# Transcription Activation at Bacterial Promoters

**Stephen JW Busby,** *University of Birmingham, Birmingham, UK*

**Nigel J Savery,** *University of Bristol, Bristol, UK*

Bacterial genomes encode thousands of genes whose expression is regulated in response to changes in the environment. Much of this regulation takes place at the level of the initiation of transcription of specific genes. Bacteria use hundreds of different proteins to activate transcription by a variety of mechanisms.

Introductory article

Article Contents

- Regulation at Bacterial Promoters
- Activation by $\sigma$ Factor Recruitment
- Activators that Make Direct Contact with RNA Polymerase
- Activators that Do Not Make Direct Contact with RNA Polymerase
- Cooperation between Activators at Promoters
- Regulation of Regulators
- Overview

doi: 10.1002/9780470015902.a0000855.pub2

## Regulation at Bacterial Promoters

Bacterial genomes encode thousands of proteins that are expressed in widely differing quantities. Moreover, the expression of many proteins is increased or decreased according to the circumstances of the bacterium. Much of this regulation is effected at the level of the initiation of transcription. This is a sensible choice from the point of view of the cell's economy since regulation at this stage ensures that unnecessary ribonucleic acid (RNA) is not made. This article focuses on the molecular mechanisms responsible for the activation of transcription initiation at bacterial promoters in response to changes in the environment. **See also**: Bacterial Genomes; Bacterial Transcription Regulation

In bacteria, transcription is due to the activity of a single multisubunit deoxyribonucleic acid (DNA)-dependent RNA polymerase. The form of RNA polymerase that transcribes genes is known as the core enzyme, and has a subunit composition of $\alpha_2\beta\beta'\omega$, where $\alpha$, $\beta$, $\beta'$ and $\omega$ denote different subunits. The core enzyme adopts a 'crab-claw' structure that it shares with eukaryotic RNA polymerases. The pathway of the template DNA through the crab-claw and the position of the elongating RNA chain in the enzyme's active site are understood. Although competent to make RNA, the core enzyme is unable to initiate transcription, and requires a $\sigma$ subunit to recognize promoters and to set up the initiation complex in which the strands of the template DNA are locally separated and the transcript start site is suitably placed in the RNA polymerase active site. Most $\sigma$ factors consist of 4 domains. Domains 2–4 carry determinants for the recognition of promoter DNA, while the role of Domain 1 is to mask these determinants in $\sigma$ molecules that are not attached to the core enzyme. The ensemble of core enzyme with a $\sigma$ factor is known as the RNA polymerase holoenzyme. Most bacteria contain a predominant $\sigma$ factor that is competent for recognition of promoters controlling genes needed for rapid growth. This is often referred to as the housekeeping $\sigma$ factor. **See also**: RNA Polymerases: Subunits and Functional Domains; Sigma Factors in Gene Expression

In the RNA polymerase holoenzyme, the domains of $\sigma$ are spread along the outer surface of the crab-claw structure and provide the principal determinants for RNA polymerase to recognize promoters (**Figure 1**). Thus, when holoenzyme docks to a promoter, $\sigma$ Domain 2 recognizes the promoter $-10$ element, Domain 3 recognizes specific base sequences immediately upstream (the extended $-10$ element) and Domain 4 recognizes the promoter $-35$ element. At many promoters, a further element, the UP element, that is located upstream of the promoter $-35$ region, contributes to holoenzyme binding via an interaction with the *C*-terminal domains of the two RNA polymerase $\alpha$ subunits. The initial holoenzyme–promoter complex is known as the 'closed' complex. After docking to a promoter, holoenzyme containing the housekeeping $\sigma$ (or an alternative $\sigma$ factor) undergoes a series of isomerizations that result in formation of a complex that is competent for transcript initiation (known as the 'open' complex).

It is thought that the amount of RNA polymerase and housekeeping $\sigma$ in bacteria is limited. For example, *Escherichia coli* cells contain $\sim$4000 molecules of RNA polymerase. In growing cells, much of this enzyme is involved in the transcription of a small number of essential genes, including genes encoding the cell's protein-synthesizing machinery. These genes have exceptionally strong promoters that can be used by RNA polymerase holoenzyme without the intervention of any other factor. Hence, only limited amounts of RNA polymerase are available for expression of the majority of the 4300 genes. Many of these genes encode products that are needed for growth or survival in particular conditions, and thus, bacteria have evolved mechanisms to upregulate their transcription when these conditions arise. To do this, RNA polymerase must be 'recruited' by the promoters of these genes in response to stimuli. This article describes how this occurs, with a particular focus on *E. coli* where transcriptional regulation has been most studied. It is likely that the lessons learned from *E. coli* apply to other bacteria, though this is still debatable. Many of the mechanisms that we know

