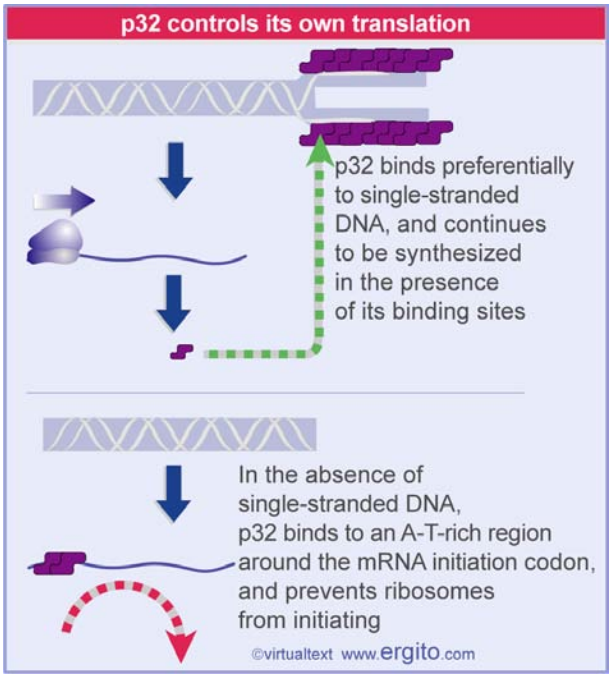

Figure 11.19 Excess gene 32 protein (p32) binds to its own mRNA to prevent ribosomes from initiating translation.

Two features of the binding of p32 to the site on mRNA are required to make the control loop work effectively:

- The affinity of p32 for the site on gene *32* mRNA must be significantly lower than its affinity for single-stranded DNA. The equilibrium constant for binding RNA is in fact almost two orders of magnitude below that for single-stranded DNA.
- But the affinity of p32 for the mRNA must be significantly greater than the affinity for other RNA sequences. It is influenced by base composition and by secondary structure; an important aspect of the binding to gene *32* mRNA is that the regulatory region has an extended sequence lacking secondary structure.

Using the known equilibrium constants, we can plot the binding of p32 to its target sites as a function of protein concentration. **Figure 11.20** shows that at concentrations below 10^{-6} M, p32 binds to single-stranded DNA. At concentrations $>10^{-6}$ M, it binds to gene *32* mRNA. At yet greater concentrations, it binds to other mRNA sequences, with a range of affinities.

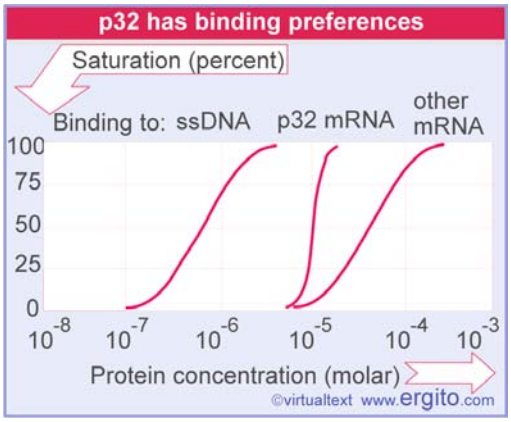

Figure 11.20 Gene 32 protein binds to various substrates with different affinities, in the order single-stranded DNA, its own mRNA, and other mRNAs. Binding to its own mRNA prevents the level of p32 from rising $>10^{-6}$ M.

These results imply that the level of p32 should be autoregulated to be $<10^{-6}$ M, which corresponds to ~2000 molecules per bacterium. This fits well with the measured level of 1000-2000 molecules/cell.

A feature of autogenous control is that each regulatory interaction is unique: a protein acts only on the mRNA responsible for its own synthesis. Phage T4 provides an example of a more general translational regulator, coded by the gene *regA*, which represses the expression of several genes that are transcribed during early infection. RegA protein prevents the translation of mRNAs for these genes by competing with 30S subunits for the initiation sites on the mRNA. Its action is a direct counterpart to the function of a repressor protein that binds multiple operators.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.11.12

REGULATORY CIRCUITS

3.11.13 Autogenous regulation is often used to control synthesis of macromolecular assemblies

Key Concepts

- The precursor to microtubules, free tubulin protein, inhibits translation of tubulin mRNA.

Autogenous regulation is a common type of control among proteins that are incorporated into macromolecular assemblies. The assembled particle itself may be unsuitable as a regulator, because it is too large, too numerous, or too restricted in its location. But the need for synthesis of its components may be reflected in the pool of free precursor subunits. If the assembly pathway is blocked for any reason, free subunits accumulate and shut off the unnecessary synthesis of further components (for review see 98).

Eukaryotic cells have a common system in which autogenous regulation of this type occurs. Tubulin is the monomer from which microtubules, a major filamentous system of all eukaryotic cells, are synthesized. The production of tubulin mRNA is controlled by the free tubulin pool. When this pool reaches a certain concentration, the production of further tubulin mRNA is prevented. Again, the principle is the same: tubulin sequestered into its macromolecular assembly plays no part in regulation, but the level of the free precursor pool determines whether further monomers are added to it.

The target site for regulation is a short sequence at the start of the coding region. We do not know yet what role this sequence plays, but two models are illustrated in **Figure 11.21**. Tubulin may bind directly to the mRNA; or it may bind to the nascent polypeptide representing this region. Whichever model applies, excess tubulin causes tubulin mRNA that is located on polysomes to be degraded, so the consequence of the reaction is to make the tubulin mRNA unstable.

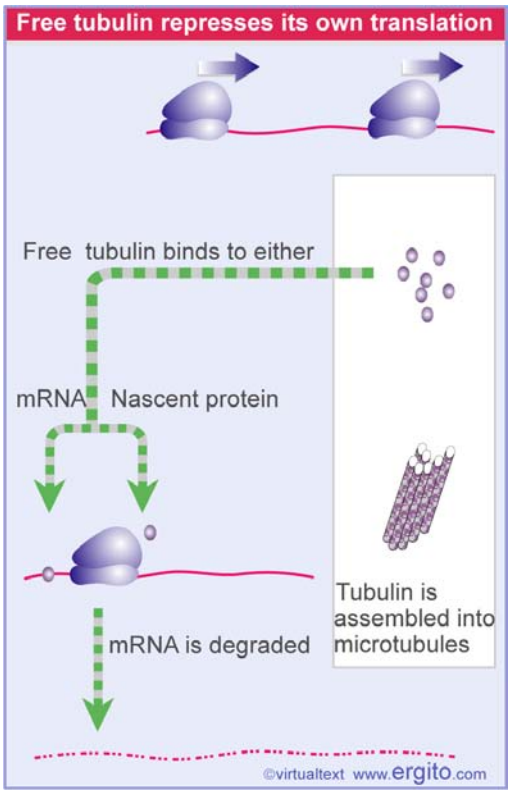

Figure 11.21 Tubulin is assembled into microtubules when it is synthesized. Accumulation of excess free tubulin induces instability in the tubulin mRNA by acting at a site at the start of the reading frame in mRNA or at the corresponding position in the nascent protein.

Autogenous control is an *intrinsically* self-limiting system, by contrast with the *extrinsic* control that we discussed previously. A repressor protein's ability to bind an operator may be controlled by the level of an extraneous small molecule, which activates or inhibits its activity. But in the case of autogenous regulation, the critical parameter is the concentration of the protein itself.

Reviews

98. Gold, L. (1988). *Posttranscriptional regulatory mechanisms in E. coli.* Annu. Rev. Biochem. 57, 199-223.

REGULATORY CIRCUITS

3.11.14 Alternative secondary structures control attenuation

Key Terms

Attenuation describes the regulation of bacterial operons by controlling termination of transcription at a site located before the first structural gene.

An **attenuator** is a terminator sequence at which attenuation occurs.

Key Concepts

- Termination of transcription can be attenuated by controlling formation of the necessary hairpin structure in RNA.
- The most direct mechanisms for attenuation involve proteins that either stabilize or destabilize the hairpin.

RNA structure provides an opportunity for regulation in both prokaryotes and eukaryotes. Its most common role occurs when an RNA molecule can take up alternative secondary structures by utilizing different schemes for intramolecular base pairing. The properties of the alternative conformations may be different. This type of mechanism can be used to regulate the termination of transcription, when the alternative structures differ in whether they permit termination. Another means of controlling conformation (and thereby function) is provided by the cleavage of an RNA; by removing one segment of an RNA, the conformation of the rest may be altered. It is possible also for a (small) RNA molecule to control the activity of a target RNA by base pairing with it; the role of the small RNA is directly analogous to that of a regulator protein (see *Molecular Biology 3.11.19 Small RNA molecules can regulate translation*). The ability of an RNA to shift between different conformations with regulatory consequences is the nucleic acid's alternative to the allosteric changes of conformation that regulate protein function. Both these mechanisms allow an interaction at one site in the molecule to affect the structure of another site.

Several operons are regulated by **attenuation**, a mechanism that controls the ability of RNA polymerase to read through an **attenuator**, which is an intrinsic terminator located at the beginning of a transcription unit. *The principle of attenuation is that some external event controls the formation of the hairpin needed for intrinsic termination.* If the hairpin is allowed to form, termination prevents RNA polymerase from transcribing the structural genes. If the hairpin is prevented from forming, RNA polymerase elongates through the terminator, and the genes are expressed. Different types of mechanisms are used in different systems for controlling the structure of the RNA.

Attenuation may be regulated by proteins that bind to RNA, either to stabilize or to destabilize formation of the hairpin required for termination. **Figure 11.22** shows an example in which a protein prevents formation of the terminator hairpin. The activity

of such a protein may be intrinsic or may respond to a small molecule in the same manner as a repressor protein responds to corepressor.

