Local and global regulation of transcription initiation in bacteria

Douglas F. Browning and Stephen J. W. Busby

Abstract | Gene expression in bacteria relies on promoter recognition by the DNA-dependent RNA polymerase and subsequent transcription initiation. Bacterial cells are able to tune their transcriptional programmes to changing environments, through numerous mechanisms that regulate the activity of RNA polymerase, or change the set of promoters to which the RNA polymerase can bind. In this Review, we outline our current understanding of the different factors that direct the regulation of transcription initiation in bacteria, whether by interacting with promoters, with RNA polymerase or with both, and we discuss the diverse molecular mechanisms that are used by these factors to regulate gene expression.

RNA polymerase core enzyme
The form of bacterial DNA-dependent RNA polymerase that lacks a sigma factor.

Template strand
The strand of the DNA duplex that acts as a template for RNA synthesis.

Open complex
The complex between RNA polymerase and a promoter after DNA duplex unwinding has occurred and the RNA polymerase is ready to start transcription.

Housekeeping sigma factor
The sigma factor in a bacterium that is responsible for the recognition of promoters that control the transcription of most genes.

Institute of Microbiology and Infection, School of Biosciences, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.
D.F.Browning@bham.ac.uk; S.J.W.Busby@bham.ac.uk

doi:10.1038/nrmicro.2016.103
Published online 8 Aug 2016

In any bacterium, at any moment, expression levels are extremely variable between individual genes. Furthermore, the expression of many genes is subject to variation in response to environmental change. Such regulation of gene expression is crucial for the overall fitness of the bacterial cell. For most genes in most bacteria, the key regulatory step that modulates gene expression seems to be promoter recognition and transcription initiation by RNA polymerase[1]. RNA polymerase core enzyme consists of the large β-subunit and β′-subunit, two α-subunits and the small ω-subunit (FIG. 1a). Each α-subunit consists of independently folded amino-terminal and carboxy-terminal domains that are joined by a flexible linker. The β-subunit and β′-subunit are assembled by binding to the N-terminal domains of the α-subunits, and form a cleft that contains the active site, whereas the ω-subunit is primarily a chaperone for the β′-subunit. However, although competent for DNA-dependent RNA synthesis, the core enzyme is unable to recognize promoters or to initiate transcription without the assistance of one of a set of additional proteins known as sigma factors. The complex that is formed by the binding of a sigma factor to the core enzyme is known as RNA polymerase holoenzyme and is able to orchestrate transcription initiation from specific promoters[2]. Interactions with the promoter position the RNA polymerase holoenzyme such that it unwinds the double-stranded DNA in the region of the transcription start site. Positions +1 and +2 of the unwound template strand (corresponding to the 5′ end of the RNA transcript) are then able to enter the active site to form the transcriptionally competent open complex[3–7]. Subsequent to initiation, the transcription cycle proceeds with escape, elongation and termination steps (FIG. 1a), which have been extensively documented and reviewed elsewhere (see REFS 8–11).

Bacterial promoters contain several discrete sequence motifs, including the –35 element, the extended –10 element, the –10 element and the discriminator region, which are recognized by the sigma factor, and the UP element, which is recognized by the C-terminal domains of the α-subunits[1–3] (BOX 1; FIG. 1b). All bacteria contain one predominant essential sigma factor, known as the housekeeping sigma factor (such as σ^{70} in *Escherichia coli*; also known as RpoD), which is responsible for recognizing most promoters[2]. These housekeeping sigma factors are composed of four structural domains that are connected to one another by flexible linkers. In the RNA polymerase holoenzyme, the sigma factor binds to the subunits of the core enzyme, such that each domain of the sigma factor is positioned to interact with a specific promoter element[2,3]. Contacts involving domain 3 and domain 4 of the sigma factor seem to have the major role in the initial positioning of the RNA polymerase, whereas domain 1 and domain 2 of the sigma factor drive the formation of the open complex[2–7,12] (FIG. 1a). A further function of housekeeping sigma factors, mediated by domain 1, is to ensure that DNA cannot enter the active site until the RNA polymerase is bound to a promoter, which triggers a conformational change that permits DNA access to the active site[13–15].

Transcription initiation can be regulated at the level of the formation of RNA polymerase holoenzyme, promoter recognition by RNA polymerase or RNA polymerase activity, all of which can be mediated by various factors. These regulatory mechanisms can either modulate the function of RNA polymerase itself, or can modulate the accessibility or affinity of promoters for RNA polymerase. In this Review, we discuss the different layers of regulation for each of these targets, focusing on the factors and mechanisms on which they rely.

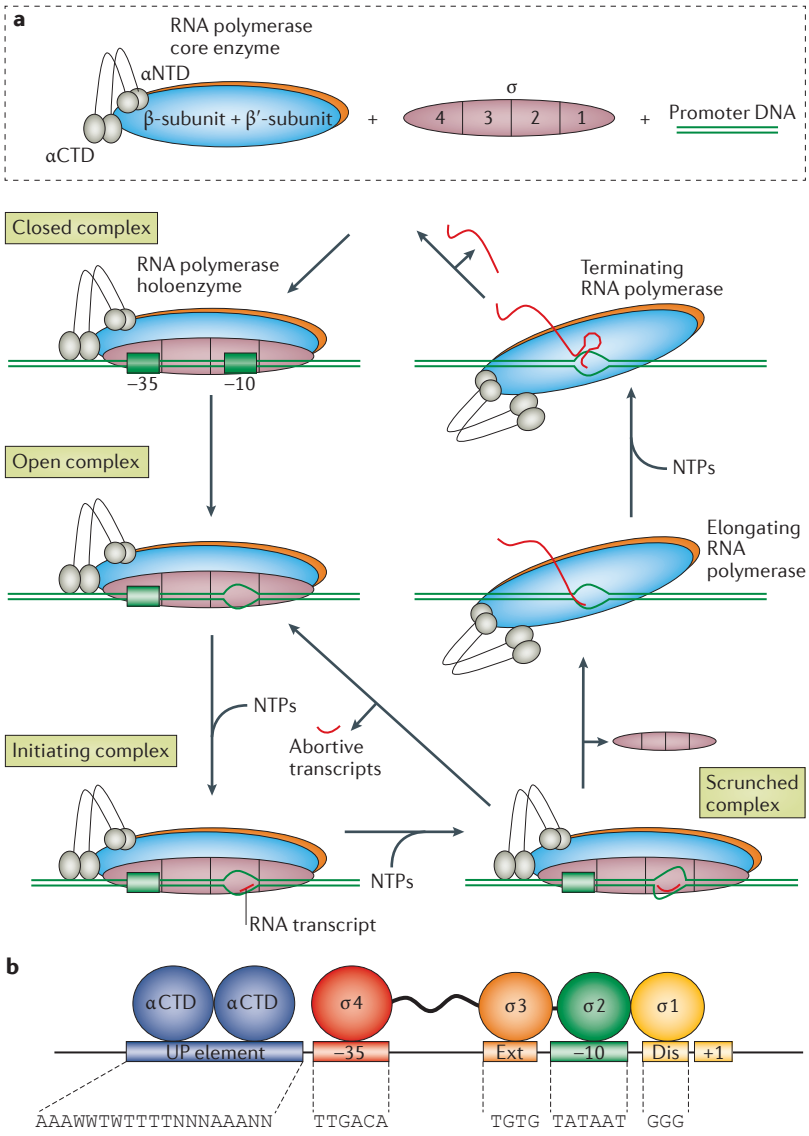


Figure 1 | **Transcription of bacterial genes. a** | The bacterial transcription cycle. RNA polymerase holoenzyme, which comprises the RNA polymerase core enzyme and a sigma factor, interacts with promoter DNA to form the closed complex. The closed complex transitions to the open complex by unwinding the DNA duplex in the region of the transcription start site. The addition of nucleoside triphosphates (NTPs) enables a further transition to the initiating complex, which synthesizes the RNA transcript. Initially, the template strand of the DNA is pulled into the initiating complex, which is a process known as 'scrunching'. The scrunched complex can be held at the promoter, which results in cycles of abortive initiation that only produce small RNA fragments. Alternatively, the RNA polymerase can escape the promoter to enter the elongation phase, leading to the release of the sigma factor and elongation of the RNA transcript using NTPs and elongation factors (not shown). Transcription proceeds until the RNA polymerase encounters a transcriptional terminator, after which the RNA transcript is released and the polymerase dissociates from the DNA template to re-engage with a sigma factor and repeat the cycle. **b** | The principal DNA elements that are recognized by RNA polymerase at bacterial promoters include the UP element (positions −37 to −58, if the transcriptional start site is denoted +1), the −35 element (positions −35 to −30), the extended −10 element (Ext; positions −17 to −14), the −10 element (positions −12 to −7) and the discriminator element (Dis; −6 to −4). Note that the exact positions of each element can vary according to the particular promoter. The regions of the sigma factor (domains 1–4) and of the carboxy-terminal domain of the α-subunit of RNA polymerase (αCTD) that contact these promoter elements are shown. The image in part **b** is adapted from REF. 161, Nature Publishing Group. αNTD, amino-terminal domain of the α-subunit of RNA polymerase.