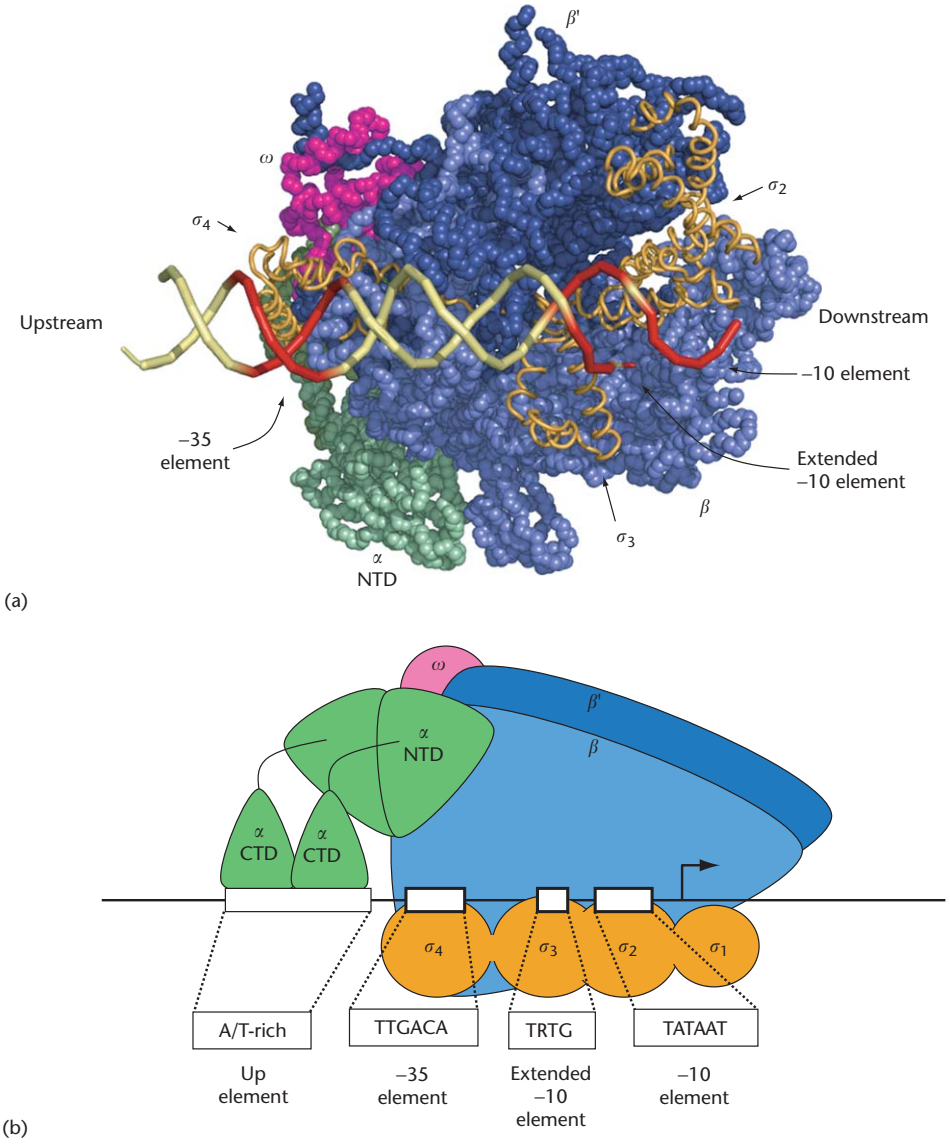



**Figure 1** RNA polymerase–promoter interactions. (a) Structure of *Thermus aquaticus* RNA polymerase holoenzyme docked to promoter DNA, from studies by (Seth Darst) and co-workers. The locations of domains 2, 3 and 4 of the $\sigma$ subunit (shown as orange ribbons) and their interactions with the −10, extended −10 and −35 elements (highlighted in red) are shown. Note that $\alpha$CTD and $\sigma$ domain 1 are not seen in the *T. aquaticus* RNA polymerase structure. (b) Scheme that illustrates the subunit structure of RNA polymerase holoenzyme and the disposition of the different subunits that make contact with a bacterial promoter. The consensus sequences of the conserved promoter DNA elements are also shown. The $\sigma$ subunit is comprised of 4 domains, and recognizes the −10, extended −10 and −35 elements. Each $\alpha$ subunit consists of an *N*-terminal domain ($\alpha$NTD) and a *C*-terminal domain ($\alpha$CTD) joined by a linker. The two $\alpha$CTDs bind to A/T-rich UP-elements at many promoters. The $\beta$ and $\beta'$ subunits comprise the active site of the enzyme. The $\omega$ subunit is thought to act as a chaperone to help the $\beta'$ subunit to fold correctly. The horizontal arrow indicates the direction of transcription and its startpoint.

about involve proteins that interact with RNA polymerase to bias its choice towards particular promoters. In most cases, either $\sigma$ factors or transcription factors are involved. The *E. coli* genome encodes seven $\sigma$ factors and nearly 250 transcription factors. The large number of genes involved in transcriptional regulation is a reflection of the prudence with which *E. coli* regulates the expression of its genes.

## Activation by $\sigma$ Factor Recruitment

As explained above, most bacteria contain one major $\sigma$ factor that confers promoter specificity on RNA polymerase holoenzyme and, in rapidly growing cells, the activity of most of the alternative $\sigma$ factors is minimal. The activity of many of these alternative factors is controlled by

an external signal (e.g. a stress such as heat or a particular chemical trigger). When the activity of a particular $\sigma$ factor is increased, a subpopulation of RNA polymerase core enzyme molecules is captured and the resulting holoenzyme becomes competent to bind promoters carrying elements that are recognized by Domains 2–4 of that particular factor. The use of alternative $\sigma$ factors is an important and ubiquitous mechanism used by bacteria to activate batteries of genes in response to external stimuli. Some $\sigma$ factors are global regulators, serving scores of promoters, while others recognize just one or two.