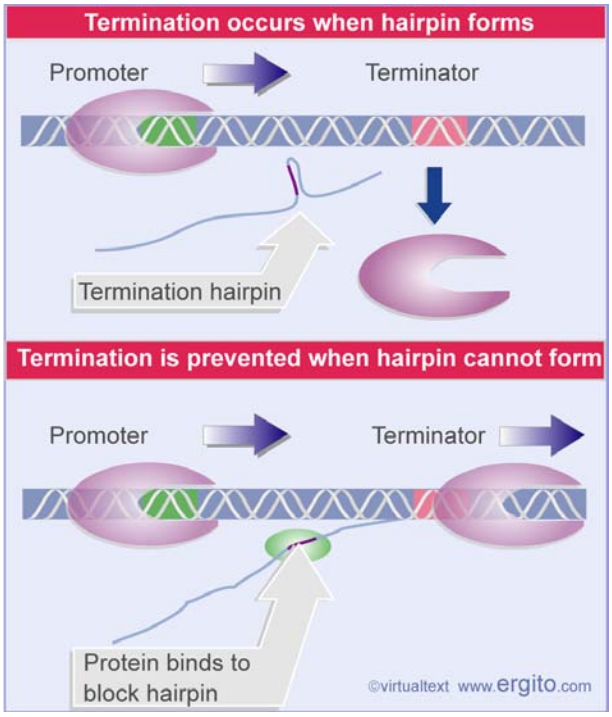

Figure 11.22 Attenuation occurs when a terminator hairpin in RNA is prevented from forming.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.11.14

REGULATORY CIRCUITS

3.11.15 Termination of B. subtilis trp genes is controlled by tryptophan and by tRNA ^{Trp}

Key Concepts

- A terminator protein called TRAP is activated by tryptophan to prevent transcription of trp genes.
- Activity of TRAP is (indirectly) inhibited by uncharged tRNA ^{Trp}.

The circuitry that controls transcription via termination can use both direct and indirect means to respond to the level of small molecule products or substrates.

In *B. subtilis*, a protein called TRAP (formerly called MtrB) is activated by tryptophan to bind to a sequence in the leader of the nascent transcript (915, 919; for review see 920). TRAP forms a multimer of 11 subunits. Each subunit binds a single tryptophan amino acid and a trinucleotide (GAG or UAG) of RNA. The RNA is wound in a circle around the protein (921). **Figure 11.23** shows that the result is to ensure the availability of the regions that are required to form the terminator hairpin. The termination of transcription then prevents production of the tryptophan biosynthetic enzymes. In effect, TRAP is a terminator protein that responds to the level of tryptophan. In the absence of TRAP, an alternative secondary structure precludes the formation of the terminator hairpin.

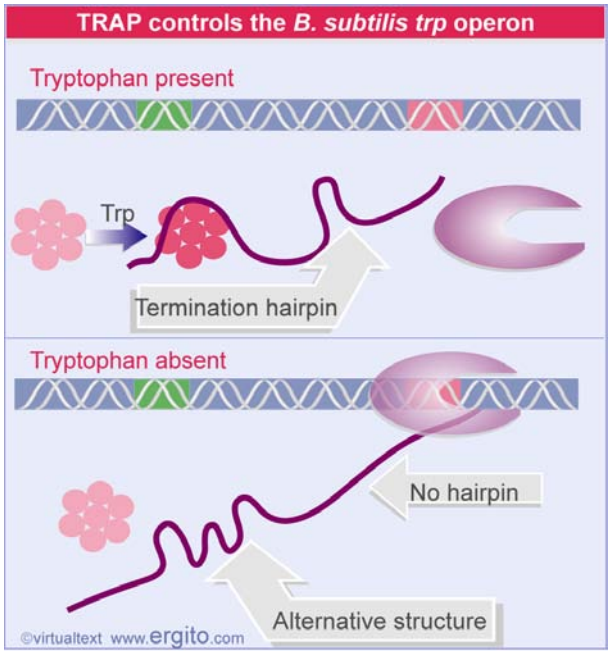

Figure 11.23 TRAP is activated by tryptophan and binds to trp mRNA. This allows the termination hairpin to form, with the result that RNA polymerase terminates, and the genes are not expressed. In the absence of tryptophan, TRAP does not bind, and the mRNA adopts a structure that prevents the terminator hairpin from forming.

However, the TRAP protein in turn is also controlled by tRNA ^{Trp}. **Figure 11.24** shows that uncharged tRNA ^{Trp} binds to the mRNA for a protein called anti-TRAP (AT). This suppresses formation of a termination hairpin in the mRNA. The uncharged tRNA ^{Trp} also increases the translation of the mRNA. The combined result of these two actions to increase synthesis of anti-TRAP, which binds to TRAP, and prevents it from repressing the tryptophan operon. By this complex series of events, the absence of tryptophan generates the uncharged tRNA, which causes synthesis of anti-TRAP, which prevents function of TRAP, which causes expression of tryptophan genes (2068).

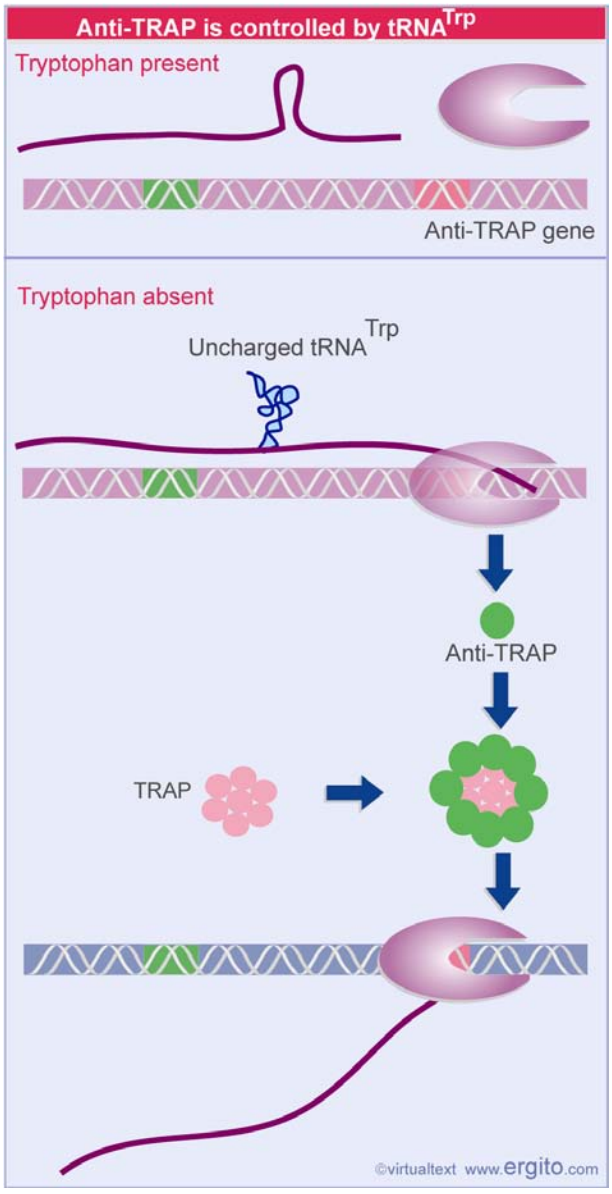

Figure 11.24 Under normal conditions (in the presence of tryptophan) transcription terminates before the anti-TRAP gene. When tryptophan is absent, uncharged $tRNA^{Trp}$ base pairs with the anti-TRAP mRNA, preventing formation of the terminator hairpin, thus causing expression of anti-TRAP.

Expression of the *B. subtilis trp* genes is therefore controlled by both tryptophan and $tRNA^{Trp}$. When tryptophan is present, there is no need for it to be synthesized. This is accomplished when tryptophan activates TRAP and therefore inhibits expression of the enzymes that synthesize tryptophan. The presence of uncharged tRNATrp indicates that there is a shortage of tryptophan. The uncharged tRNA activates the anti-TRAP, and thereby activates transcription of the *trp* genes.

Last updated on 7-21-2003

Reviews

920. Gollnick, P. (1994). *Regulation of the B. subtilis trp operon by an RNA-binding protein.* Mol. Microbiol. 11, 991-997.

References

915. Otridge, J. and Gollnick, P. (1993). *MtrB from B. subtilis binds specifically to trp leader RNA in a tryptophan-dependent manner.* Proc. Natl. Acad. Sci. USA 90, 128-132.

919. Babitzke, P. and Yanoksy, C. (1993). *Reconstitution of B. subtilis trp attenuation in vitro with TRAP, the trp RNA-binding attenuation protein.* Proc. Natl. Acad. Sci. USA 90, 133-137.

921. Antson, A. A. et al. (1999). *Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA.* Nature 401, 235-242.

2068. Valbuzzi, A. and Yanofsky, C. (2001). *Inhibition of the B. subtilis regulatory protein TRAP by the TRAP-inhibitory protein, AT.* Science 293, 2057-2059.

REGULATORY CIRCUITS

3.11.16 The *E. coli tryptophan* operon is controlled by attenuation

Key Concepts

- An attenuator (intrinsic terminator) is located between the promoter and the first gene of the *trp* cluster.
- The absence of tryptophan suppresses termination and results in a 10× increase in transcription.

A complex regulatory system is used in *E. coli* (where attenuation was originally discovered). The changes in secondary structure that control attenuation are determined by the position of the ribosome on mRNA. **Figure 11.25** shows that termination requires that *the ribosome can translate a leader segment that precedes the trp genes in the mRNA.* When the ribosome translates the leader region, a termination hairpin forms at terminator 1. But when the ribosome is prevented from translating the leader, the termination hairpin does not form, and RNA polymerase transcribes the coding region. *This mechanism of antitermination therefore depends upon the ability of external circumstances to influence ribosome movement in the leader region* (for review see 89).