Although we mainly describe findings from *E. coli* and other model bacterial species, we also consider regulation in non-model bacterial species, and we suggest that understanding the diversity of these mechanisms may shed light on the parallel evolution of other strategies for gene regulation.

RNA polymerase-centred regulation

Many factors interact directly with bacterial RNA polymerase to influence its activity at different promoters. These factors include sigma factors and various other proteins and ligands that regulate either the formation of RNA polymerase holoenzyme or its activity or promoter preferences.

Sigma factors. The role of sigma factors is to guide the positioning of RNA polymerase molecules at promoters and then to orchestrate the formation of the open complex (FIG. 1a). Thus, the promoter specificity of RNA polymerase depends on its sigma factor[2]. Although the housekeeping sigma factor recruits RNA polymerase to the majority of promoters, nearly all bacteria have one or more alternative sigma factors that guide the RNA polymerase to different promoters, which provides a simple mechanism for reprogramming transcription to a different set of genes[2,16]. The available evidence suggests that the housekeeping sigma factor and alternative sigma factors bind to the same site on the surface of the RNA polymerase core enzyme but that, under most conditions, the housekeeping sigma factor is more abundant and thus able to outcompete alternative sigma factors[15,17–21]. However, when the abundance of an alternative sigma factor increases, it can then compete with and displace the housekeeping sigma factor to reprogramme a subset of RNA polymerase molecules[22].

Alternative sigma factors can be grouped into a small number of classes, according to their domain structure, which is usually simpler than that of housekeeping sigma factors[2] (one class has just two domains, corresponding to domain 2 and domain 4 of housekeeping sigma factors[23]). Another difference between alternative sigma factors and housekeeping sigma factors is that most alternative sigma factors bind to a smaller set of promoters, have more stringent sequence specificities and associate with fewer transcription factors than housekeeping sigma factors[24–26], which results in a narrower range of transcriptional outputs. A possible explanation for these differences is that many alternative sigma factors have evolved to switch on the concerted expression of sets of genes that have stress-responsive functions, for which fine-tuning is not required (but see REF. 27 for exceptions).

The activity of alternative sigma factors can be controlled by several mechanisms, including covalent modification, sequestration by cognate anti-sigma factors, subcellular localization, the rate of synthesis and proteolytic turnover[16,28]. In some cases, regulation by anti-sigma factors is countered by anti-anti-sigma factors that sequester anti-sigma factors in response to a particular trigger, in a process termed 'partner switching' (REFS 23,28,29). Some anti-anti-sigma factors share

Box 1 | **Recognition of promoters by RNA polymerase**

Optimizing promoter strength

Bacterial promoter sequences contain motifs known as the –35 element, the extended –10 element, the –10 element, the discriminator region, the UP element and the core recognition element. The activity of a promoter is determined by these elements, with stronger promoters often having elements with sequences that better resemble the consensus sequences (FIG. 1b). Bacteria have evolved a 'mix-and-match' approach to setting promoter strength, in which the strengths of individual promoter elements combine to determine the overall activity of the promoter. This modular organization probably has the benefit of enabling promoters to easily evolve to produce different output levels. Most naturally occurring promoters have sequences for each element that are suboptimal for maximum activity[129], although the strength of each element is often not uniform across the promoter, such that different combinations of elements can generate promoters of similar strength[130]. An additional feature that contributes to promoter strength is the initiating nucleotide, which, for example, is important for transcription initiation at rRNA promoters, at which cellular concentrations of initiating nucleotides (either ATP or GTP) influence the formation and stability of the open complex[68,69].

Double-stranded or single-stranded?

For many promoter elements, recognition by RNA polymerase is specific for double-stranded DNA (dsDNA). For example, the carboxy-terminal domains of the α-subunits recognize the UP element by docking into adjacent minor grooves of the dsDNA[131], the –35 element is recognized specifically as dsDNA by a helix–turn–helix motif in domain 4 of the sigma factor, and the dsDNA of the extended –10 element is bound by a long helix in domain 3 of the sigma factor[2,3]. However, recent structural studies have shown that the –10 element and the discriminator region are both recognized as single-stranded DNA (ssDNA); that is, once the DNA has been unwound during the formation of the transcription bubble[4–7]. To bind to the RNA polymerase, specific bases from the non-template strand of the –10 element (the A and T bases at positions –11 and –7) and the discriminator region (the G at position –6) are flipped out and inserted into pockets in domain 2 and domain 1 of the sigma factor, respectively. The core recognition element, which is formed by bases around the transcription start site, is recognized by residues in the β-subunit and β′-subunit of the RNA polymerase[5–7]. The order in which each element is recognized by RNA polymerase during transcription initiation has been examined using rapid kinetic methods to study the A1 promoter of the T7 phage, which showed that the UP element and the –35 element are bound first, followed by unwinding of the DNA and recognition of the –10 element[132]. Other studies of the *lacUV5* and λP_R promoters showed that RNA polymerase seems to bind to dsDNA at the –10 element and downstream DNA, before the DNA is unwound to produce ssDNA[133,134].

Promoters as 'bottlenecks'

By visualizing fluorescently labelled RNA polymerase in live cells[135] or by measuring transcript levels as a proxy for RNA polymerase activity[136], studies in *Escherichia coli* have shown that a pool of RNA polymerase holoenzyme that scans for promoters seems to be present on bacterial DNA in most conditions. However, as a promoter can only be occupied by a single RNA polymerase at any time, transcript formation at many promoters is subject to a 'bottleneck' effect, in which the formation of the open complex and transition to the elongation step of the transcription cycle become rate-limiting[137–139]. Accordingly, experiments that used chromatin immunoprecipitation (ChIP) to measure promoter occupancy have shown that RNA polymerase is stalled at many promoters, which results in decreased flux of RNA polymerase through specific transcription units[140,141].

structural similarities with the cognate sigma factor, which suggests that they have evolved to be mimics[30,31]. All of these regulatory mechanisms combine to form complex regulatory circuits that adjust the number of RNA polymerase molecules that are programmed by a particular alternative sigma factor according to the environmental conditions.

Despite our knowledge of this circuitry, many details of the mechanisms that mediate sigma factor exchange remain unclear. For example, we do not know how many alternative sigma factors, which have weak binding affinities for the RNA polymerase core enzyme, are able to recruit sufficient RNA polymerase molecules to have a discernible effect on the transcriptional programme[22]. One well-studied case of sigma factor exchange occurs between σ^{38} (also known as RpoS) and the housekeeping sigma factor, σ^{70}, in *E. coli*, in response to certain stresses or at entry into stationary phase[20,21]. Displacement of σ^{70} following an increase in the abundance of σ^{38} is facilitated by Crl[32–34], which interacts with σ^{38}, and by Rsd[35–37], an anti-sigma factor that reduces the level of functional σ^{70}, thereby biasing the competition between the two sigma factors in favour of σ^{38}. In addition, the small guanine nucleotide guanosine tetraphosphate (ppGpp), which undergoes a sharp increase in abundance in response to certain metabolic responses, may have a role in the exchange of some sigma factors[38,39].

Nearly all alternative sigma factors are evolutionarily related to housekeeping sigma factors, consisting of two, three or four domains that retain common functions in their different contexts[2]. However, most bacteria contain one additional alternative sigma factor that is 'in a class of its own' and belongs to the family of σ^{54} (also known as RpoN) sigma factors, which is evolutionarily unrelated to other sigma factors[40,41]. These sigma factors recognize different elements at target promoters to other sigma factors, with specificity determined by elements at positions –24 and –12, rather than –35 and –10 (REF. 42). Crucially, unlike the complex that is formed by σ^{70}, RNA polymerase holoenzyme that contains σ^{54} is unable to proceed to the open complex and requires activation by a special class of ATP-dependent activators (see below)[42]. Recent structural data has shown that the site at which the template strand needs to be located for the formation of the open complex is blocked by parts of σ^{54} (REF. 19).

Chromatin immunoprecipitation
(ChIP). A method whereby antibodies are used to isolate DNA fragments that have been cross-linked to a specific protein.

Stationary phase
The period when bacteria have stopped growing.