Different bacterial genomes encode different numbers of $\sigma$ factors that play different roles. For example, *Streptomyces coelicolor* contains over 60 such factors that control adaptation to many different environments, while in *Bacillus subtilis*, the sequential expression of alternative $\sigma$ factors drives cellular differentiation and spore formation. As a general rule, the number of $\sigma$ factors encoded in a bacterial genome is a reflection of the complexity of the environments in which it can grow. The mechanisms by which the activities of alternative $\sigma$ factors are controlled are complex. In some cases, their concentration in the cell is tightly regulated, while, in other instances, activity is modulated by the binding of anti-$\sigma$ factors that themselves can be regulated by anti-anti-$\sigma$ factors (see the section on Regulation of Regulators). **See also**: Sigma Factors in Gene Expression

## Activators that Make Direct Contact with RNA Polymerase

Many bacterial promoters are unable to recruit RNA polymerase in the face of competition from other promoters. Transcription activators remedy this problem by binding at target promoters, thereby recruiting RNA polymerase. In most cases, this recruitment is due to a direct contact between the activator and RNA polymerase holoenzyme.

Transcription activators are nearly all DNA-binding proteins with base sequence specificity determined by a DNA-binding module. Some activators are global controllers of transcription and interact at scores of promoters. For example, the *E. coli* cyclic adenosine monophosphate (AMP) receptor protein, CRP, interacts at nearly 200 different promoters and regulates their expression in response to certain stress conditions. Other activators interact at just one or two targets. In all cases, their activity must be regulated, and, as with $\sigma$ factors, this regulation can be complex (see the section on Regulation of Regulators).

Bacterial transcription activators that function by directly recruiting RNA polymerase holoenzyme to target promoters must contain a surface that interacts with RNA polymerase, as well as a DNA-binding determinant. Such surfaces are known as activating regions. In principle, an activator could contact any surface of RNA polymerase. After the activator–RNA polymerase contact has formed, RNA polymerase isomerizes to an open complex just as it would at an activator-independent promoter.

### Class I activation

A large set of bacterial activators bind upstream of position $-50$ at promoters such that they make contact with the *C*-terminal domain of one of the RNA polymerase $\alpha$ subunits. The contact recruits the $\alpha$ subunit *C*-terminal domain and hence the rest of the RNA polymerase to the DNA. Such activators are referred to as Class I activators (**Figure 2a**). The best-understood example is *E. coli* CRP at the lactose operon promoter. CRP is a dimer of two identical subunits that binds to a 22 base pair DNA site, centred between base pairs $-61$ and $-62$ upstream of the lactose promoter transcription startpoint (i.e. position $-61.5$). Each CRP subunit contains a DNA-binding helix-turn-helix and an activating region, formed from several side-chains of a small beta-turn, located just adjacent to the helix-turn-helix. This activating region (known as Activating Region 1) in the downstream subunit of CRP contacts a surface-exposed determinant (known as the 287 determinant) in one of the RNA polymerase $\alpha$ subunit *C*-terminal domains (**Figure 3a, b**).

Class I activators can activate transcription from many different locations upstream of target promoters, and there is great variety in the architecture of promoters that are subject to Class I activation. The reason for this is that the $\alpha$ subunit C-terminal domain is connected to the $\alpha$ subunit *N*-terminal domain by a very flexible linker and thus, RNA polymerase can be productively guided to a promoter by recruitment of the $\alpha$ subunit *C*-terminal domain at many different locations. Additionally, distortions in upstream DNA (e.g. by bending) can contribute to the flexibility in the possible location of binding sites for Class I activators.

### Class II activation

Another big set of activators, known as Class II activators, bind to sites that overlap the $-35$ element at target promoters. These activators can, potentially, contact several different parts of RNA polymerase holoenzyme. One consequence of the binding of an activator at this position is that the *C*-terminal domains of the RNA polymerase $\alpha$ subunits are displaced and, in some cases, bind to promoter DNA immediately upstream of the activator. Many such activators make contact with Domain 4 of the RNA polymerase $\sigma$ subunit (**Figure 2b**). The best characterized example is the bacteriophage $\lambda$ cI protein at the bacteriophage $\lambda$ $P_{RM}$ where a negatively charged activating region in cI protein contacts positively charged residues