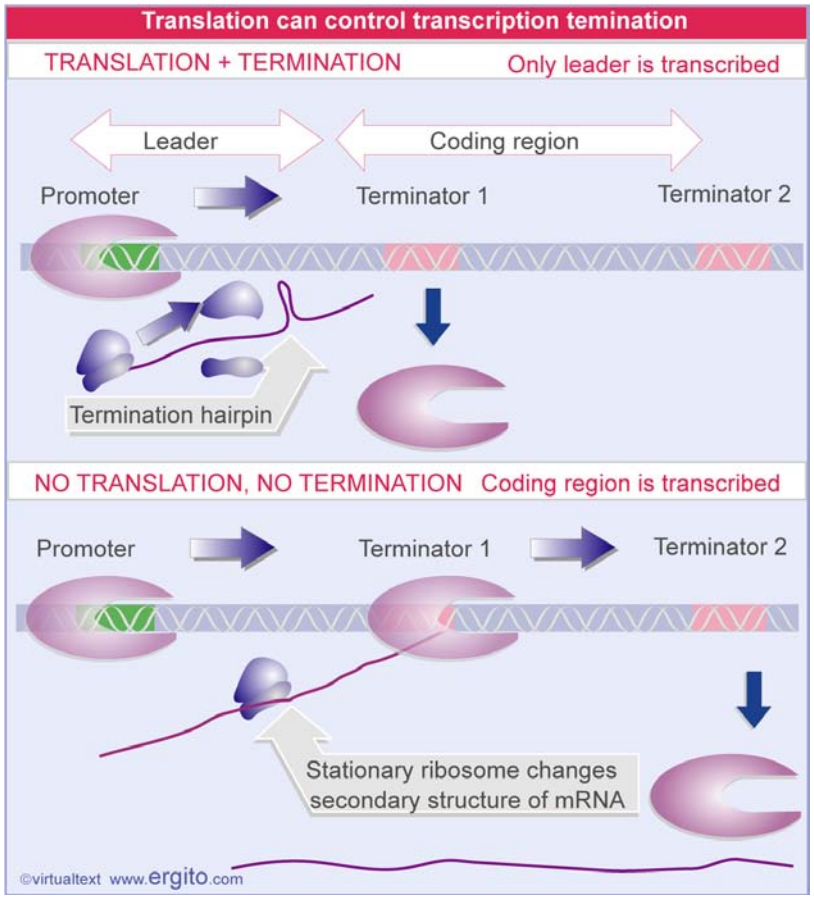

Figure 11.25 Termination can be controlled via changes in RNA secondary structure that are determined by ribosome movement.

The *trp* operon consists of five structural genes arranged in a contiguous series, coding for the three enzymes that convert chorismic acid to tryptophan. **Figure 11.26** shows that transcription starts at a promoter at the left end of the cluster. *trp* operon expression is controlled by two separate mechanisms. Repression of expression is exercised by a repressor protein (coded by the unlinked gene *trpR*) that binds to an operator that is adjacent to the promoter. Attenuation controls the progress of RNA polymerase into the operon by regulating whether termination occurs at a site preceding the first structural gene.

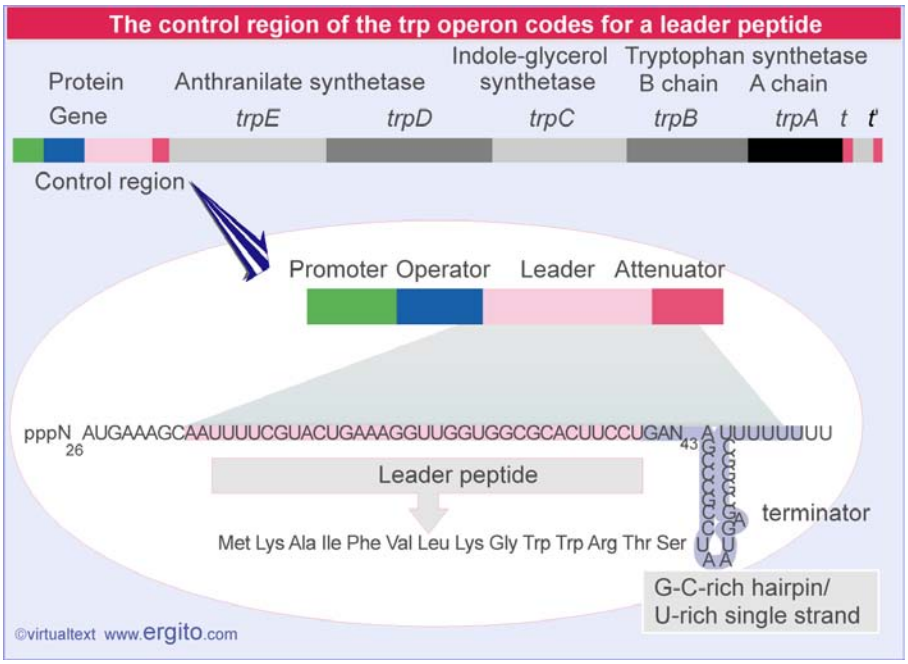

Figure 11.26 The *trp* operon consists of five contiguous structural genes preceded by a control region that includes a promoter, operator, leader peptide coding region, and attenuator.

Attenuation was first revealed by the observation that deleting a sequence between the operator and the *trpE* coding region can increase the expression of the structural genes. This effect is independent of repression: both the basal and derepressed levels of transcription are increased. So this site influences events that occur *after* RNA polymerase has set out from the promoter (irrespective of the conditions prevailing at initiation).

An attenuator (intrinsic terminator) is located between the promoter and the *trpE* gene. It provides a barrier to transcription into the structural genes. RNA polymerase terminates there, either *in vivo* or *in vitro*, to produce a 140-base transcript.

Termination at the attenuator responds to the level of tryptophan, as illustrated in **Figure 11.27**. In the presence of adequate amounts of tryptophan, termination is efficient. But in the absence of tryptophan, RNA polymerase can continue into the structural genes.

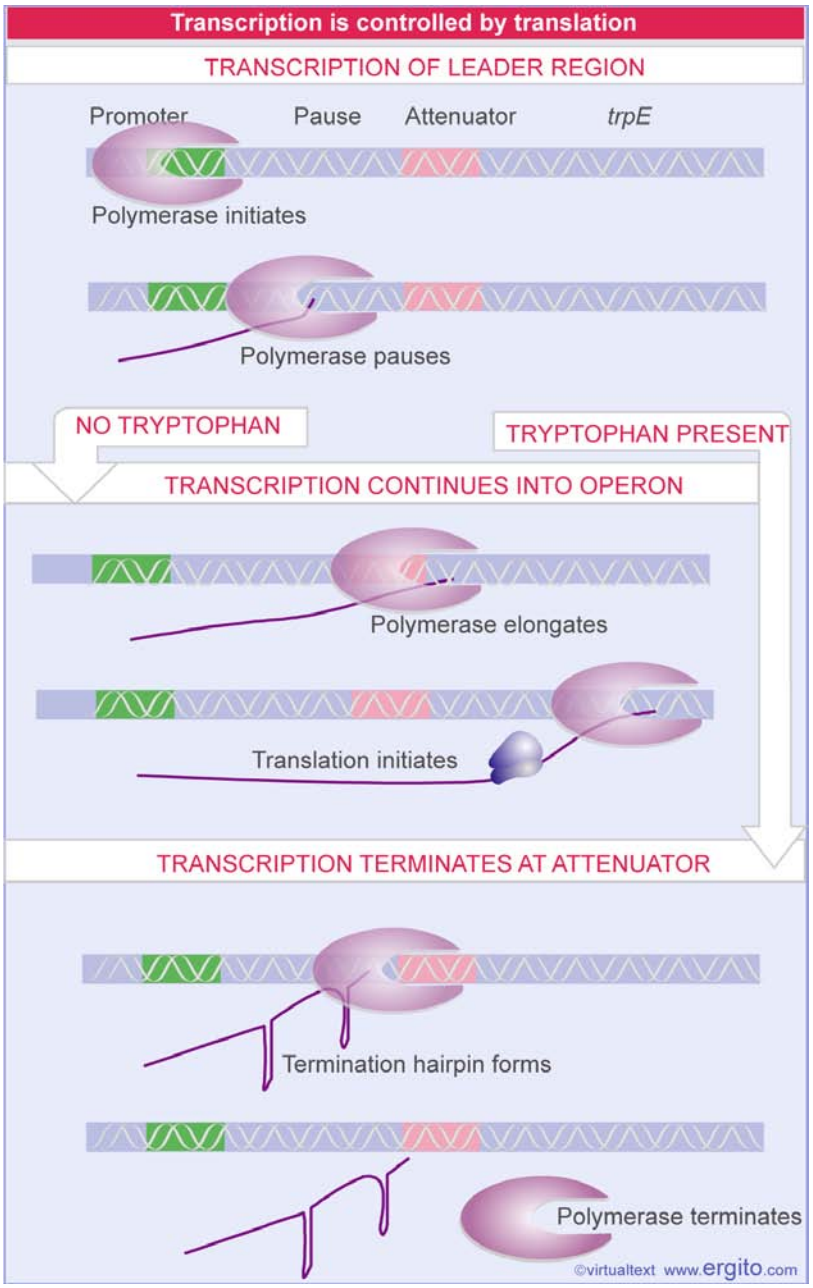

Figure 11.27 An attenuator controls the progression of RNA polymerase into the *trp* genes. RNA polymerase initiates at the promoter and then proceeds to position 90, where it pauses before proceeding to the attenuator at position 140. In the absence of tryptophan, the polymerase continues into the structural genes (*trpE* starts at +163). In the presence of tryptophan there is ~90% probability of termination to release the 140-base leader RNA.

Repression and attenuation respond in the same way to the level of tryptophan. When tryptophan is present, the operon is repressed; and most of the RNA polymerases that escape from the promoter then terminate at the attenuator. When tryptophan is removed, RNA polymerase has free access to the promoter, and also is no longer compelled to terminate prematurely.

Attenuation has ~10× effect on transcription. When tryptophan is present, termination is effective, and the attenuator allows only ~10% of the RNA polymerases to proceed. In the absence of tryptophan, attenuation allows virtually all of the polymerases to proceed. Together with the ~70× increase in initiation of transcription that results from the release of repression, this allows an ~700-fold range of regulation of the operon.

Reviews

89. Yanofsky, C. (1981). *Attenuation in the control of expression of bacterial operons.* Nature 289, 751-758.

REGULATORY CIRCUITS

3.11.17 Attenuation can be controlled by translation

Key Terms

The **leader peptide** is the product that would result from translation of a short coding sequence used to regulate transcription of the tryptophan operon by controlling ribosome movement.

Ribosome stalling describes the inhibition of movement that occurs when a ribosome reaches a codon for which there is no corresponding charged aminoacyl-tRNA.