Guanosine tetraphosphate
(ppGpp). A small molecule that is synthesized in response to certain stresses. ppGpp is often referred to as 'magic spot', which is a term that also refers to guanosine pentaphosphate (pppGpp).

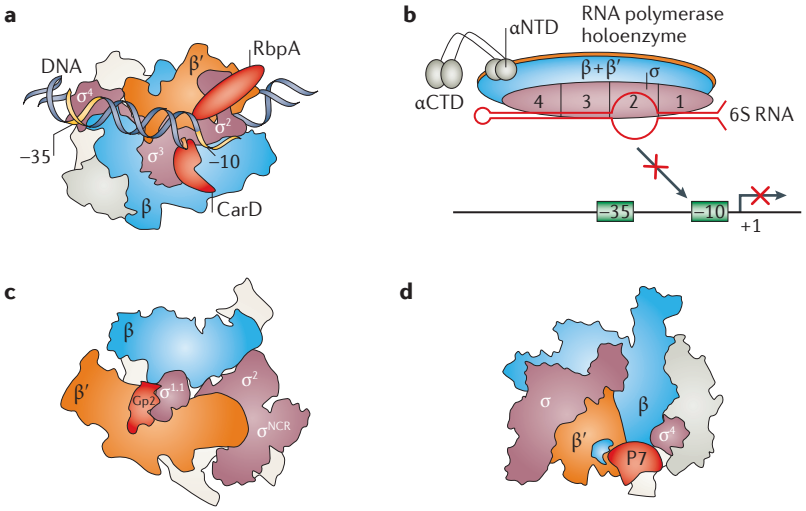


Figure 2 | **Modulation of RNA polymerase activity. a** | Open complex stabilizers. Both RbpA and CarD are global transcription factors that bind directly to RNA polymerase to regulate transcription at promoters. RbpA contacts domain 2 of the housekeeping sigma factor, whereas CarD binds to the β-subunit of RNA polymerase. Both are positioned to stabilize the open complex by interacting with promoter DNA in addition to the RNA polymerase. **b** | Regulation by a promoter mimic. The structure of the 6S RNA mimics that of a promoter in an open complex, which enables the 6S RNA to sequester RNA polymerase holoenzyme that contains the housekeeping sigma factor and thus downregulate global transcription. **c** | Sequestration of a sigma factor domain. On infection of *Escherichia coli* by phage T7, the phage Gp2 protein binds to two regions of the *E. coli* RNA polymerase holoenzyme: the channel that contains the active site and domain 1 of the sigma factor. These interactions inhibit transcription initiation by preventing DNA from entering the active site and by stopping the movement of domain 1 of the sigma factor, respectively. **d** | Sigma factor displacement. On infection of *Xanthomonas oryzae* by phage Xp10, the phage protein P7 binds to the β-subunit and β′-subunit of the bacterial RNA polymerase, which inhibits transcription by displacing domain 4 of the sigma factor from the holoenzyme when the RNA polymerase is associated with promoter DNA. The image in part **a** is adapted with permission from REF. 44, National Academy of Sciences. The image in part **c** is a adapted with permission from REF. 14, National Academy of Sciences. The image in part **d** is adapted from Liu, B. *et al.* A bacteriophage transcription regulator inhibits bacterial transcription initiation by sigma-factor displacement. *Nucleic Acids Res.* (2014), **42**, 7, 4294–4305, by permission of Oxford University Press. αCTD, carboxy-terminal domain of the α-subunit of RNA polymerase. σNCR, sigma factor non-conserved region. αNTD, amino-terminal domain of the α-subunit of RNA polymerase.

Interestingly, different activators are required at different target promoters and, at least in *E. coli*, levels of σ^{54} seem to be unregulated[21]. Together, these features of σ^{54} suggest that, in keeping with being in a class of its own, σ^{54} uses a different regulatory strategy to other sigma factors, in that transcription initiation is regulated by the presence of an activator rather than the abundance of the sigma factor.

Regulators of holoenzyme activity. The simplicity and ubiquity of regulation by sigma factors have detracted attention from numerous other regulatory factors that interact with RNA polymerase (FIG. 2). Although many of these regulatory factors act at the level of transcript elongation and termination (such as N-utilization substance (Nus) factors)[9], and are, as such, beyond the scope of this Review article, a small number are, similarly to sigma factors, regulators of transcription initiation. Some of these regulators are restricted to particular groups of bacteria, such as RbpA and CarD, which bind to and stabilize open complexes in Actinomycetes[43] (FIG. 2a). Recent studies have shown that RbpA binds directly to domain 2 of the housekeeping sigma factor and is positioned to make direct stabilizing contacts with the upstream end of unwound DNA in open complexes[44]. CarD has a C-terminal domain that makes stabilizing contacts that are complementary to those made by RbpA, which suggests that RbpA and CarD function coordinately, whereas the N-terminal domain of CarD interacts with the RNA polymerase β-subunit[45,46].

Although *E. coli* and related members of the Enterobacteriaceae lack RbpA and CarD, these bacteria use other factors to modulate the stability of open complexes. For example, the extended coiled-coil protein motif in DksA inserts into a narrow channel (known as the secondary channel) that leads from the surface of the RNA polymerase to the active site[47–50]. The insertion of DksA, in cooperation with ppGpp, selectively stabilizes or destabilizes open RNA polymerase–promoter complexes, depending on the promoter, which represents a sigma factor-free mechanism to reprogramme the transcriptome in response to increased levels of ppGpp, which can result from certain metabolic stresses[51,52]. The mechanism of regulation by DksA awaits elucidation but probably involves a lowering of the energy barrier between the RNA polymerase–promoter closed complex and open complex[47–52]. The structure of DksA and its mode of binding to the polymerase resemble those of the GreA and GreB transcription elongation factors[47,53], and recent studies suggest that DksA, GreA and GreB all function as 'inspectors' that continually probe the enzyme active site by making rapid transient 'visits' to the secondary channel[54–56].

Other factors decrease the number of RNA polymerase molecules that are available for transcription by sequestering the holoenzyme. One example is 6S RNA, an approximately 180-nucleotide non-coding RNA that is synthesized in response to slow growth and forms a 1/1 complex with the RNA polymerase holoenzyme[57] (FIG. 2b). In *E. coli*, 6S RNA is a mimetic for the DNA of promoters that are targets for the housekeeping RNA polymerase holoenzyme (that is, holoenzyme that contains σ^{70}). Consequently, the proportion of functional (that is, non-sequestered) RNA polymerase holoenzymes that contain σ^{38} increases, but the total transcription potential of the cell is reduced, as 6S RNA sequesters the housekeeping RNA polymerase holoenzyme[57]. 6S RNA seems to work cooperatively with the Rsd anti-sigma factor (see above) to couple gene expression to growth, as the levels of both factors increase as cell growth slows[37].

Finally, some phages have factors that inhibit the activity of bacterial RNA polymerases to favour the activity of their own bespoke RNA polymerases. For example, during infection of *E. coli* cells, the Gp2 protein of phage T7 induces a conformational change in a part of domain 1 of σ^{70} that blocks the access of template DNA to the active site[14] (FIG. 2c). By contrast, the P7 protein

Actinomycetes
A class of soil bacteria with a particular morphology.

Coiled-coil
An extended motif found in proteins.

Closed complex
The complex between RNA polymerase and a promoter before DNA duplex unwinding has occurred.

of the related phage Xp10, which infects *Xanthomonas oryzae*, inhibits the activity of RNA polymerase by displacing the sigma factor from the holoenzyme; the association of P7 with RNA polymerase is stable, which prevents the recruitment of a replacement sigma factor to reform the holoenzyme[58] (FIG. 2d).

Other factors that regulate RNA polymerase holoenzyme. Some regulators of RNA polymerase holoenzyme, known as 'appropriators', remodel parts of the polymerase to alter promoter preferences (FIG. 3), and thus have a similar function to alternative sigma factors. The most well-studied example of a phage appropriator is the AsiA protein of phage T4, which infects *E. coli*. As phage T4 does not encode its own RNA polymerase, its invasion strategy is to appropriate, rather than to silence, the host RNA polymerase. AsiA is encoded by a gene that is expressed during the early stage of the T4 life cycle, and remodels and repositions domain 4 of σ^{70} so that it is unable to recognize –35 elements[59,60] and instead becomes a target for the T4 MotA transcription activator. MotA, which is encoded by early-stage genes, is the essential activator for transcription of genes that are expressed during the middle stage of the T4 life cycle[61] (FIG. 3a). Thus, AsiA and MotA redirect the housekeeping RNA polymerase holoenzyme from the transcription of host genes to the transcription of phage genes. Interestingly, the T4 proteins Alt and ModA also alter the promoter preferences of the host RNA polymerase, although the mechanism by which they do so is to ADP-ribosylate residue R265 in the C terminus of either one or both RNA polymerase α-subunits. This modification prevents the RNA polymerase from recognizing UP elements and so helps divert the polymerase from transcribing host genes to transcribing T4 genes[61] (FIG. 3b).