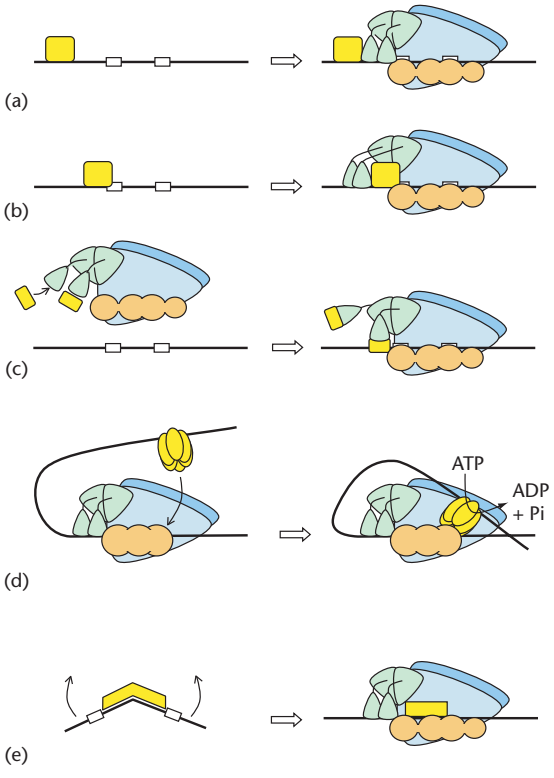



**Figure 2** Mechanisms of transcription activation. (a) Class I activation. Class I activators bind upstream of their target promoters and recruit RNA polymerase by contacting the *C*-terminal domain of one of its $\alpha$ subunits. (b) Class II activation. Class II activators bind close to the −35 element of target promoters and activate transcription by contacting domain 4 of the $\sigma$ subunit. (c) Activation by pre-recruitment. Some activators bind directly to RNA polymerase before it binds to promoter DNA. Such activators can alter the DNA-binding specificity of RNA polymerase, for example by masking the DNA-binding surface of the $\alpha$ subunit and replacing it with the DNA-binding domain of the activator. (d) Activation at $\sigma^{54}$-dependent promoters. RNA polymerase containing $\sigma^{54}$ can form a closed complex at a target promoter but is unable to form an open complex. Activators that function with this form of RNA polymerase bind to the DNA well upstream of the promoter, and, in order for them to contact RNA polymerase, a loop must form in the intervening DNA. The activators are ATPases, and direct contacts between the activator and $\sigma^{54}$ drive a series of ATP-driven conformational changes that promote open-complex formation. (e) Activation by restructuring DNA. If the −10 and −35 elements are not properly aligned, RNA polymerase cannot initiate transcription. Transcription is stimulated by activators that bind between the −10 and −35 elements and alter the shape of the intervening DNA so that correct alignment is restored.

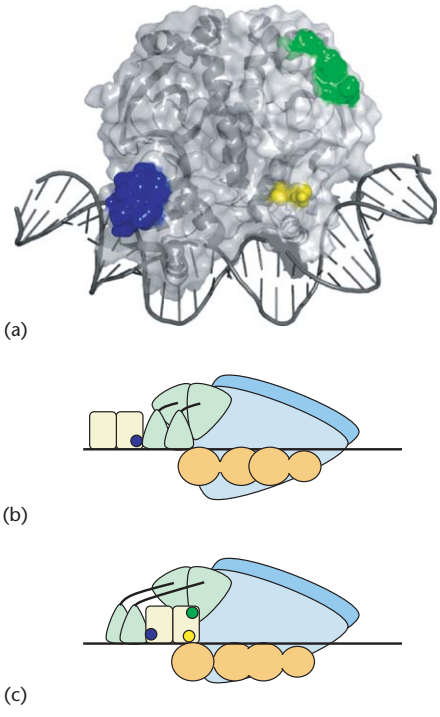



**Figure 3** Activation by CRP. (a) Structure of *E. coli* CRP bound to its target DNA, from studies by (Tom Steitz) and co-workers. The figure shows a CRP dimer, with the three activating regions, that can contact RNA polymerase, highlighted. In this view, AR1 (shown in blue) is visible on the left-hand monomer, and AR2 (shown in green) and AR3 (shown in yellow) are visible on the right-hand monomer. (b) Activation by CRP at a Class I promoter. The CRP-binding site is located upstream of the RNA polymerase-binding site. AR1 of the downstream CRP subunit interacts with the *C*-terminal domain of one of the RNA polymerase $\alpha$ subunits and so recruits RNA polymerase to the promoter. (c) Activation by CRP at a Class II promoter. The CRP-binding site is located at position −41.5. AR1 of the upstream CRP subunit contacts the *C*-terminal domain of one of the RNA polymerase $\alpha$ subunits, AR2 of the downstream subunit contacts the *N*-terminal domain of one of the RNA polymerase $\alpha$ subunits, and AR3 of the downstream subunit contacts domain 4 of the $\sigma$ subunit.

that are adjacent to the DNA-binding surface of $\sigma$ subunit Domain 4.

Another well-studied example is *E. coli* CRP at the galactose operon promoter, where it binds to a DNA site centred at position −41.5 upstream of the transcription startpoint, and the *C*-terminal domains of the RNA polymerase $\alpha$ subunits bind upstream. Activating Region 1 of the upstream subunit of bound CRP interacts with one of the $\alpha$ *C*-terminal domains. Two other surfaces in the downstream subunit of CRP, known as Activating Region 2 and Activating Region 3, contact determinants in the *N*-terminal domain of the RNA polymerase $\alpha$ subunit and Domain 4 of the $\sigma$ subunit, respectively (**Figure 3a, c**). Note that CRP, like many other activators, is 'ambidextrous', containing more than one activating region that can, apparently simultaneously, contact different targets in RNA polymerase. Kinetic studies have shown that, in such cases, the different activator–RNA polymerase interactions are likely to affect different steps in the transcription initiation process. Note too that *E. coli* CRP, like many global regulators, does not use the same mechanism at all of its target promoters and it brings different combinations of its activating regions into play at different promoters. Hence, while it functions as a Class I activator at the lactose promoter, it becomes a Class II activator at the galactose promoter.