Key Concepts

- The leader region of the *trp* operon has a 14 codon open reading frame that includes two codons for tryptophan.
- The structure of RNA at the attenuator depends on whether this reading frame is translated.
- In the presence of tryptophan, the leader is translated, and the attenuator is able to form the hairpin that causes termination.
- In the absence of tryptophan, the ribosome stalls at the tryptophan codons and an alternative secondary structure prevents formation of the hairpin, so that transcription continues.

How can termination of transcription at the attenuator respond to the level of tryptophan? The sequence of the leader region suggests a mechanism. It has a short coding sequence that could represent a **leader peptide** of 14 amino acids. **Figure 11.26** shows that it contains a ribosome binding site whose AUG codon is followed by a short coding region that contains two successive codons for tryptophan. When the cell runs out of tryptophan, ribosomes initiate translation of the leader peptide, but stop when they reach the Trp codons. The sequence of the mRNA suggests that this **ribosome stalling** influences termination at the attenuator.

The leader sequence can be written in alternative base-paired structures. The ability of the ribosome to proceed through the leader region controls transitions between these structures. The structure determines whether the mRNA can provide the features needed for termination.

Figure 11.28 draws these structures. In the first, region *1* pairs with region *2*; and region *3* pairs with region *4*. The pairing of regions *3* and *4* generates the hairpin that precedes the U_8 sequence: this is the essential signal for intrinsic termination. Probably the RNA would take up this structure in lieu of any outside intervention.

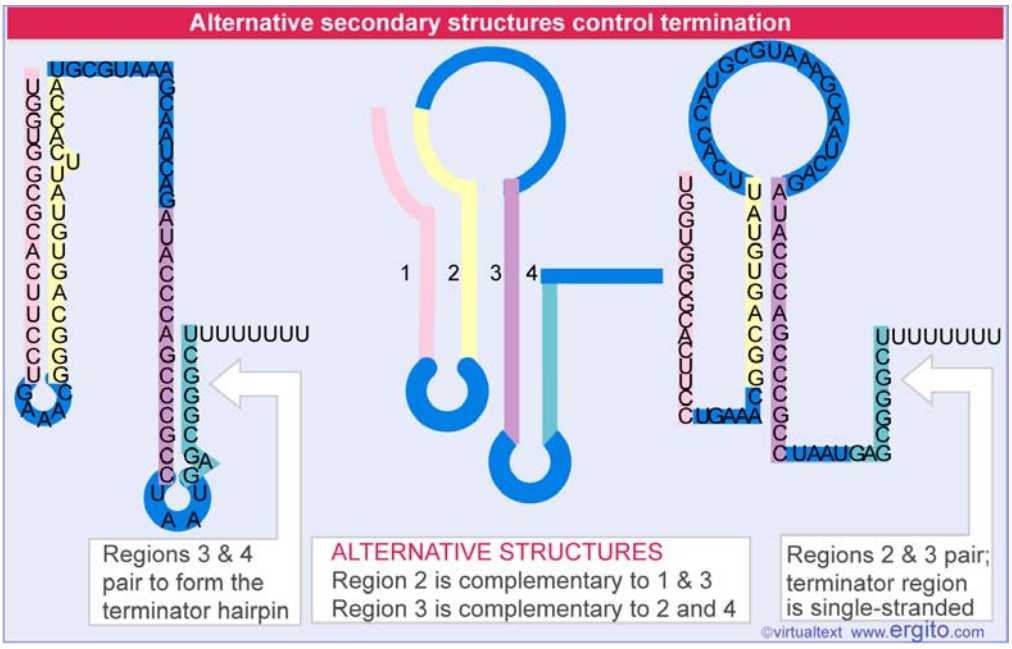

Figure 11.28 The trp leader region can exist in alternative base-paired conformations. The center shows the four regions that can base pair. Region 1 is complementary to region 2, which is complementary to region 3, which is complementary to region 4. On the left is the conformation produced when region 1 pairs with region 2, and region 3 pairs with region 4. On the right is the conformation when region 2 pairs with region 3, leaving regions 1 and 4 unpaired.

A different structure is formed if region *1* is prevented from pairing with region *2*. In this case, region *2* is free to pair with region *3*. Then region *4* has no available pairing partner; so it is compelled to remain single-stranded. So the terminator hairpin cannot be formed.

Figure 11.29 shows that the position of the ribosome can determine which structure is formed, in such a way that termination is attenuated only in the absence of tryptophan. The crucial feature is the position of the Trp codons in the leader peptide coding sequence.

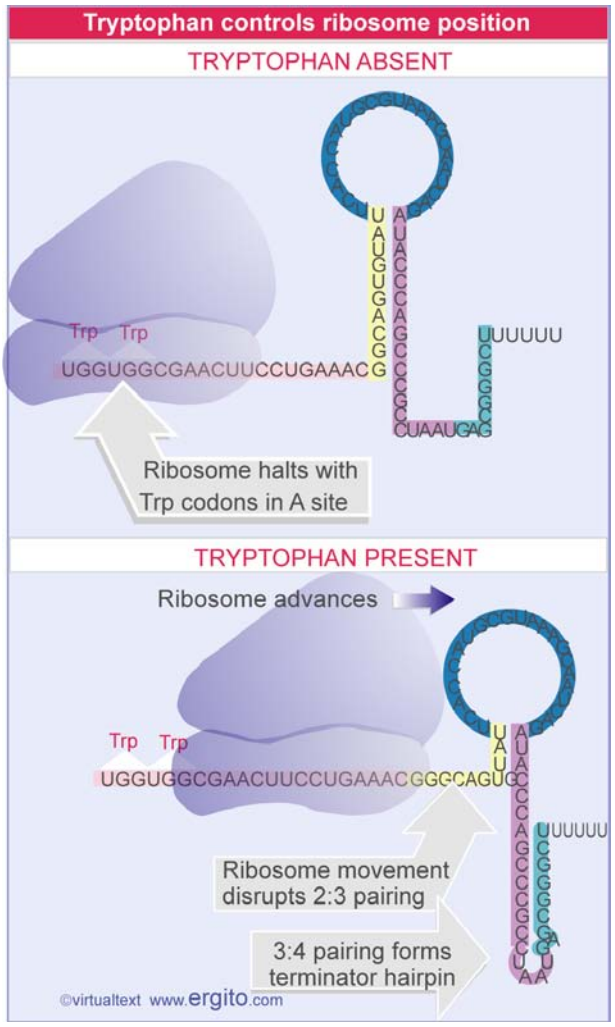

Figure 11.29 The alternatives for RNA polymerase at the attenuator depend on the location of the ribosome, which determines whether regions 3 and 4 can pair to form the terminator hairpin.

When tryptophan is present, ribosomes are able to synthesize the leader peptide. They continue along the leader section of the mRNA to the UGA codon, which lies between regions *1* and *2*. As shown in the lower part of the figure, by progressing to this point, the ribosomes extend over region *2* and prevent it from base pairing. The result is that region *3* is available to base pair with region *4*, generating the terminator hairpin. Under these conditions, therefore, RNA polymerase terminates at the attenuator.

When there is no tryptophan, ribosomes stall at the Trp codons, which are part of region *1*, as shown in the upper part of the figure. So region *1* is sequestered within the ribosome and cannot base pair with region *2*. This means that regions *2* and *3* become base paired before region *4* has been transcribed. This compels region *4* to remain in a single-stranded form. In the absence of the terminator hairpin, RNA polymerase continues transcription past the attenuator (522; 523; for review see 96).

Control by attenuation requires a precise timing of events. For ribosome movement to determine formation of alternative secondary structures that control termination, *translation of the leader must occur at the same time when RNA polymerase approaches the terminator site.* A critical event in controlling the timing is the presence of a site that causes the RNA polymerase to pause at base 90 along the leader. The RNA polymerase remains paused until a ribosome translates the leader peptide. Then the polymerase is released and moves off toward the attenuation site. By the time it arrives there, secondary structure of the attenuation region has been determined.

Figure 11.30summarizes the role of Trp-tRNA in controlling expression of the operon. *By providing a mechanism to sense the inadequacy of the supply of Trp-tRNA, attenuation responds directly to the need of the cell for tryptophan in protein synthesis* (for review see 97).

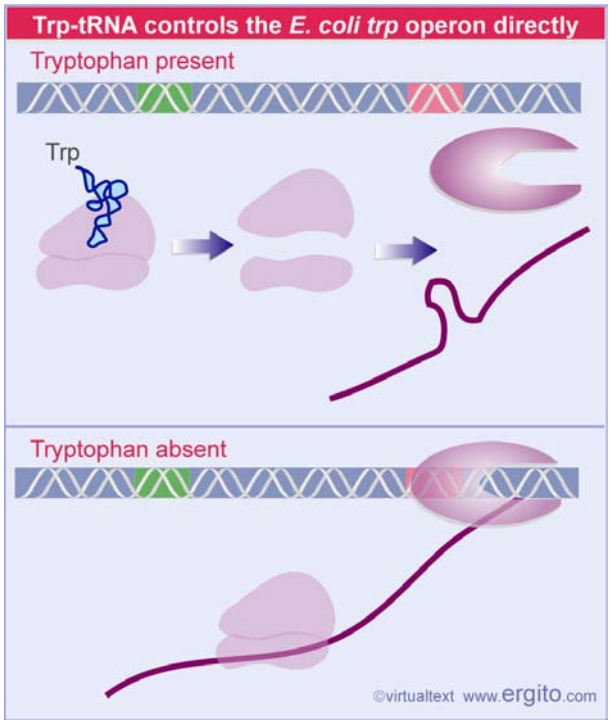

Figure 11.30 In the presence of tryptophan tRNA, ribosomes translate the leader peptide and are released. This allows hairpin formation, so that RNA polymerase terminates. In the absence of tryptophan tRNA, the ribosome is blocked, the termination hairpin cannot form, and RNA polymerase continues.

How widespread is the use of attenuation as a control mechanism for bacterial operons? It is used in at least six operons that code for enzymes concerned with the biosynthesis of amino acids. So a feedback from the level of the amino acid available for protein synthesis (as represented by the availability of aminoacyl-tRNA) to the production of the enzymes may be common (for review see 91).