The two examples of host-encoded appropriators that have been studied in the most detail are SoxS in *E. coli* and Spx in *Bacillus subtilis*. SoxS targets RNA polymerase holoenzyme to promoters that have upstream Sox-box sequences and is essential for the induction of dozens of genes in response to the sensing of oxidative stress by the SoxR repressor, which triggers an increase in the abundance of SoxS. Evidence suggests that SoxS guides the RNA polymerase holoenzyme to target promoters using a 'pre-recruitment' mechanism; that is, SoxS contacts the DNA-recognition surfaces of the C-terminal domain of each RNA polymerase α-subunit before the binding of the SoxS–RNA polymerase holoenzyme complex to the DNA[62,63]. The SoxS–RNA polymerase holoenzyme complex is unable to recognize UP elements but instead recognizes Sox-boxes[63] (FIG. 3c). Similarly, in response to oxidative stress, Spx binds to the C-terminal domains of the RNA polymerase holoenzyme α-subunits, which alters the promoter preference such that certain promoters are favoured and others are disfavoured[64] (FIG. 3d). Structures of Spx in complex with the C-terminal domain of an α-subunit suggest that the formation of the complex does not produce any major structural change in either Spx or the α-subunit[65,66].

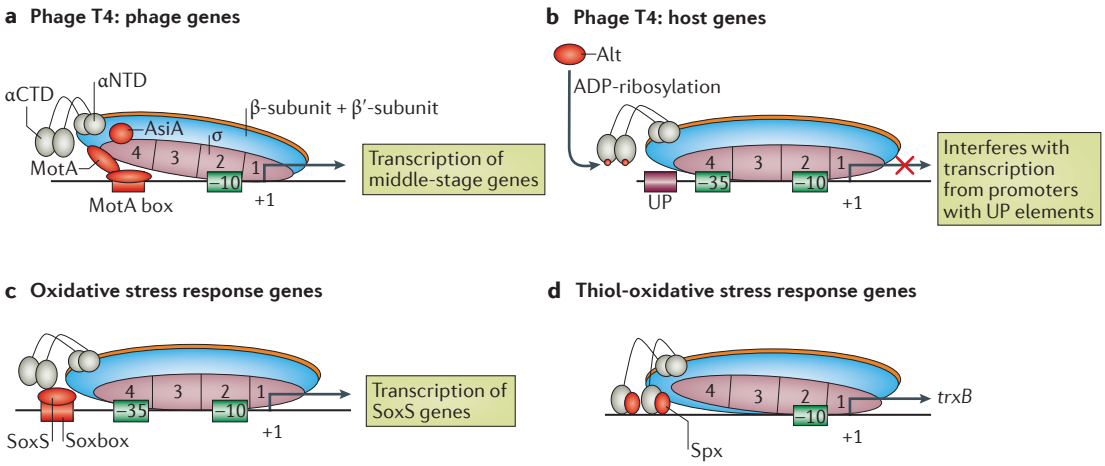


Figure 3 | **Appropriation of RNA polymerase for specific transcriptional programmes. a** | During infection, the AsiA protein of phage T4 remodels the structure of domain 4 of the *Escherichia coli* housekeeping sigma factor. The remodelled sigma factor interacts directly with another phage protein, MotA, which is a transcription activator that binds to promoters of middle-stage genes of the T4 infection cycle. The interaction between the remodelled sigma factor and MotA recruits RNA polymerase to activate the transcription of these genes. **b** | Early during infection, the phage T4 protein Alt ADP-ribosylates amino acid residue R265 on the carboxy-terminal domains of the α-subunits of *E. coli* RNA polymerase (αCTD). This modification modulates the recognition of UP element sequences in bacterial promoters, thereby increasing the availability of RNA polymerase for transcription of phage genes. **c** | In *E. coli*, SoxS is expressed in response to oxidative stress. By binding to the C-terminal domains of the α-subunits of RNA polymerase, SoxS directs RNA polymerase to promoters that contain 'Sox-box' SoxS-recognition sites in upstream regions, which leads to the expression of genes that are controlled by these promoters. **d** | In *Bacillus subtilis*, the Spx protein is activated in response to thiol-oxidative stress. Activated Spx is able to associate with the C-terminal domains of the α-subunits of RNA polymerase, which modulates the binding preferences of RNA polymerase such that the expression of genes such as *trxB* is induced to combat thiol-oxidative stress. αNTD, the amino-terminal domains of the α-subunits of RNA polymerase.

Michaelis constant
The concentration of a substrate at which the reaction catalysed by an enzyme proceeds at half of its maximum speed.

Initiating nucleotide
The 5′ nucleotide of a transcript.

Nucleoid
The structure that forms after a bacterial chromosome is compacted inside a bacterium.

Superhelical density
The measure of the degree to which the winding of one DNA strand around the other differs from the periodicity of the Watson–Crick structure.

Finally, in addition to regulation by proteins, the activity of RNA polymerase holoenzyme can be regulated by fluctuations in the levels of its four nucleoside triphosphate (NTP) substrates. Note that the Michaelis constant for the initiating nucleotide is higher than that for subsequently added NTPs[67]; that is, for a given rate of transcription, the required concentration of the initiating NTP is higher than that of subsequent NTPs, which means that the concentration of the initiating NTP is most crucial to the activity of the RNA polymerase. As the initiating NTP for rRNA transcripts in *E. coli* is ATP, transcription initiation at rRNA promoters is expected to be sensitive to the cellular concentration of ATP, which increases as cells leave stationary phase with a concomitant burst of rRNA synthesis[68,69]. In *B. subtilis*, the initiating NTP for rRNA is GTP, which decreases in abundance when it is consumed by the synthesis of ppGpp. Therefore, ppGpp levels in *B. subtilis* indirectly couple metabolism to RNA polymerase activity through corresponding changes to GTP levels[70,71]. This contrasts with *E. coli*, in which the cooperative action of ppGpp and DksA directly regulates the stability of RNA polymerase holoenzyme open complexes[51] (see above).

More complex effects of NTP levels are observed during transcription initiation at *E. coli* promoters that control genes that are involved in pyrimidine biosynthesis. At many of these promoters, fluctuations in the levels of UTP or CTP affect transcript stability and translation efficiency. This can occur either by changing the location of the transcript start site or by non-templated insertion of extra bases into the nascent transcript[72]. As UTP and CTP are both pyrimidine-based nucleotides, this form of regulation produces a feedback mechanism in which pyrimidine biosynthesis can be adjusted in response to changes in the levels of these nucleotides[72].

Promoter-centred regulation

Regulation by factors that bind directly to the RNA polymerase is complemented by factors that directly target the promoter DNA (FIG. 4). Such regulation can be mediated by many different mechanisms, which range from supercoiling to transcription factors that simply block access to the promoter. In this scenario, first imagined by François Jacob and collaborators (BOX 2), transcription factors have structural motifs that bind to promoters that contain cognate 'operators', which are specific sequences of base pairs. The operators for most bacterial DNA-binding structural motifs, such as the helix–turn–helix motif, have 4–5 base pairs. As any given 5-base pair sequence will arise, on average, once every 4^5 base pairs, further specificity is required to target transcription factors only to the desired promoters. Such specificity can be achieved using various mechanisms, including homodimerization (or higher order multimerization) of the transcription factor, association of the transcription factor with another DNA-binding factor, or the incorporation of several DNA-binding structural motifs into a single transcription factor. Of these mechanisms, dimerization or further multimerization is the most common, and therefore most operators contain direct or inverted repeats of a 4–5 base pair sequence.

The key role of transcription factors is to couple promoter activity to environmental cues, so their activity has to be regulated in response to these cues. In many transcription factors, regulatory modules mediate DNA binding in response to the cues that are sensed by the binding of a small ligand or protein, or covalent modification. The activities of other transcription factors depend on abundance and availability, which can be regulated by synthesis, turnover or sequestration. Although some transcription factors regulate only a single promoter, most transcription factors regulate many promoters. In addition, at least in *E. coli*, most promoters are regulated by more than one factor. Furthermore, many transcription factors are expressed from promoters that are themselves regulated by other transcription factors. Promoter regulation by transcription factors therefore generates a complex regulatory network in which the concerted activities of specific, global and master regulators orchestrate the distribution of RNA polymerase to the various transcription units that are present in the genome[73–76].