## Activators that bind to free RNA polymerase holoenzyme

The textbook view that all bacterial transcription activators function by first binding at promoters and then contacting RNA polymerase has recently been challenged by the discovery that the *E. coli* MarA and SoxS proteins contact RNA polymerase holoenzyme before it binds to target promoters. MarA and SoxS are single-domain DNA-binding proteins and, when complexed with RNA polymerase, add a DNA–binding determinant to the holoenzyme. MarA and SoxS control the expression of genes needed for *E. coli* to manage certain chemical stresses and their expression is induced by these stresses. Thus, in response to stress, the levels of MarA or SoxS rise and this results in a subpopulation of RNA polymerase holoenzyme being captured and binding to promoters carrying cognate target sequences. This mechanism, dubbed pre-recruitment (**Figure 2c**), is similar to that used by alternative $\sigma$ factors. Remarkably, MarA and SoxS appear to bind to RNA polymerase by contacting the DNA-binding surface of the $\alpha$ subunit *C*-terminal domains. Thus, in times of stress, the cell appears to replace the $\alpha$ subunit DNA-binding determinant with the MarA/SoxS determinant. This theme recurs with the *B. subtilis* Spx protein, which binds to the *C*-terminal domains of the RNA polymerase $\alpha$ subunits and causes activation at certain promoters and inhibition at others. Scores of different proteins have now been shown to interact directly with free bacterial RNA polymerases, and regulation by these protein ligands is likely to be more significant than currently suggested by textbooks.

## A second paradigm for transcription activation: $\sigma^{54}$-dependent promoters

Nearly all $\sigma$ factors have a similar domain structure to *E. coli* $\sigma^{70}$, and their presence in holoenzyme orchestrates a series of isomerizations that convert the 'closed' RNA polymerase–promoter complex into a transcriptionally competent complex that can initiate RNA synthesis. Because these isomerizations can take place without the intervention of any other factor, the key determinant of the activity of any promoter is the recruitment of RNA polymerase, and it is for this reason that the activation mechanisms described above mostly intervene at the level of RNA polymerase recruitment. However, a small number of $\sigma$ factors do not resemble *E. coli* $\sigma^{70}$ and belong to a separate family, the $\sigma^{54}$ family. Unlike RNA polymerase carrying $\sigma^{70}$ and its homologues, RNA polymerase carrying $\sigma^{54}$ and its homologues is unable to isomerize from the closed polymerase–promoter complex to the transcriptionally competent open complex. Hence, RNA polymerase containing $\sigma^{54}$ recognizes target promoters but remains stalled at the closed complex. Progression depends upon binding of a special type of transcription factor, that binds upstream and then makes a direct interaction with $\sigma^{54}$, which pushes the stalled RNA polymerase into an open complex in a process driven by adenosine triphosphate (ATP) hydrolysis. These transcription factors all contain an ATP motor domain (known as the AAA domain) that couples the hydrolysis of ATP to the movement of surface-exposed loops, which interact with $\sigma^{54}$ and somehow drive the formation of a transcriptionally competent complex. Most of these factors also contain a regulatory domain and a DNA-binding domain and they are functional as multimers, often binding far upstream of the target promoter.

This mechanism is a second paradigm for transcription activation (**Figure 2d**) since the transcription factor acts on RNA polymerase that is already bound at targets rather than on the recruitment of RNA polymerase to those targets. Note that, at least in *E. coli*, the level of the $\sigma^{54}$ protein does not vary significantly when conditions change, and hence, regulation is due to modulation of the activity of the AAA domain-containing transcription factor.

# Activators that Do Not Make Direct Contact with RNA Polymerase

At some promoters, the proteins responsible for transcription activation make no direct interaction with RNA polymerase. Sometimes, activation is indirect and is due to the relief of repression. For example, promoters whose activity is repressed by local DNA folding may be activated if binding of a particular protein disrupts this structure. Such repression can be caused by different bacterial nucleoid-associated proteins such as HNS (histone for nucleoid structuring), IHF (integration host factor) or FIS (factor for inversion stimulation). This type of indirect activation is very common since bacterial DNA has to be highly compacted *in vivo*.

Promoters where a transcription factor activates transcription directly, without contacting RNA polymerase, are rare. Activators at these promoters function by altering promoter DNA conformation. Proteins of the MerR family provide the paradigm for this type of activation. Most of these bind to sites located between the $-35$ and $-10$ elements at target promoters (**Figure 2e**). Ligand binding then triggers a conformation change that results in distortion of the target DNA, aligning the $-35$ and $-10$ elements so that RNA polymerase holoenzyme can initiate transcription. MerR protein itself is encoded by the mercury resistance locus of transposon Tn501. When triggered by mercuric ions, it activates the transcription of genes encoding determinants for resistance to the toxic effects of mercury.

At other promoters where transcription activators function solely by altering DNA conformation, the activator binds well upstream of target promoter elements and induces a sharp bend at its target. Two mechanisms for activation have been suggested: either the activator-induced

DNA bend permits upstream sequences to wrap around the back of RNA polymerase, making interactions that promote transcription initiation, or activator binding induces conformational changes in the downstream promoter sequences that facilitate initiation. An especially interesting case is the *E. coli ilvG* promoter, where upstream-bound IHF protein relays a conformational change the promoter −10 element. Another intriguing case is the *E. coli tyrT* promoter that is activated by FIS protein that binds to three upstream sites and induces sharp bends at these sites. Although direct FIS–RNA polymerase interactions account for some of the activation, it has been suggested that the bound FIS molecules stabilize a DNA microdomain that facilitates several steps in the transcription-initiation process.

## Cooperation between Activators at Promoters

The activity of many bacterial promoters is co-dependent on the function of two or more activators, with co-dependence coupling promoter activity to different physiological signals. To date, five models have been found to account for such co-dependence.