The use of the ribosome to control RNA secondary structure in response to the availability of an aminoacyl-tRNA establishes an inverse relationship between the presence of aminoacyl-tRNA and the transcription of the operon, equivalent to a

situation in which aminoacyl-tRNA functions as a corepressor of transcription. Since the regulatory mechanism is mediated by changes in the formation of duplex regions, attenuation provides a striking example of the importance of secondary structure in the termination event, and of its use in regulation.

E. coli and *B. subtilis* therefore use the same types of mechanisms, involving control of mRNA structure in response to the presence or absence of a tRNA, but they have combined the individual interactions in different ways. The end result is the same: to inhibit production of the enzymes when there is an excess supply of the amino acid, and to activate production when a shortage is indicated by the accumulation of uncharged $tRNA^{Trp}$.

Last updated on 10-17-2001

Reviews

91. Bauer, C. E. et al. (1983). *Attenuation in bacterial operons*. Gene Function in Prokaryotes, 65-89.

96. Landick, R. and Yanofsky, C. (1987). *Transcription attenuation In E. coli and S. typhimurium*. E. coli and S. typhimurium, 1276-1301.

97. Yanofsky, C. and Crawford, I. P. (1987). *The tryptophan operon*. E. coli and S. typhimurium, 1453-1472.

References

522. Lee, F. and Yanofsky, C. (1977). *Transcription termination at the trp operon attenuators of E. coli and S. typhimurium: RNA secondary structure and regulation of termination*. Proc. Natl. Acad. Sci. USA 74, 4365-4368.

523. Zurawski, G. et al. (1978). *Translational control of transcription termination at the attenuator of the E. coli tryptophan operon*. Proc. Natl. Acad. Sci. USA 75, 5988-5991.

REGULATORY CIRCUITS

3.11.18 Antisense RNA can be used to inactivate gene expression

Key Terms

An **antisense gene** codes for an (antisense) RNA that has a complementary sequence to an RNA that is its target.

Key Concepts

- Antisense genes block expression of their targets when introduced into eukaryotic cells.

Base pairing offers a powerful means for one RNA to control the activity of another. There are many cases in both prokaryotes and eukaryotes where a (usually rather short) single-stranded RNA base pairs with a complementary region of an mRNA, and as a result prevents expression of the mRNA. One of the early illustrations of this effect was provided by an artificial situation, in which **antisense genes** were introduced into eukaryotic cells.

Antisense genes are constructed by reversing the orientation of a gene with regard to its promoter, so that the "antisense" strand is transcribed, as illustrated in **Figure 11.31**. Synthesis of antisense RNA can inactivate a target RNA in either prokaryotic or eukaryotic cells. An antisense RNA is in effect a synthetic RNA regulator. An antisense thymidine kinase gene inhibits synthesis of thymidine kinase from the endogenous gene. Quantitation of the effect is not entirely reliable, but it seems that an excess (perhaps a considerable excess) of the antisense RNA may be necessary (524).

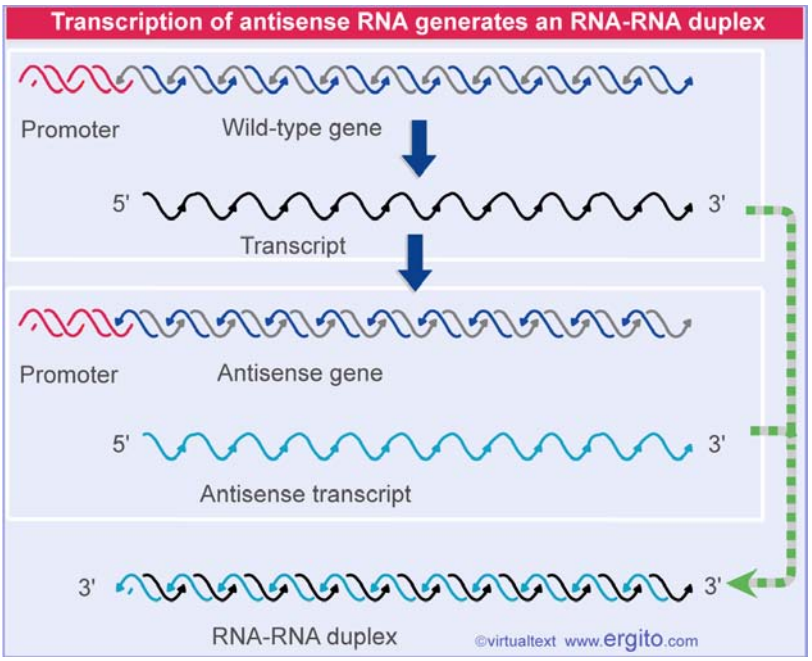

Figure 11.31 Antisense RNA can be generated by reversing the orientation of a gene with respect to its promoter, and can anneal with the wild-type transcript to form duplex RNA.

At what level does the antisense RNA inhibit expression? It could in principle prevent transcription of the authentic gene, processing of its RNA product, or translation of the messenger. Results with different systems show that the inhibition depends on formation of RNA·RNA duplex molecules, but this can occur either in the nucleus or in the cytoplasm. In the case of an antisense gene stably carried by a cultured cell, sense-antisense RNA duplexes form in the nucleus, preventing normal processing and/or transport of the sense RNA. In another case, injection of antisense RNA into the cytoplasm inhibits translation by forming duplex RNA in the 5′ region of the mRNA.

This technique offers a powerful approach for turning off genes at will; for example, the function of a regulatory gene can be investigated by introducing an antisense version. An extension of this technique is to place the antisense gene under control of a promoter itself subject to regulation. Then the target gene can be turned off and on by regulating the production of antisense RNA. This technique allows investigation of the importance of the timing of expression of the target gene.

References

524. Izant, J. G. and Weintraub, H. (1984). *Inhibition of thymidine kinase gene expression by antisense RNA: a molecular approach to genetic analysis*. Cell 36, 1007-1015.

REGULATORY CIRCUITS

3.11.19 Small RNA molecules can regulate translation

Key Concepts

- A regulator RNA functions by forming a duplex region with a target RNA.
- The duplex may block initiation of translation, cause termination of transcription, or create a target for an endonuclease.

Repressors and activators are *trans*-acting proteins. Yet the formal circuitry of a regulatory network could equally well be constructed by using an RNA as regulator. In fact, the original model for the operon left open the question of whether the regulator might be RNA or protein. Indeed, the construction of synthetic antisense RNAs turns out to mimic a class of RNA regulators that is becoming of increasing importance.

Like a protein regulator, a small regulator RNA is an independently synthesized molecule that diffuses to a target site consisting of a specific nucleotide sequence. The target for a regulator RNA is a single-stranded nucleic acid sequence. The regulator RNA functions by complementarity with its target, at which it can form a double-stranded region.

We can imagine two general mechanisms for the action of a regulator RNA:

- Formation of a duplex region with the target nucleic acid directly prevents its ability to function, by forming or sequestering a specific site. **Figure 11.32** illustrates the situation in which a protein that binds to single-stranded RNA is prevented from acting by formation of a duplex. **Figure 11.33** shows the opposite type of relationship in which the formation of a double-stranded region creates a target site for an endonuclease that destroys the RNA target.
- Formation of a duplex region in one part of the target molecule changes the conformation of another region, thus indirectly affecting its function. **Figure 11.34** shows an example. The mechanism is essentially similar to the use of secondary structure in attenuation (see *Molecular Biology 3.11.14 Alternative secondary structures control attenuation*), except that the interacting regions are on different RNA molecules instead of being part of the same RNA molecule.

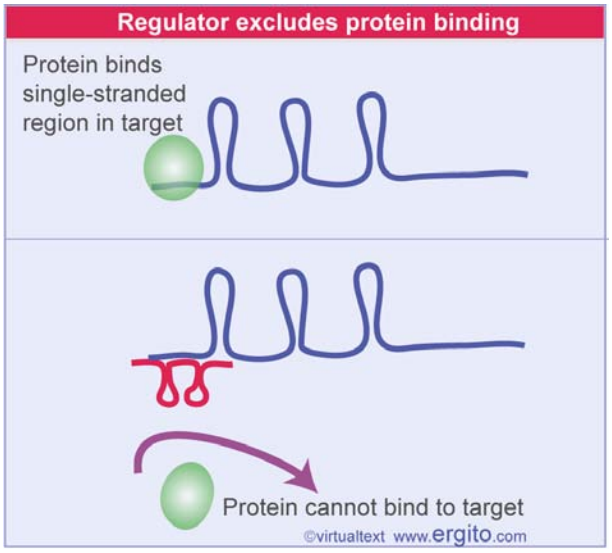

Figure 11.32 A protein that binds to a single-stranded region in a target RNA could be excluded by a regulator RNA that forms a duplex in this region.

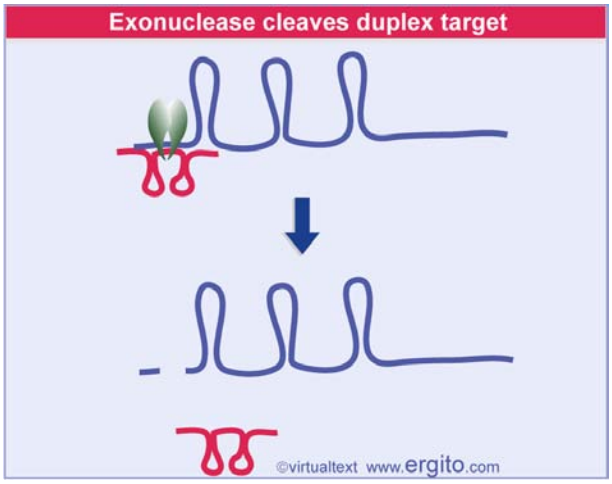

Figure 11.33 By binding to a target RNA to form a duplex region, a regulator RNA may create a site that is attacked by a nuclease.