It is important to appreciate that transcription initiation occurs in the context of the bacterial nucleoid, and that the compaction of the bacterial chromosome that occurs during the formation of the nucleoid is thought to have an overall negative effect on promoter activity. Therefore, modulation of the local structure of the nucleoid to change the level of compaction may result in changes in transcription. Compaction is primarily due to supercoiling and to the contortions that are induced in the DNA by nucleoid-associated proteins (NAPs), which 'sculpt' the bacterial chromosome by bending, wrapping, looping and twisting DNA such that it fits inside the bacterial cell[77] (BOX 3). Many NAPs are abundant and bind to numerous target sites using a relaxed sequence specificity[78–80]. By contrast, some NAPs, such as factor for inversion stimulation (Fis) and integration host factor (IHF), clearly do have sequence specificity, which enables them to mediate promoter-specific effects on the local DNA structure, such that they can be regarded as transcription factors[81]. For example, Fis represses the activity of the promoter that controls the expression of the major stationary phase DNA-binding protein, Dps, so that the promoter is active in stationary phase but inactive in stress-free rapidly growing cells[82]. In parallel, global alterations in supercoiling, which are probably mediated by fluctuations in the levels of ATP, lead to upregulation and downregulation at scores of promoters; in some cases, this results in local superhelical density changes that modulate the transcription activity at specific target promoters[83–88].

Recent observations that the activities of certain promoters vary according to their precise location in the folded bacterial chromosome may be best explained by the suggestion that the availability of free RNA polymerase holoenzyme, or at least the degree to which the enzyme can access promoter regions, is not the same at every location[89]. Thus, the long-standing assumption that promoter activity in bacteria is determined only

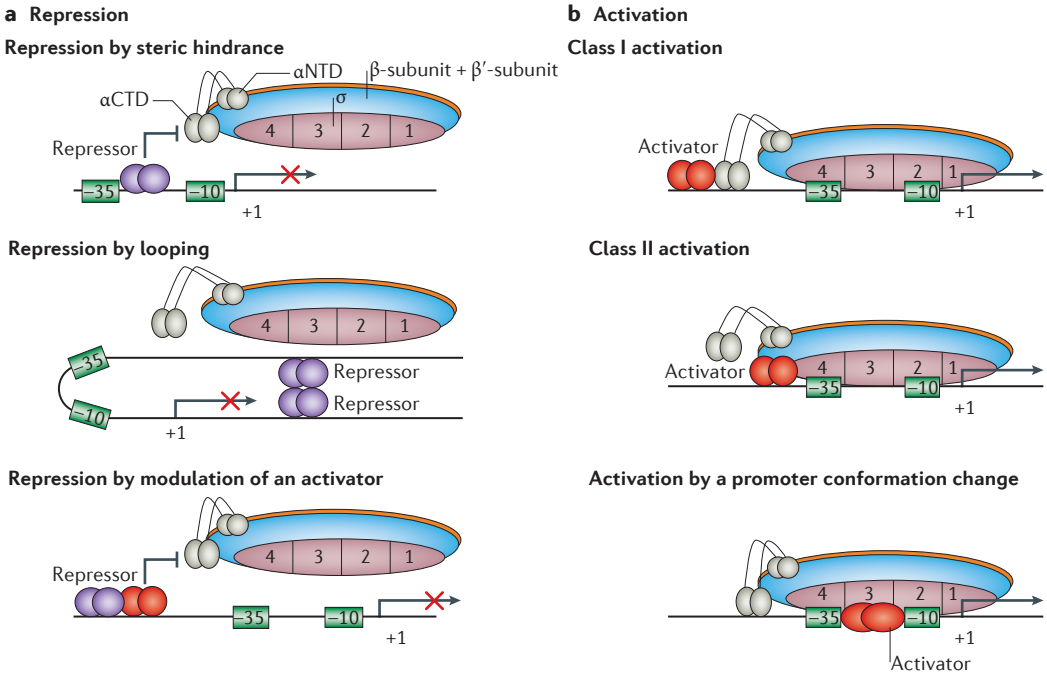


Figure 4 | **Repression or activation at promoters by transcription factors.** Transcription factors known as repressors and activators use one of several mechanisms to repress and activate transcription initiation, respectively. The principal mechanisms are shown here, using the example of the RNA polymerase holoenzyme that contains σ^{70} in *Escherichia coli*. **a** | In repression by steric hindrance, the repressor binds to a site that overlaps the core elements of the promoter that is recognized by RNA polymerase and thus blocks the binding of RNA polymerase to the promoter. In repression by looping, protein–protein interactions are formed between repressors that bind to sites upstream and downstream of the promoter, inducing looping of the DNA between the two repressor binding sites. DNA looping prevents the recognition of promoter elements by RNA polymerase, and thus transcription is repressed. For those genes that require activators for transcription, repressors can modulate activators to prevent recruitment of RNA polymerase. Although these repressors indirectly modulate RNA polymerase activity by targeting activators, in contrast to repressors that directly modulate RNA polymerase activity by masking promoter elements, the ultimate effect remains to prevent recognition of the promoter and thus repress transcription. **b** | In class I activation, the activator binds to a site upstream of the promoter and recruits RNA polymerase to the promoter by contacting the carboxy-terminal domain of the α-subunit (αCTD). In class II activation, the activator binds to a site in the promoter adjacent to (or overlapping with) the –35 element, where it recruits RNA polymerase through direct interactions with domain 4 of the sigma factor. Rather than directly recruit RNA polymerase, some activators induce a conformational change in the promoter DNA to activate transcription. These activators bind at, or near to, the core RNA polymerase recognition elements of the promoter and often realign the –10 and –35 elements so that they can be recognized by RNA polymerase, thereby enabling the recruitment of RNA polymerase to the promoter and activation of transcription. αNTD, the amino-terminal domains of the α-subunits of RNA polymerase.

by local transcription factors and by how sequences of promoter elements correspond to consensus sequences (BOX 1) is an oversimplification.

Repressive transcription factors. Several mechanisms have been described for the repression of transcription initiation by repressive transcription factors (known as 'repressors'). At many promoters, repression of transcription initiation occurs simply by steric hindrance, in which repressors bind to operators that overlap the –10 or –35 elements of the promoter and thus block access of RNA polymerase to the promoter[1] (FIG. 4a). Many promoters contain arrays of operators that are organized adjacently to each other. For these promoters, a larger number of bound repressors usually increases the strength of repression. At other promoters, the simultaneous binding of a repressor to operators that are located distally to one another gives rise to a loop in the local DNA structure that prevents binding by RNA polymerase[90] (FIG. 4a). The best-documented case of repression by DNA looping is mediated by the GalR repressor of the galactose operon in *E. coli*. The operators that are recognized by GalR are located upstream and downstream of the different promoter elements of the *gal* operon, but do not overlap with these elements[90,91]. Some promoters can be bound (usually independently) by two or more different repressors, which can lead to very tight repression, such as at promoters that control the expression of colicin[92,93].

At some promoters, repression is more complicated and is mediated by 'anti-activator' repressors (FIG. 4a), which counter the function of activating transcription factors ('activators'; see below). For example, the CytR repressor is as an anti-activator that simultaneously interacts with its operator and an adjacent activator, the cyclic AMP (cAMP) receptor protein (CRP; also known as

Box 2 | **50 years of studying gene regulatory mechanisms**

In 1965, François Jacob, André Lwoff and Jacques Monod received a Nobel Prize for 'discoveries concerning the genetic regulation of enzyme and virus syntheses'. Together, they had discovered the mechanism by which λ phages that infect *Escherichia coli* are silenced as lysogens until exposure to UV light, and the mechanism by which β-galactosidase activity is 'silenced' until *E. coli* is grown in a medium containing lactose. In both cases, the mechanism centred on a transcriptional repressor. As the fundamentals of molecular biology had only just been established, the regulation of transcription had not been considered as a possible mechanism when the work was started. Indeed, the idea that the solution was a transcriptional repressor was resisted by Monod[142] until he was persuaded by the clarity of genetic experiments[143]. The discovery that these functions are regulated at the level of transcription led to the development of a research area that covered promoters, transcription start sites and operators, together with all of the protein factors and other components that mediate the induction and repression of gene expression. From these studies emerged the concept that what matters is not the genes one has, but how one expresses them, and it is no exaggeration to state that all work in this area has its origins in the discovery of transcriptional repressors. However, although his work represented fundamental contributions to biological research, two shortcomings are notable in the writings of Monod. First, he assumed that all regulation would be explained by the activity of repressors and, unlike Jacob, was never persuaded that transcriptional regulation might also rely on activators[144]. Second, the notion popularized by Monod that 'anything found to be true of *E. coli* must also be true of elephants' gave undue prominence to *E. coli* and its *lac* operon as a model for molecular biology, the limitations of which were appreciated by Jacob[121] and, more recently, articulated by Victor de Lorenzo as the black cat–white cat principle[145].

catabolite activator protein (CAP)). At some promoters, CytR binding requires a combination of CytR–CRP and CytR–DNA interactions to prevent the binding of RNA polymerase[94]. Another complex mechanism of repression can occur when the repressor interacts directly with bound RNA polymerase. For example, the p4 protein encoded by phage φ29, which infects *B. subtilis*, simultaneously binds to the C-terminal domain of the α-subunit of RNA polymerase and to the DNA upstream of the polymerase, thereby preventing promoter clearance[95].