First, the binding of one activator may be dependent on the binding of another and vice versa (**Figure 4a**). Surprisingly, this very simple mechanism is rare at bacterial promoters (though it is common at eukaryotic promoters).

Second, the binding of one activator may trigger the repositioning of the other, moving it from a location where it is unable to activate transcription, to a location where it is able to activate transcription (**Figure 4b**). This is the case for CRP and the MalT protein at the *E. coli malK* promoter. CRP triggers the repositioning of MalT to a location where it can interact with RNA polymerase (likely with Domain 4 of $\sigma^{70}$).

Third, the binding of one activator to the promoter may alter the trajectory of promoter DNA such that the other activator is able to make contact with RNA polymerase (**Figure 4c**). This is the case for IHF at a number of promoters served by RNA polymerase holoenzyme containing $\sigma^{54}$. IHF binds to the DNA between RNA polymerase that is stalled at the promoter and the upstream-bound AAA domain-containing activator, and induces a bend that facilitates interaction between the activator and $\sigma^{54}$. Similarly, at the *E. coli narG* promoter, IHF permits NarL, that is bound around 200 base pairs upstream, to interact with other components of the transcription-initiation complex.

Fourth, at some promoters, two or more contacts are needed to provide sufficient determinants for RNA polymerase to be recruited, and these contacts are provided by two or more transcription factors. This mechanism does not require any interaction between the different transcription factors and, for this reason, is found frequently at

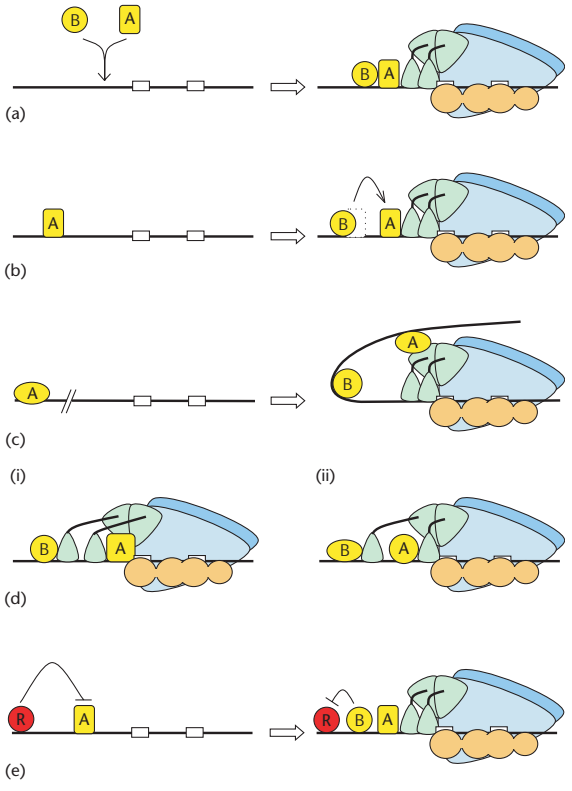
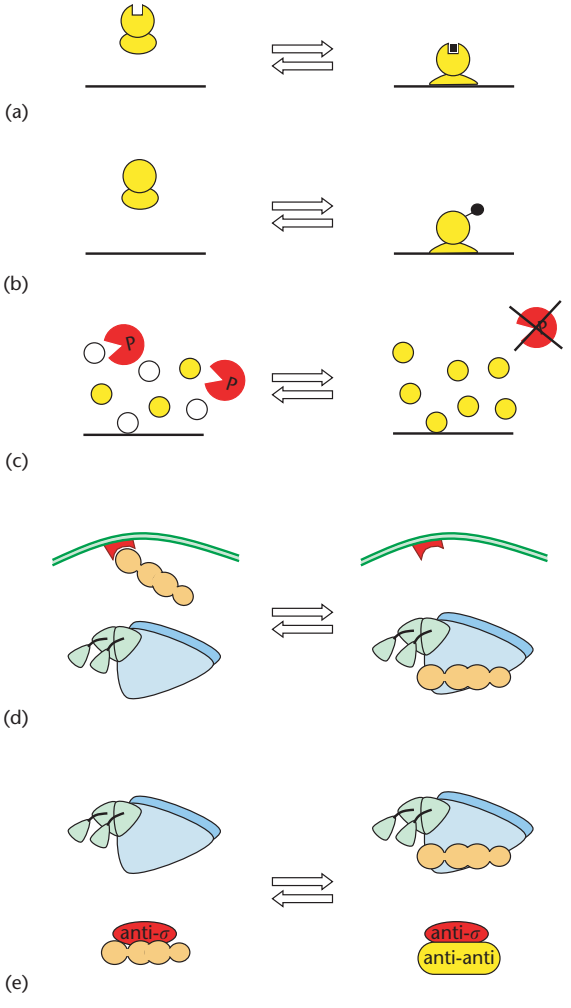



**Figure 4** Co-dependence mechanisms. Transcription initiation at many promoters requires two or more activators. The figure illustrates different mechanisms by which two activators (A and B) can cooperate to regulate a single promoter. (a) The binding of each activator to its target site in the DNA is dependent on the binding of the other. Only when both activators are present will a stable activator:DNA complex be formed. (b) The binding of one activator (B) may trigger the repositioning of the other (A), moving it from a location where it is unable to activate transcription, to a location where it is able to activate transcription. (c) The binding of one activator (B) to the promoter may alter the trajectory of promoter DNA such that the second activator (A) is able to make contact with RNA polymerase. (d) Two or more determinants are needed to provide sufficient contacts for RNA polymerase to be recruited, and these are provided by two or more transcription factors. e.g. a Class I activator can act in concert with a Class II activator (i), or with a second Class I activator (ii). (e) The second transcription activator (B) may counteract the action of a repressor (R) that is interfering with the function of the first activator (A).