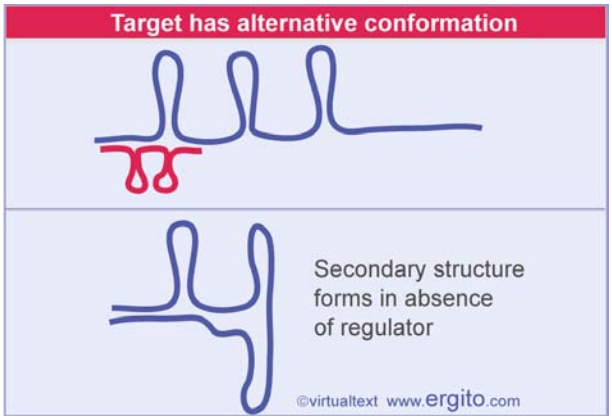

Figure 11.34 The secondary structure formed by base pairing between two regions of the target RNA may be prevented from forming by base pairing with a regulator RNA. In this example, the ability of the 3′ end of the RNA to pair with the 5′ end is prevented by the regulator.

The feature common to both types of RNA-mediated regulation is that changes in secondary structure of the target control its activity.

A difference between RNA regulators and the proteins that repress operons is that the RNA does not have allosteric properties; it cannot respond to other small molecules by changing its ability to recognize its target. It can be turned on by controlling transcription of its gene or it could be turned off by an enzyme that degrades the RNA regulator product.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.3.11.19

REGULATORY CIRCUITS

3.11.20 Bacteria contain regulator RNAs

Key Terms

An **sRNA** is a small bacterial RNA that functions as a regulator of gene expression.

Key Concepts

- Bacterial regulator RNAs are called sRNAs.
- Several of the sRNAs are bound by the protein Hfq, which increases their effectiveness.
- The OxyS sRNA activates or represses expression of >10 loci at the post-transcriptional level.

In bacteria, regulator RNAs are short molecules, collectively known as **sRNAs**; *E. coli* contains at least 17 different sRNAs (2945). Some of the sRNAs are general regulators that affect many target genes (for review see 3234). They function by base pairing with target RNAs (typically mRNAs) to control either their stability or function.

An example of stability control is provided by the small antisense regulator RyhB, which regulates 6 mRNAs coding for proteins concerned with iron storage in *E. coli*. It base pairs with each of the target mRNAs to form double-stranded regions that are substrates for RNAase E. An interesting feature of the circuit is that the ribonuclease destroys the regulator RNA as well as the mRNA (4521).

Oxidative stress provides an interesting example of a general control system in which RNA is the regulator. When exposed to reactive oxygen species, bacteria respond by inducing antioxidant defense genes. Hydrogen peroxide activates the transcription activator OxyR, which controls the expression of several inducible genes. One of these genes is *oxyS*, which codes for a small RNA.

Figure 11.35 shows two salient features of the control of *oxyS* expression. In a wild-type bacterium under normal conditions, it is not expressed. The pair of gels on the left side of the figure shows that it is expressed at high levels in a mutant bacterium with a constitutively active *oxyR* gene. This identifies *oxyS* as a target for activation by *oxyR*. The pair of gels on the right side of the figure show that OxyS RNA is transcribed within 1 minute of exposure to hydrogen peroxide (2944).

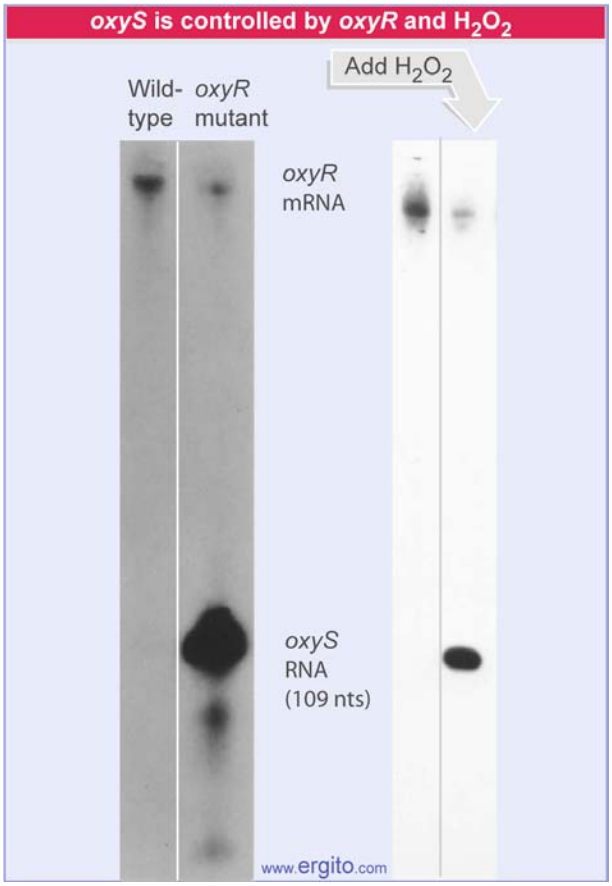

Figure 11.35 The gels on the left show that *oxyS* RNA is induced in an *oxyR* constitutive mutant. The gels on the right show that *oxyS* RNA is induced within 1 minute of adding hydrogen peroxide to a wild-type culture. Photograph kindly provided by Gisela Storz (see 2944).

The OxyS RNA is a short sequence (109 nucleotides) that does not code for protein. It is a *trans*-acting regulator that affects gene expression at post-transcriptional levels. It has >10 target loci; at some of them, it activates expression, at others it represses expression. **Figure 11.36**shows the mechanism of repression of one target, the FlhA mRNA. Three stem-loop structures protrude in the secondary structure of OxyR mRNA, and the loop close to the 3′ terminus is complementary to a sequence just preceding the initiation codon of FlhA mRNA. Base pairing between OxyS RNA and FlhA RNA prevents the ribosome from binding to the initiation codon, and therefore represses translation (2943). There is also a second pairing interaction that involves a sequence within the coding region of FlhA.

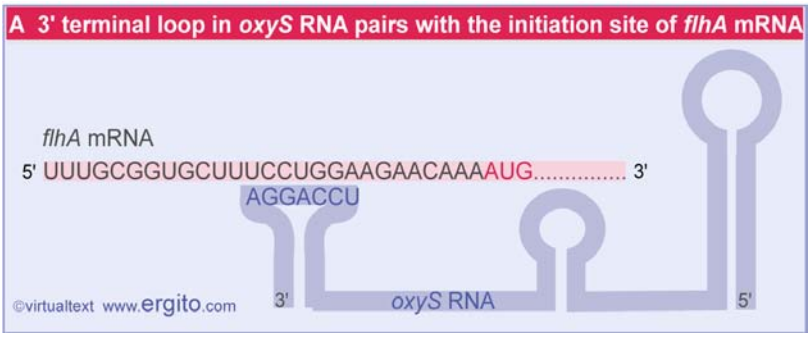

Figure 11.36 *oxyS* RNA inhibits translation of *flhA* mRNA by base pairing with a sequence just upstream of the AUG initiation codon.

Another target for *oxyS* is *rpoS*, the gene coding for an alternative sigma factor (which activates a general stress response). By inhibiting production of the sigma factor, *oxyS* ensures that the specific response to oxidative stress does not trigger the response that is appropriate for other stress conditions. The *rpoS* gene is also regulated by two other sRNAs (DsrA and RprA), which activate it. These three sRNAs appear to be global regulators that coordinate responses to various environmental conditions.

The actions of all three sRNAs are assisted by an RNA-binding protein called Hfq. The Hfq protein was originally identified as a bacterial host factor needed for replication of the RNA bacteriophage Q β. It is related to the Sm proteins of eukaryotes that bind to many of the snRNAs (small nuclear RNAs) that have regulatory roles in gene expression (see *Molecular Biology 5.24.5 snRNAs are required for splicing*) (2941; 2942). Mutations in its gene have many effects, identifying it as a pleiotropic protein. Hfq binds to many of the sRNAs of *E. coli*. It increases the effectiveness of OxyS RNA by enhancing its ability to bind to its target mRNAs. The effect of Hfq is probably mediated by causing a small change in the secondary structure of OxyS RNA that improves the exposure of the single-stranded sequences that pair with the target mRNAs.

Last updated on January 7, 2004

Reviews

3234. Gottesman, S. (2002). *Stealth regulation: biological circuits with small RNA switches.* Genes Dev. 16, 2829-2842.

References

2941. Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C., and Storz, G. (2002). *The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs.* Mol. Cell 9, 11-22.

2942. Moller, T., Franch, T., Hojrup, P., Keene, D. R., Bachinger, H. P., Brennan, R. G., and Valentin-Hansen, P. (2002). *Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction.* Mol. Cell 9, 23-30.

2943. Altuvia, S., Zhang, A., Argaman, L., Tiwari, A., and Storz, G. (1998). *The E. coli OxyS regulatory RNA represses fhlA translation by blocking ribosome binding.* EMBO J. 17, 6069-6075.

2944. Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L., and Storz, G. (1997). *A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator.* Cell 90, 43-53.

2945. Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001). *Identification of novel small RNAs using comparative genomics and microarrays.* Genes Dev. 15, 1637-1651.

4521. Massé, E., Escorcia, F. E., and Gottesman, S. (2003). *Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli.* Genes Dev. 17, 2374-2383.

REGULATORY CIRCUITS

3.11.21 MicroRNAs are regulators in many eukaryotes

Key Terms

MicroRNAs are very short RNAs that may regulate gene expression.

Key Concepts

- Animal and plant genomes code for many short (~22 base) RNA molecules, called microRNAs.
- MicroRNAs regulate gene expression by base pairing with complementary sequences in target mRNAs.

Very small RNAs are gene regulators in many eukaryotes. The first example was discovered in the nematode *C. elegans* as the result of the interaction between the regulator gene *lin4* and its target gene, *lin14*. **Figure 11.37** illustrates the behavior of this regulatory system. The *lin14* target gene regulates larval development. Expression of *lin14* is controlled by *lin4*, which codes for a small transcript of 22 nucleotides (2195; 2196). The *lin4* transcripts are complementary to a 10 base sequence that is repeated 7 times in the 3′ nontranslated region of *lin14*. Expression of *lin4* represses expression of *lin14* post-transcriptionally, most likely because the base pairing reaction between the two RNAs leads to degradation of the mRNA. This system is especially interesting in implicating the 3′ end as a site for regulation.