Activating transcription factors. Activating transcription factors (known as 'activators') increase the activity of promoters, often from low basal levels, by one of three mechanisms: class I activation, class II activation or activation by a conformational change[96] (FIG. 4b). In class I activation, the activator binds to an operator that is located upstream of the promoter elements and then recruits RNA polymerase to the promoter through interactions formed between a small surface-exposed patch on the activator (known as an activating region) and the C-terminal domain of the RNA polymerase α-subunit[96]. Structural analysis of activation by CRP shows that this interaction occurs without any structural change in either partner, and, as such, is often referred to as 'velcro' (REF. 97). Furthermore, class I activation occurs at promoters that require activators to recruit RNA polymerase because one or more of the promoter elements has a sequence that is suboptimal for the binding of RNA polymerase. Thus, for those promoters that are dependent on class I activation, RNA polymerase is recruited by protein–protein interactions rather than protein–DNA interactions. Optimal class I activation occurs when the activator binds to the same face of the DNA helix as the RNA polymerase α-subunit[98], which facilitates the interaction between the two proteins such that activation is reduced if the operator is displaced by 1–5 base pairs (corresponding to up to half a turn of the helix) but is restored if the displacement is approximately 10 base pairs (corresponding to a full turn of the helix). In some cases, decreased activation due to suboptimal placement of an activator can be reversed if the bases of one or more of the promoter elements are modified to enhance the binding of RNA polymerase[99].

In class II activation, the activator binds to an operator that overlaps the –35 element of the promoter. Once bound to the operator, discrete activating regions in the activator recruit the RNA polymerase by forming direct interactions with domain 4 of the sigma factor, with the N-terminal domain of the α-subunit or with other components of the RNA polymerase[96] (FIG. 4b). At some promoters, the interactions between the activator and the RNA polymerase are more favourable in the open complex, which facilitates the transition from the closed to the open complex[100–103]. One of the consequences of class II activation is that the C-terminal domains of the RNA polymerase α-subunits are unable to bind to their preferred binding site immediately upstream of the –35 promoter element and instead bind immediately upstream of the activator[96]. This permits class II activators to function together with class I activators, and this combination is responsible for synergy at many bacterial promoters in which activity is co-dependent on two signals[1].

A third mechanism by which a single activator can drive transcription initiation at a promoter involves the binding of the activator to an operator that is located between the –35 and –10 elements of the promoter (FIG. 4b). Promoters that are thus regulated usually have non-optimal spacing between the two elements, and the current view is that activation involves distortion of the DNA so that the –35 and –10 elements of the promoter are better positioned for RNA polymerase binding[104,105]. The available structural evidence indicates that this distortion is unevenly spread across the DNA between the –35 and –10 elements[105,106].

In the above examples of 'direct' activation, the activator alters the target promoter to make it more 'attractive' to the RNA polymerase. However, activation can also occur indirectly, usually by removing a repressor, and in these cases the activator can be thought of as an anti-repressor[1]. For example, in enterohaemorrhagic *E. coli*, the expression of one of the principal virulence determinants is repressed by histone-like nucleoid-structuring protein (H-NS), but the expression of Ler, an H-NS homologue, disrupts the repression and triggers virulence[107,108]. A more complex example occurs at the *nir* promoter, which controls the expression of a nitrite reductase in *E. coli*. Class II activation, repression by NAPs and indirect activation by anti-repressors all compete to regulate the activity of this promoter: class II activation of *nir* is mediated by an activator known as Fnr (fumarate and nitrate reduction regulatory protein) in response to oxygen deprivation, but is suppressed by the NAPs IHF and Fis, which bind to specific target sites that are located just upstream of the DNA binding site for Fnr[109,110]. However, this suppression is disrupted by either NarL or

Enterohaemorrhagic *E. coli*
A virulent strain of *Escherichia coli* that causes bloody diarrhoea.

Box 3 | **Nucleoid-associated proteins**

Although not enclosed by a membrane in a nucleus, bacterial chromosomes are compacted and highly folded, forming a structure known as the nucleoid. This compaction enables large bacterial chromosomes to fit inside cells and is achieved by the combined effects of DNA supercoiling, molecular crowding, and the presence of RNA and nucleoid-associated proteins (NAPs). Initially, it was thought that NAPs would have similar structures to eukaryotic histones, but it is now clear that NAPs are instead a diverse group of proteins that recognize DNA target sites using many different structural motifs[77,146]. *Escherichia coli* has at least 12 different NAPs, including factor for inversion stimulation (Fis), integration host factor (IHF), Histone-like protein originally isolated from an *E. coli* strain labelled U93 (HU), histone-like nucleoid-structuring protein (H-NS)) and leucine-responsive protein (LRP), all of which have been studied extensively[77,79]. The most studied example is probably H-NS, which forms filaments along AT-rich segments of DNA; these filaments then form bridges between different segments to repress the activity of any associated promoter[147]. How the activity of H-NS is modulated is not fully understood, but one mechanism that is used to relieve repression is the expression of H-NS homologues, such as Ler, that are incorporated into the filaments but are unable to form the repressive bridges[148].

Much of what we understand about NAPs comes from studies of *E. coli*; however, the set of NAPs that are present in a bacterium varies according to the species. For example, Fis is restricted to members of the Gammaproteobacteria, whereas HU is found in almost all bacteria and is the principal NAP in *Bacillus subtilis*, in which it is an essential protein[77,149,150]. Some species in the Actinomycetes have a NAP, Hlp, that has an amino-terminal domain that is similar to HU, but a carboxy-terminal domain that resembles eukaryotic histone H1 (REF. 151). Adding further complexity, the levels of some NAPs change in response to growth conditions and growth phase, which means that the composition of the nucleoid of a cell can change over time. For example, rapidly growing *E. coli* cells at early exponential phase contain more than 50,000 molecules of Fis, whereas the number of Fis molecules in cells at stationary phase is almost zero[78,152].

Characterization of the binding preferences of individual NAPs has shown that some of these proteins, such as IHF and Fis, are sequence-specific DNA-binding proteins, whereas other NAPs, such as H-NS and HU, bind to DNA nonspecifically, often to distorted DNA structures[79]. Biochemical, biophysical and single-molecule experiments have shown that NAPs package DNA by several different mechanisms. For example, IHF and Fis package DNA by bending the target site, whereas, as mentioned above, H-NS forms bridges between DNA molecules[77,146,153–155]. Chromatin immunoprecipitation (ChIP) studies have determined the genome-wide binding profiles of some NAPs, which have shown that many NAPs bind to hundreds of target sites[156–158]. Many of these target sites are located in intergenic regulatory regions, to which several different NAPs may bind[156]. Thus, it is not surprising that many examples have been described in which NAPs regulate gene expression by directly modulating promoter activity[81]. For example, the formation of extended filaments of H-NS on DNA silences many promoters[147], and even prevents the unwanted expression of horizontally acquired AT-rich genes, which often encode virulence factors[159]. By contrast, Fis is able to repress transcription initiation through various mechanisms and can also activate transcription directly by interacting with RNA polymerase holoenzyme using an activating region[127,128]. Other NAPs can prevent transcription across a much larger segment of a bacterial chromosome and a genome-wide survey of the *E. coli* chromosome identified more than 100 loci at which transcription seemed to be silenced[160]. The ability of NAPs to regulate transcription suggests that evolution has blurred the distinction between NAPs and transcription factors so as to ensure both nucleoid integrity and coordinated gene regulation[124].