complex promoters. It has the potential to generate many different promoter architectures since RNA polymerase has many potential sites that can be targets for activators (**Figure 4d**). In one arrangement, the first activator binds to a site that overlaps the promoter −35 element and contacts Domain 4 of the holoenzyme $\sigma$ subunit or other targets in RNA polymerase. However, these contacts are insufficient and recruitment of RNA polymerase requires the second activator that binds upstream and contacts one or both of the $\alpha$ subunit *C*-terminal domains. Such promoters, which are quite common in *E. coli*, are, thus, regulated by a Class I and a Class II activator functioning cooperatively. Note that, due to the flexible linker between the *N*- and

*C*-terminal domains of the α subunit, the upstream Class I activator can function from many locations. Note also that promoters that are dependent on two or more contacts with activators can be served by two molecules of a factor such as CRP, that can function as both a Class I and a Class II activator. Another possible arrangement for promoters dependent on two independent contacts with activators arises because the two RNA polymerase α subunit *C*-terminal domains can be recruited independently to a promoter. Hence, some co-dependent promoters are activated by two independently binding Class I transcription activators that each contact one of the two α *C*-terminal domains (**Figure 4d**).

Finally, the role of the second transcription activator may be to counteract the action of a repressor that is interfering with the function of the first activator (**Figure 4e**). Thus, in this case, there is only one true activator, with the second activator functioning as an antirepressor. An example is the *E. coli nir* promoter that is dependent on FNR binding to a target that overlaps the promoter −35 region. Activation by FNR is repressed by upstream-bound IHF and FIS, but this repression is counteracted by the binding of NarL or NarP to a DNA site located at position −69.5 upstream of the transcription start.




## Regulation of Regulators

The activities of most transcription factors and sigma factors are highly regulated and this permits the control of gene expression in response to changes in the environment. In addition, when bacteria undergo programmed morphogenic changes, e.g. during the sporulation of *Bacillus subtilis*, transcription factors and/or sigma factors are turned on and off in a defined sequence.

Many different mechanisms are used to control the activities of transcription factors. One commonly found mechanism is regulation by small molecules, in which the presence of a ligand alters the conformation of a transcription factor from an inactive to an active form or vice versa (**Figure 5a**). In many cases, gene-regulatory proteins contain separate ligand- and DNA-binding domains and the binding of the ligand toggles the factor between

**Figure 5** Regulation of transcription factors. For transcription activation to act as a 'genetic switch', the activity of each transcription activator must be controlled. The figure illustrates different mechanisms used to ensure that activators switch on genes in response to the correct signal. (a) Regulation by binding of a small ligand. Some activators are toggled between an active and inactive form by binding a small signal molecule, such as cyclic adenosine monophosphate (cAMP). Typically, the small molecule causes a conformational change that affects an essential function such as DNA-binding specificity. In this example, interaction with a small ligand converts the activator from an inactive form that is unable to bind to its target sequence to an active, sequence-specific DNA-binding form. (b) Regulation by covalent modification. Some activators are converted between active and inactive forms by covalent modification (e.g. phosphorylation by a kinase that senses a particular signal). Typically, the covalent modification affects an essential function such as DNA-binding specificity. In this example, the activator is converted from an inactive form, that is unable to bind to its target sequence, to an active, sequence-specific DNA-binding form by phosphorylation of its receiver module. (c) Regulation by intracellular protein concentration. The ability of some activators to regulate gene expression is controlled by their concentration. In this example, the activator is produced constitutively and its steady state concentration is kept low because the activator is rapidly degraded by proteases (P). If the proteases are prevented from degrading the activator its concentration in the cell rapidly rises. (d) Regulation by restriction of cellular location. The activity of some regulatory proteins is controlled by proteins that sequester them to locations from which they cannot activate transcription. In this example, $\sigma^{28}$ is prevented from binding to RNA polymerase by a complex of proteins that hold it at the cell membrane. In response to an appropriate signal, the complex releases $\sigma^{28}$, allowing it to activate transcription. (e) Regulation by anti-σ factors. Anti-σ factors regulate σ factor activity by forming a complex that prevents the σ factor binding to RNA polymerase. In response to an appropriate signal, the anti-σ factor releases the σ factor, allowing it to activate transcription of its target genes. The mechanism by which the σ:anti-σ complex is disrupted often involves a third protein, termed an anti-anti-σ, which binds to the anti-σ factor more tightly than the σ factor. This mechanism is often referred to as 'partner switching'.

DNA- and non-DNA-binding conformations. The best example of this is the *E. coli lac* operon repressor that cannot bind strongly to its DNA operator sequence when it interacts with allolactose. Recent studies of how cohorts of bacteria cooperate together have highlighted the importance of diffusible quorum-sensing molecules (e.g. homoserine lactones) that permit communication between bacteria in a population. These molecules trigger gene-regulatory proteins to activate transcription initiation when the ligand concentration reaches a certain threshold.