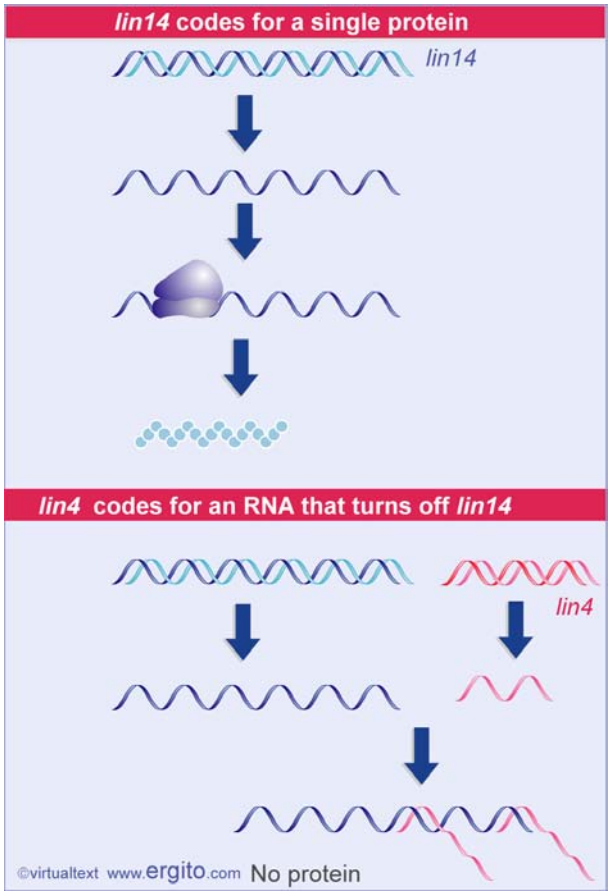

Figure 11.37 *lin4* RNA regulates expression of *lin14* by binding to the 3′ nontranslated region.

The *lin4* RNA is an example of a **microRNA** (miRNA). There are ~80 genes in the *C. elegans* genome coding for microRNAs of 21-24 nucleotide length (2193; 2194). They have varying patterns of expression during development and are likely to be regulators of gene expression. Many of the microRNAs of *C. elegans* are contained in a large (15S) ribonucleoprotein particle (2510).

Many of the *C. elegans* microRNAs have homologues in mammals, so the mechanism may be widespread. They are also found in plants. Of 16 microRNAs in *Arabidopsis*, 8 are completely conserved in rice, suggesting widespread conservation of this regulatory mechanism (3027).

The mechanism of production of the microRNAs is also widely conserved. In the example of *lin4*, the gene is transcribed into a transcript that forms a double-stranded region that becomes a target for a nuclease called Dicer. This has an N-terminal helicase activity, enabling it to unwind the double-stranded region, and two nuclease domains that are related to the bacterial ribonuclease III. Related enzymes are found in flies, worms, and plants (3028; 3029; 3027). Cleavage of the initial transcript generates the active microRNA. Interfering with the enzyme activity blocks the production of microRNAs and causes developmental defects.

Last updated on 10-16-2002

References

2193. Lau, N. C., Lim, l. e. E. P., Weinstein, E. G., and Bartel, d. a. V. P. (2001). *An abundant class of tiny RNAs with probable regulatory roles in C. elegans.* Science 294, 858-862.

2194. Lee, R. C. and Ambros, V. (2001). *An extensive class of small RNAs in C. eleganss.* Science 294, 862-864.

2195. Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell 75, 843-854.

2196. Wightman, B., Ha, I., and Ruvkun, G. (1993). *Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans.* Cell 75, 855-862.

2510. Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). *miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs.* Genes Dev. 16, 720-728.

3027. Reinhart,B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., and Bartel, D. P. (2002). *MicroRNAs in plants.* Genes Dev. 16, 1616-1626.

3028. Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). *Role for a bidentate ribonuclease in the initiation step of RNA interference.* Nature 409, 363-366.

3029. Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J., and Plasterk, R. H. (2001). *Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans.* Genes Dev. 15, 2654-2659.

REGULATORY CIRCUITS

3.11.22 RNA interference is related to gene silencing

Key Terms

RNA interference (RNAi) describes the technique in which double-strand RNA is introduced into cells to eliminate or reduce the activity of a target gene. It is caused by using sequences complementary to the double-stranded RNA sequences to trigger degradation of the mRNA of the gene.

RNA silencing describes the ability of a dsRNA to suppress expression of the corresponding gene systemically in a plant.

Cosuppression describes the ability of a transgene (usually in plants) to inhibit expression of the corresponding endogenous gene.

Key Concepts

- RNA interference triggers degradation of mRNAs complementary to either strand of a short dsRNA.
- dsRNA may cause silencing of host genes.

The regulation of mRNAs by microRNAs is mimicked by the phenomenon of **RNA interference** (RNAi). This was discovered when it was observed that antisense and sense RNAs can be equally effective in inhibiting gene expression (for review see 2077). The reason is that preparations of either type of (supposedly) single-stranded RNA are actually contaminated by small amounts of double-stranded RNA (1189).

Work with an *in vitro* system shows that the dsRNA is degraded by ATP-dependent cleavage to give oligonucleotides of 21-23 bases. The short RNA is sometimes called siRNA (short interfering RNA). **Figure 11.38** shows that the mechanism of cleavage involves making breaks relative to each 3′ end of a long dsRNA to generate siRNA fragments with short (2 base) protruding 3′ ends. The same enzyme (Dicer) that generates microRNAs is responsible for the cleavage.

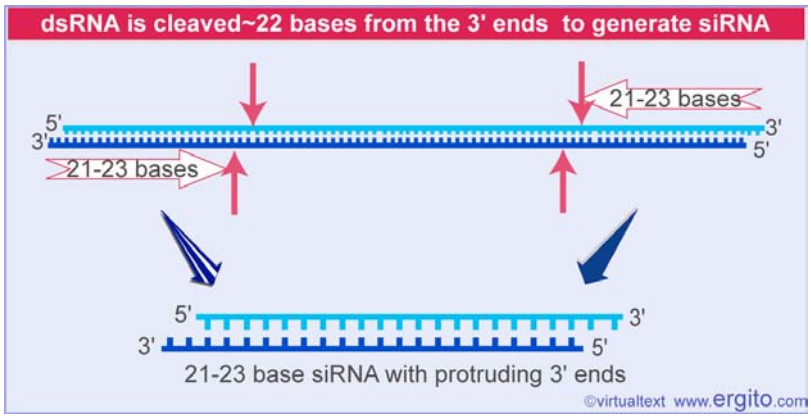

Figure 11.38 siRNA that mediates RNA interference is generated by cleaving dsRNA into smaller fragments. The cleavage reaction occurs 21-23 nucleotides from a 3′ end. The siRNA product has protruding bases on its 3′ ends.

RNAi occurs post-transcriptionally when an siRNA induces degradation of a complementary mRNA (1190; 1191). **Figure 11.39** suggests that the siRNA may provide a template that directs a nuclease to degrade mRNAs that are complementary to one or both strands, perhaps by a process in which the mRNA pairs with the fragments (1192). It is likely that a helicase is required to assist the pairing reaction. The siRNA directs cleavage of the mRNA in the middle of the paired segment. These reactions occur within a ribonucleoprotein complex called RISC (RNA-induced silencing complex) (for review see 2511).

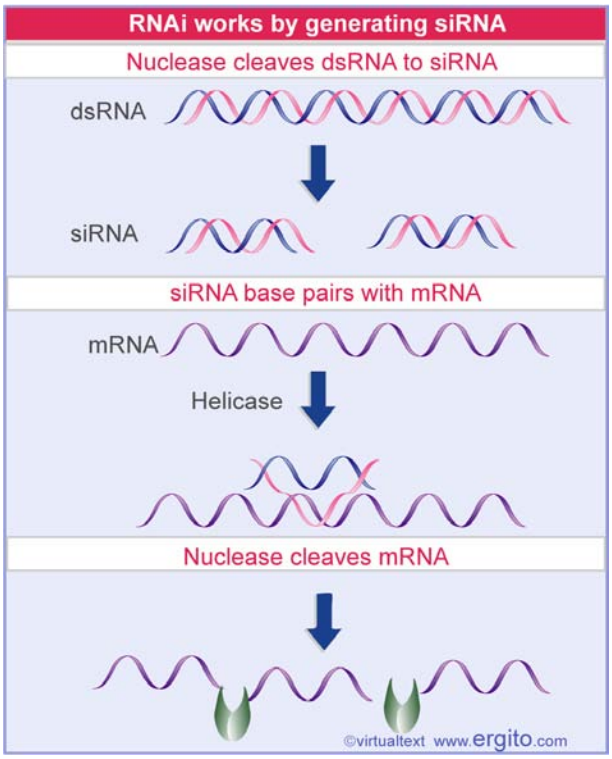

Figure 11.39 RNAi occurs when a dsRNA is cleaved into fragments that direct cleavage of the corresponding mRNA.

RNAi has become a powerful technique for ablating the expression of a specific target gene in invertebrate cells, especially in *C. elegans* and *D. melanogaster*. However, the technique has been limited in mammalian cells, which have a more generalized response to dsRNA of shutting down protein synthesis and degrading mRNA. **Figure 11.40** shows that this happens because of two reactions. The dsRNA activates the enzyme PKR, which inactivates the translation initiation factor eIF2a by phosphorylating it. And it activates 2′5′ oligoadenylate synthetase, whose product activates RNAase L, which degrades all mRNAs. However, it turns out that these reactions require dsRNA that is longer than 26 nucleotides. If shorter dsRNA (21-23 nucleotides) is introduced into mammalian cells, it triggers the specific degradation of complementary RNAs just as with the RNAi technique in worms and flies (1905). With this advance, it seems likely that RNAi will become the universal mechanism of choice for turning off the expression of a specific gene.