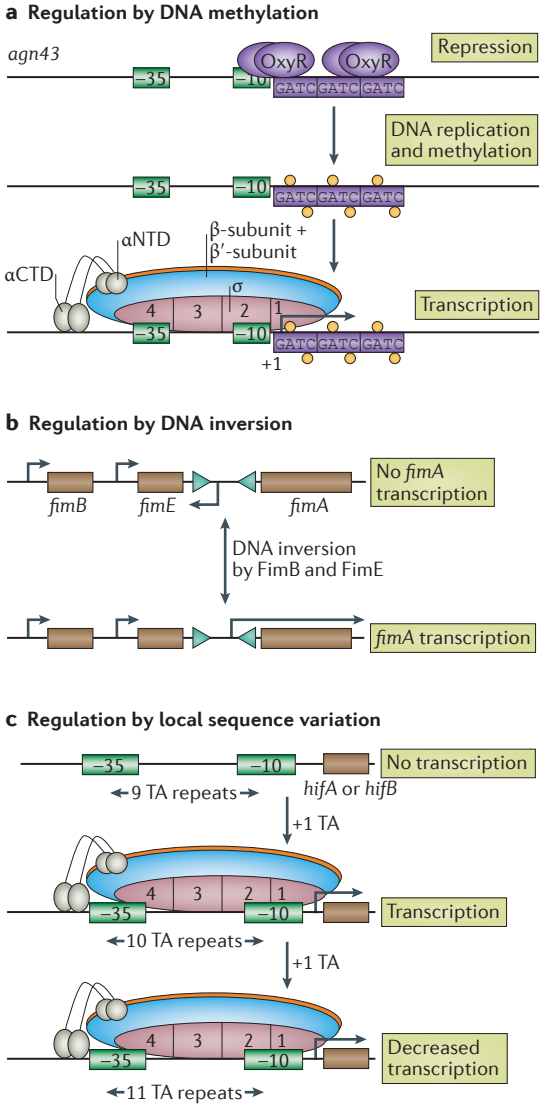
NarP, which are anti-repressors that have activities that are regulated by levels of extracellular nitrate and nitrite ions[111]. Thus, expression from Fnr-regulated promoters integrates two different environmental signals, oxygen and nitrate (or nitrite).

The primary role of many activation mechanisms is simply to facilitate the recruitment of RNA polymerase to the promoter, rather than to promote the transition of the RNA polymerase closed complex to the open complex, which, in any case, does not require an activator for most RNA polymerase holoenzymes. However RNA polymerase holoenzymes that contain σ^{54}-family sigma factors are unable to make this transition without the assistance of activators known as enhancer binding proteins (EBPs), which use ATP hydrolysis to induce a conformational change that drives the formation of the open complex and enables transcriptionally competent complexes to be formed[112]. As with activators that mediate class I activation, EBPs bind to DNA upstream of promoter regions; however, as EBPs form interactions with the sigma factor — rather than the more proximal α-subunits — of the RNA polymerase holoenzyme, looping of upstream DNA has to occur to 'deliver' the EBP to the sigma factor, and this is often assisted by DNA-bending proteins such as IHF[112,113].

Promoter modifications. Other mechanisms by which the activity of bacterial promoters can be regulated involve the modification of specific bases or changes to the base sequence. These modifications can either alter the affinity of an operator for a transcription factor or modulate the binding or affinity of RNA polymerase. The most common chemical modification is DNA methylation, which can result in changes in the binding affinities of specific transcription factors to operators that can potentially lead to global changes in the transcriptional programme[114–118]. For example, methylation of the operators of two *E. coli* genes, *pap* and *agn43*, by the DNA adenine methylase (Dam) results in the loss of repressor binding, which leads to the induction of gene expression. Repression is re-established in daughter cells, as the DNA that is synthesized during replication is unmethylated (FIG. 5a).

The most extreme examples of base sequence changes at regulatory regions are inversions of DNA segments that reverse the orientation of the promoter in response to a signal. This creates promoters that are either 'on' or 'off', depending on whether they are oriented towards or away from the genes that they control, which contrasts with promoters that are directly controlled by transcription factors, for which output can be proportional to the input signal[117,118]. An example of regulation by DNA inversion is the promoter that controls the expression of the *fim* operon, which encodes type 1 fimbriae in *E. coli* (FIG. 5b). As might be expected, the FimB and FimE enzymes that catalyse DNA inversion at this operon are themselves tightly regulated.

Local sequence variation at promoters provides a more sophisticated alternative strategy to DNA inversion. This is usually driven by tracts of variable length that repeat a single nucleotide (or dinucleotide), often in the vicinity of the −35 element of the promoter (FIG. 5c). For example, the promoters of *hifA* and *hifB* in *Haemophilus influenzae* have variable numbers of TA repeats, which modulates the activity of the promoters. Individual bacterial cells in a population have a different number of repeats in the variable region of the promoter, with each tract length corresponding to a different level of transcriptional activity. Thus, at any moment, a subset of cells will have the optimal level of transcriptional activity at the promoter for the given conditions[117–120]. This regulation 'by lottery', which is driven by repetitive sequences that differ from one generation to another,



◀ Figure 5 | **Regulation by promoter DNA modification.** **a** | Regulation by DNA methylation. To repress transcription of *agn43*, the transcription factor OxyR binds to sites in the promoter region of the *agn43* locus that contain unmethylated GATC sequences. This results in transcriptional repression because OxyR blocks the access of RNA polymerase to the promoter. After DNA replication, the DNA adenine methylase (Dam) can methylate the adenines of these GATC sequences on one or both strands, which prevents OxyR binding. Once the GATC sites are no longer occupied by OxyR, RNA polymerase is able to bind to the promoter and initiate transcription. **b** | Regulation by DNA inversion. The expression of *fimA* and other genes of the *fim* operon is regulated by inversion of a DNA element in the promoter region. The inversion is mediated by the recombinases FimB and FimE and switches the orientation of the promoter between off and on states. In the 'off' state, the promoter is oriented away from *fimA*, which results in the production of non-coding transcripts, whereas, in the 'on' state, the orientation of the promoter enables the production of *fimA* transcripts. The *fim* locus is not drawn to scale. **c** | Regulation by local sequence variation. In *Haemophilus influenzae*, the promoters of *hifA* and *hifB*, which encode fimbrial components, have a variable number of TA-dinucleotide repeats, which alter the spacing between the −10 and −35 elements. As the spacing between these elements is a determinant of RNA polymerase holoenzyme binding, and thus promoter strength, variation of the number of repeats produces changes in gene expression, such that expression can be switched off entirely or adjusted to different levels. αCTD, carboxy-terminal domain of the α-subunit of RNA polymerase. αNTD, the amino-terminal domains of the α-subunits of RNA polymerase.

contrasts sharply with regulation by transcription factors, which drive specific responses to specific environmental cues. Therefore, transcription factors may not always be essential for bacterial survival, as other strategies can be used by bacteria to adapt their transcriptional programmes to changing environments.

Evolutionary origins

François Jacob stated that 'evolution was a tinkerer' and nowhere is this famous statement more evident than in the regulation of transcription initiation in bacteria[121]. From the perspective of evolutionary history, we can assume that transcription was present in the first cells and it is easy to imagine that this transcription might have been much less specific and much less regulated than in modern-day cells. Indeed, it has been suggested that the widely observed, unregulated, pervasive transcription that is seen in many bacteria is an evolutionary relic from the time before control was exerted[122]. Furthermore, the RNA polymerase core enzyme, which is common to all three domains of life, can make transcripts, but lacks the ability to select specific start sites or efficiently initiate transcription[123]. In bacteria, these functions were acquired through the evolution of sigma factors, whereas different mechanisms evolved to solve this problem in archaea and eukaryotes[123].

Following the acquisition of sigma factors, the bacterial RNA polymerase had to overcome the compacted DNA structure of the nucleoid to access the many potential promoters that were probably blocked by NAPs (BOX 3). In fact, NAPs may have been important in the evolution of transcription regulation, which has been argued to have involved the co-option of NAPs to regulate the repression or activation of specific genes[124]. Therefore, the reversal of NAP-mediated repression (such as that mediated by Ler, NarL or NarP) may well reflect the oldest form of activation, with the acquisition of activating regions that recruit RNA polymerase occurring later in evolution. Evidence for a common origin for transcription factors and NAPs comes from findings that DNA binding by transcription factors does not always produce a measurable effect on transcription[125,126], and that many NAPs can function as regulators of transcription by using activating regions to recruit RNA polymerase[127,128].

When considering the various regulatory mechanisms of sigma factors, transcription factors, appropriators and other factors, one can conclude that the complexity of the regulation of transcription initiation in bacteria has probably evolved through the sequential addition of layers of regulation. Furthermore, the evolution of these regulatory layers continues today, as can be

Pervasive transcription
The synthesis of transcripts that seem not to correspond to any functional genetic unit.

seen in the different extents to which different bacteria use sigma factors and transcription factors, which underscores the danger of making generalized conclusions based solely on studies using *E. coli* as a model.