The activity of other gene regulatory proteins is controlled by covalent modification (**Figure 5b**). For example, members of the response-regulator family of transcription factors carry a conserved receiver module, that can be phosphorylated, coupled to an output module that usually functions as a DNA-binding transcription factor. Phosphorylation is usually mediated by transmembrane-located sensor-kinases that are composed of a sensing module, a transmembrane segment and a kinase module, located on the cytoplasmic face of the bacterial inner membrane. Pairs of sensor kinases and response regulators have evolved so that external signals can activate gene expression. Hence, in many cases, the sensor module is located on the outside face of the bacterial membrane. In response to an external signal, it activates the kinase module that then covalently modifies the receiver domain of the cognate response regulator. This causes a conformational change in the response regulator to toggle it into its active form.

Although the activities of many bacterial transcription factors are controlled by ligand binding in response to signals inside the cell, or by covalent modification in response to signals outside the cell, for many factors, activity is fixed by their level (**Figure 5c**). This can be set either by the transcription of their own gene or by their turnover. Hence, some factors are made constitutively but are continually being degraded and full activity is revealed only when the cycle of degradation is interrupted. Many such transcription factors have a simple domain structure since they lack a regulatory module that binds a ligand or is a target for covalent modification. An interesting variation of this is found with factors that are actively sequestered at the cell membrane and thus rendered inactive (**Figure 5d**). Such is the case for the *E. coli* $\sigma^{28}$ protein, that is essential for the expression of many genes needed for survival in the face of stresses such as high temperature. A complex set of inner membrane proteins ensures that $\sigma^{28}$ is sequestered and hence inactive, but is released in response to stresses sensed outside the cell. This provides another mechanism for gene expression to be coupled to events outside of the bacterium. A further variation of the sequestration theme is found with many other $\sigma$ factors whose activity is regulated by binding to an anti-$\sigma$ factor. In these cases, the activity of the $\sigma$ factor is determined by the activity of the anti-$\sigma$ factor, which itself may be regulated by binding to another partner (**Figure 5e**). Combinations of the mechanisms described in **Figure 5** can generate cascades of transcription factors, often with global regulators at the head of the cascade. These cascades often incorporate complex feedback loops and are responsible for most processes of bacterial differentiation and global adaptation.

## Overview

At first sight, gene regulation in bacteria appears simple compared to the situation in eukaryotes. However, we now know that bacterial promoters are extremely complex and their regulation depends on networks of interactions involving many different factors. The basic reason for this is that, after all, bacteria are not so simple and, in order to survive in the natural environment, they regulate the expression of their genes with precision. The key to this regulation is a large army of transcription factors and $\sigma$ factors.

To date, most of what we know has come from studies on laboratory strains of *E. coli* and a few other model strains. Although the models outlined here, and the principles we have described, are likely to apply to other bacteria, some caution is needed as there may be some surprises waiting for us. The ease with which the whole genome sequences of new strains can be determined and analysed has opened up vast areas of virgin territory for investigation.

In discussing the topic of the activation of bacterial transcription, it is important to realize that the emphasis on transcription factors and $\sigma$ factors comes from the historical legacy of bacterial genetics and mutational analysis. Models derived from these studies usually assume that the DNA template is neutral. However, we know that bacterial DNA is negatively supercoiled and the supercoils may not be evenly distributed. In addition, at some promoters, activity is modulated by methylation of adenine residues. Finally, bacterial DNA is highly compacted by a set of nucleoid–associated proteins. The resulting folded chromosome is unlikely to be a static structure. And, intuitively, it is easy to see how local changes might up- or down-regulate transcription initiation at particular promoters. To date, we have few details concerning such regulation, but this is likely to be a big focus for future efforts.

## Further Reading

Baumberg S (1998) *Procaryotic Gene Expression*. In *Frontiers of Molecular Biology* series, vol. 21. Oxford: Oxford University Press.

Browning D and Busby S (2004) The regulation of bacterial transcription initiation. *Nature Reviews Microbiology* **2**: 57–65.

Gross C, Chan C, Dombroski A *et al*. (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harbor Symposium on Quantitative Biology* **43**: 141–155.

Hughes K and Mathee K (1998) The anti-sigma factors. *Annual Review of Microbiology* **52**: 231–286.

Lawson C, Swigon D, Murakami K *et al*. (2004) Catabolite activator protein: DNA binding and transcription activation. *Current Opinion in Structural Biology* **14**: 10–20.

Lloyd G, Landini P and Busby S (2001) Activation and repression of transcription initiation in bacteria. *Essays in Biochemistry*, vol. 37, chap. 2, pp. 17–32. London: Portland Press.

Murakami K and Darst S (2003) Bacterial RNA polymerases: the whole story. *Current Opinion in Structural Biology* **13**: 1–9.

Murakami K, Masuda S, Campbell E, Muzzin O and Darst S (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme–DNA complex. *Science* **295**: 1285–1290.

Murakami K, Masuda S and Darst S (2002) Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **295**: 1280–1284.

Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst S *et al.* (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**: 1285–1290, pdb accession code 1L9Z.

Ptashne M and Gann A (2002) *Genes and Signals*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schultz SC, Shields GC and Steitz TA (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* **253**: 1001–1007, pdb accession code 1CGP.

Wagner R (2000) *Transcription Regulation in Procaryotes*. Oxford: Oxford University Press.

Vassylyev D, Sekine S, Laptenko O *et al.* (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**: 712–719.

Zhang X, Chaney M, Wigneshweraraj S *et al.* (2002) Mechanochemical ATPases and transcriptional activation. *Molecular Microbiology* **45**: 895–903.