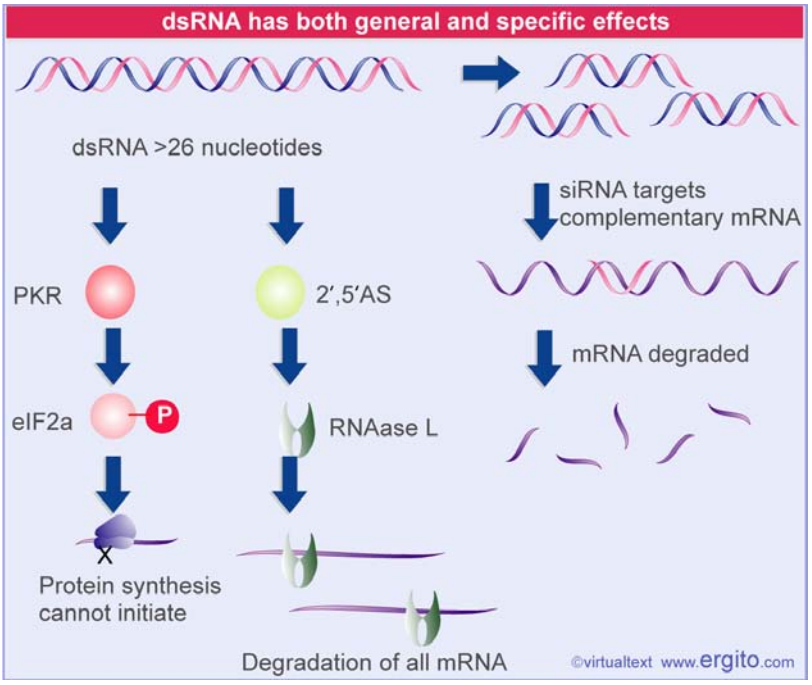

Figure 11.40 dsRNA inhibits protein synthesis and triggers degradation of all mRNA in mammalian cells as well as having sequence-specific effects.

As an example of the progress being made with the technique, it has been possible to use RNAi for a systematic analysis of gene expression in *C. elegans*. Loss of function phenotypes can be generated by feeding worms with bacteria expressing a dsRNA that is homologous to a target gene. By making a library of bacteria in which each bacterium expresses a dsRNA corresponding to a different gene, worms have been screened for the effects of knocking out most (86%) of the genes (3323).

RNA interference is related to natural processes in which gene expression is silenced. Plants and fungi show **RNA silencing** (sometimes called post-transcriptional gene silencing) in which dsRNA inhibits expression of a gene (for review see 3242). The most common source of the RNA is a replicating virus. This mechanism may have evolved as a defense against viral infection. When a virus infects a plant cell, the formation of dsRNA triggers the suppression of expression from the plant genome (1396). RNA silencing has the further remarkable feature that it is not limited to the

cell in which the viral infection occurs: it can spread throughout the plant systemically (1397). Presumably the propagation of the signal involves passage of RNA or fragments of RNA. It may require some of the same features that are involved in movement of the virus itself. It is possible that RNA silencing involves an amplification of the signal by an RNA-dependent RNA synthesis process in which a novel polymerase uses the siRNA as a primer to synthesize more RNA on a template of complementary RNA (for review see 2490).

A related process is the phenomenon of **cosuppression** in which introduction of a transgene causes the corresponding endogenous gene to be silenced. This has been largely characterized in plants (1193). The implication is that the transgene must make both antisense and sense RNA copies, and this inhibits expression of the endogenous gene.

Silencing takes place by RNA-RNA interactions. It is also possible that dsRNA may inhibit gene expression by interacting with the DNA. If a DNA copy of a viroid RNA sequence is inserted into a plant genome, it becomes methylated when the viroid RNA replicates (2149). This suggests that the RNA sequence could be inducing methylation of the DNA sequence. Similar targeting of methylation of DNA corresponding to sequences represented in dsRNA has been detected in plant cells (2150). Methylation of DNA is associated with repression of transcription, so this could be another means of silencing genes represented in dsRNA (see *Molecular Biology 5.21.18 Gene expression is associated with demethylation*). Nothing is known about the mechanism (for review see 2077; 2078).

Last updated on 10-16-2002

Reviews

2077. Sharp, P. A. (2001). *RNA interference--2001.* Genes Dev. 15, 485-490.

2078. Matzke, M., Matzke, A. J., and Kooter, J. M. (2001). *RNA: guiding gene silencing.* Science 293, 1080-1083.

2490. Ahlquist, P. (2002). *RNA-Dependent RNA Polymerases, Viruses, and RNA Silencing.* Science 296, 1270-1273.

2511. Schwartz, D. S. and Zamore, P. D. (2002). *Why do miRNAs live in the miRNP?* Genes Dev. 16, 1025-1031.

3242. Tijsterman, M., Ketting, R. F., and Plasterk, R. H. (2002). *The genetics of RNA silencing.* Annu. Rev. Genet. 36, 489-519.

References

1189. Fire, A.Xu, S.Montgomery, M. K.Kostas, S. A.Driver, and S. E.Mello, C. C. (1998). *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature 391, 806-811.

1190. Montgomery, M. K., Xu, S., and Fire, A. (1998). *RNA as a target of double-stranded RNA-mediated genetic interference in C. elegans.* Proc. Natl. Acad. Sci. USA 95, 15502-15507.

1191. Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). *Double-stranded RNA induces mRNA degradation in Trypanosoma brucei.* Proc. Natl. Acad. Sci. USA 95, 14687-14692.

1192. Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). *RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals.* Cell 101, 25-33.

1193. Waterhouse, P. M., Graham, M. W., and Wang, M. B. (1998). *Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA.* Proc. Natl. Acad. Sci. USA 95, 13959-13964.

1396. Hamilton, A. J. and Baulcombe, D. C. (1999). *A species of small antisense RNA in posttranscriptional gene silencing in plants*. Science 286, 950-952.

1397. Voinnet, O., Pinto, Y. M., and Baulcombe, D. C. (1999). *Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants.* Proc. Natl. Acad. Sci. USA 96, 14147-14152.

1905. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). *Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells.* Nature 411, 494-498.

2149. Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H. L. (1994). *RNA-directed de novo methylation of genomic sequences in plants.* Cell 76, 567-576.

2150. Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A., and Matzke, A. J. (2000). *Transcriptional silencing and promoter methylation triggered by double-stranded RNA.* EMBO J. 19, 5194-5201.

3323. Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P., and Ahringer, J. (2003). *Systematic functional analysis of the C. elegans genome using RNAi.* Nature 421, 231-237.

REGULATORY CIRCUITS

3.11.23 Summary

Gene expression can be regulated positively by factors that activate a gene or negatively by factors that repress a gene. The first and most common level of control is at the initiation of transcription, but termination of transcription may be also be controlled. Translation may be controlled by regulators that interact with mRNA. The regulatory products may be proteins, which often are controlled by allosteric interactions in response to the environment, or RNAs, which function by base pairing with the target RNA to change its secondary structure. Regulatory networks can be created by linking regulators so that the production or activity of one regulator is controlled by another.

Bacteria respond to the supply of glucose by repressing the production of the enzyme systems that catabolize alternative carbon sources. Inducer exclusion is a major component of the response, and works by inhibiting the uptake of the other sugars into the bacterium, with the result that the operons coding for their enzyme systems fail to be turned on. Increase in glucose levels also may lead to a reduction in the level of the small nucleotide cyclic AMP, although this is now controversial.

Some promoters cannot be recognized by RNA polymerase (or are recognized only poorly) unless a specific activator protein is present. Activator proteins also may be regulated by small molecules. The CRP activator becomes able to bind to target sequences in the presence of cyclic AMP. All promoters that respond to CRP have at least one copy of the target sequence. Binding of CRP to its target involves bending DNA. Direct contact between one subunit of CRP and RNA polymerase is required to activate transcription.

A common means for controlling translation is for a regulator protein to bind to a site on the mRNA that overlaps the ribosome binding site at the initiation codon. This prevents ribosomes from initiating translation. RegA of T4 is a general regulator that functions on several target mRNAs at the level of translation. Most proteins that repress translation possess this capacity in addition to other functional roles; in particular, translation is controlled in some cases of autogenous regulation, when a gene product regulates expression of the operon containing its own gene.

The level of protein synthesis itself provides an important coordinating signal. Deficiency in aminoacyl-tRNA causes an idling reaction on the ribosome, which leads to the synthesis of the unusual nucleotide ppGpp. This is an effector that inhibits initiation of transcription at certain promoters; it also has a general effect in inhibiting elongation on all templates.

Attenuation is a mechanism that relies on regulation of termination to control transcription through bacterial operons. It is commonly used in operons that code for enzymes involved in biosynthesis of an amino acid. The polycistronic mRNA of the operon starts with a sequence that can form alternative secondary structures. One of the structures has a hairpin loop that provides an intrinsic terminator upstream of the structural genes; the alternative structure lacks the hairpin. Various types of interaction can be used to determine whether the hairpin forms. One is for a protein

to bind to the mRNA to prevent formation of the alternative structure. In the *trp* operon of *B. subtilis*, the TRAP protein has this function; it is controlled by the anti-TRAP protein, whose production in turn is controlled by the level of uncharged aminoacyl-tRNA^{Trp}. In the *trp* operon of *E. coli*, the choice of which structure forms is controlled by the progress of translation through a short leader sequence that includes codons for the amino acid(s) that are the product of the system. In the presence of aminoacyl-tRNA bearing such amino acid(s), ribosomes translate the leader peptide, allowing a secondary structure to form that supports termination. In the absence of this aminoacyl-tRNA, the ribosome stalls, resulting in a new secondary structure in which the hairpin needed for termination cannot form. The supply of aminoacyl-tRNA therefore (inversely) controls amino acid biosynthesis.

Small regulator RNAs are found in both bacteria and eukaryotes. *E. coli* has ~17 sRNA species. The *oxyS* sRNA controls about 10 target loci at the post-transcriptional level; some of them are repressed, and others are activated. Repression is caused when the sRNA binds to a target mRNA to form a duplex region that includes the ribosome-binding site. MicroRNAs are ~22 bases long and are produced in many eukaryotes by cleavage of a longer transcript. They function by base pairing with target mRNAs to form duplex regions that are susceptible to cleavage by endonucleases. The degradation of the mRNA prevents its expression. The technique of RNA interference is becoming the method of choice for inactivating eukaryotic genes. It uses the introduction of short dsRNA sequences with one strand complementary to the target RNA, and it works by inducing degradation of the targets. This may be related to a natural defense system in plants called RNA silencing.