Outlook

Since 2004, when we last reviewed this topic[1], enormous progress has been made in understanding the molecular details of transcription initiation, together with mechanisms that regulate its activation and repression. Fundamental to this progress has been an improved understanding of the structure and dynamics of the bacterial RNA polymerase, which has enabled further studies to elucidate the mechanisms of many of the factors that interact with the polymerase. In parallel, the availability of omics data made possible by the development of next-generation sequencing and other high-throughput methods has alerted us to the diversity of factors that regulate RNA polymerase in different bacteria and to the breadth of the regulatory repertoire. However, despite these advances, we are still unable to predict transcriptomes and regulatory patterns from first principles. Therefore, we believe that an important aspiration for the years ahead is to develop our understanding of regulatory mechanisms such that the many layers of regulation can be modelled to accurately predict transcriptional outcomes from a bacterial genome sequence.

1. Browning, D. F. & Busby, S. J. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65 (2004).
2. Feklistov, A., Sharon, B. D., Darst, S. A. & Gross, C. A. Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* **68**, 357–376 (2014).
3. Murakami, K. S. & Darst, S. A. Bacterial RNA polymerases: the wholo story. *Curr. Opin. Struct. Biol.* **13**, 31–39 (2003).
4. Feklistov, A. & Darst, S. A. Structural basis for promoter −10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* **147**, 1257–1269 (2011).
5. Zhang, Y. *et al.* Structural basis of transcription initiation. *Science* **338**, 1076–1080 (2012).
This work complements previous structural work detailed in reference 3 by revealing the interactions of bacterial holoenzyme containing σ^{70} with the downstream end of the initiation 'bubble'. The results introduce us to unstacking of certain bases and their insertion into pockets in the sigma factor, and to specific interactions between side chains in the core enzyme and certain bases near the transcript start (the core recognition element (CRE)).
6. Zuo, Y. & Steitz, T. A. Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol. Cell* **58**, 534–540 (2015).
7. Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R. & Darst, S. A. Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife* **4**, e08504 (2015).
8. Ruff, E. F., Record, M. T. & Artsimovitch, I. Initial events in bacterial transcription initiation. *Biomolecules* **5**, 1035–1062 (2015).
9. Washburn, R. S. & Gottesman, M. E. Regulation of transcription elongation and termination. *Biomolecules* **5**, 1063–1078 (2015).
10. Roberts, J. W., Shankar, S. & Filter, J. J. RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62**, 211–233 (2008).
11. Zhang, J. & Landick, R. A two-way street: regulatory interplay between RNA polymerase and nascent RNA structure. *Trends Biochem. Sci.* **41**, 293–310 (2016).
12. Campbell, E. A. *et al.* Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* **9**, 527–539 (2002).
13. Mekler, V. *et al.* Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase–promoter open complex. *Cell* **108**, 599–614 (2002).
14. Bae, B. *et al.* Phage T7 Gp2 inhibition of *Escherichia coli* RNA polymerase involves misappropriation of σ70 domain 1.1. *Proc. Natl Acad. Sci. USA* **110**, 19772–19777 (2013).
15. Murakami, K. S. X-ray crystal structure of *Escherichia coli* RNA polymerase σ70 holoenzyme. *J. Biol. Chem.* **288**, 9126–9134 (2013).
16. Gruber, T. M. & Gross, C. A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441–466 (2003).
17. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O. & Darst, S. A. Structural basis of transcription initiation: an RNA polymerase holoenzyme–DNA complex. *Science* **296**, 1285–1290 (2002).
18. Vassylyev, D. G. *et al.* Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**, 712–719 (2002).
19. Yang, Y. *et al.* Structures of the RNA polymerase–σ54 reveal new and conserved regulatory strategies. *Science* **349**, 882–885 (2015).
This paper presents the long awaited structure of σ^{54} and its interactions with RNA polymerase. The structure shows how σ^{54} really is different from σ^{70}, why it is incompetent for transcription initiation, and suggests how its activators might work.
20. Jishage, M. & Ishihama, A. Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of σ70 and σ38. *J. Bacteriol.* **177**, 6832–6835 (1995).
21. Jishage, M., Iwata, A., Ueda, S. & Ishihama, A. Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J. Bacteriol.* **178**, 5447–5451 (1996).
22. Maeda, H., Fujita, N. & Ishihama, A. Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. *Nucleic Acids Res.* **28**, 3497–3503 (2000).
23. Campbell, E. A., Westblade, L. F. & Darst, S. A. Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Curr. Opin. Microbiol.* **11**, 121–127 (2008).
24. Rhodius, V. A. *et al.* Design of orthogonal genetic switches based on a crosstalk map of sigmas, anti-sigmas, and promoters. *Mol. Syst. Biol.* **9**, 702 (2013).
25. Campagne, S., Marsh, M. E., Capitani, G., Vorholt, J. A. & Allain, F. H. Structural basis for −10 promoter element melting by environmentally induced sigma factors. *Nat. Struct. Mol. Biol.* **21**, 269–276 (2014).
26. Koo, B. M., Rhodius, V. A., Nonaka, G., deHaseth, P. L. & Gross, C. A. Reduced capacity of alternative sigmas to melt promoters ensures stringent promoter recognition. *Genes Dev.* **23**, 2426–2436 (2009).
27. Hollands, K., Lee, D. J., Lloyd, G. S. & Busby, S. J. Activation of σ28-dependent transcription in *Escherichia coli* by the cyclic AMP receptor protein requires an unusual promoter organization. *Mol. Microbiol.* **75**, 1098–1111 (2010).
28. Paget, M. S. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* **5**, 1245–1265 (2015).
29. Osterberg, S., del Peso-Santos, T. & Shingler, V. Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.* **65**, 37–55 (2011).
30. Francez-Charlot, A. *et al.* Sigma factor mimicry involved in regulation of general stress response. *Proc. Natl Acad. Sci. USA* **106**, 3467–3472 (2009).
Following the resolution of structures of sigma–anti-sigma complexes detailed in reference 23, this paper reports that an anti-anti-sigma factor can function by mimicking a sigma factor. Although predictable that this would be the case, its formal demonstration was a landmark for the field.
31. Herrou, J., Rotskoff, G., Luo, Y., Roux, B. & Crosson, S. Structural basis of a protein partner switch that regulates the general stress response of α-proteobacteria. *Proc. Natl Acad. Sci. USA* **109**, E1415–E1423 (2012).
32. Typas, A., Barembruch, C., Possling, A. & Hengge, R. Stationary phase reorganisation of the *Escherichia coli* transcription machinery by Crl protein, a fine-tuner of sigmas activity and levels. *EMBO J.* **26**, 1569–1578 (2007).
33. Banta, A. B. *et al.* Key features of σS required for specific recognition by Crl, a transcription factor promoting assembly of RNA polymerase holoenzyme. *Proc. Natl Acad. Sci. USA* **110**, 15955–15960 (2013).
34. Banta, A. B. *et al.* Structure of the RNA polymerase assembly factor Crl and identification of its interaction surface with σS. *J. Bacteriol.* **196**, 3279–3288 (2014).
35. Yuan, A. H. *et al.* Rsd family proteins make simultaneous interactions with regions 2 and 4 of the primary sigma factor. *Mol. Microbiol.* **70**, 1136–1151 (2008).
36. Piper, S. E., Mitchell, J. E., Lee, D. J. & Busby, S. J. A global view of *Escherichia coli* Rsd protein and its interactions. *Mol. Biosyst.* **5**, 1943–1947 (2009).
37. Sharma, U. K. & Chatterji, D. Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of σ^{70} activity. *FEMS Microbiol. Rev.* **34**, 646–657 (2010).
38. Jishage, M., Kvint, K., Shingler, V. & Nystrom, T. Regulation of sigma factor competition by the alarmone ppGpp. *Genes Dev.* **16**, 1260–1270 (2002).
39. Costanzo, A. *et al.* ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor σE in *Escherichia coli* by both direct and indirect mechanisms. *Mol. Microbiol.* **67**, 619–632 (2008).
40. Merrick, M. J. In a class of its own — the RNA polymerase sigma factor σ54 (σN). *Mol. Microbiol.* **10**, 903–909 (1993).
41. Studholme, D. J. & Buck, M. The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol. Lett.* **186**, 1–9 (2000).
42. Wigneshweraraj, S. *et al. Modus operandi* of the bacterial RNA polymerase containing the σ54 promoter-specificity factor. *Mol. Microbiol.* **68**, 538–546 (2008).
43. Flentie, K., Garner, A. L. & Stallings, C. L. The *Mycobacterium tuberculosis* transcription machinery: ready to respond to host attacks. *J. Bacteriol.* **198**, 1360–1373 (2016).
44. Hubin, E. A. *et al.* Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. *Proc. Natl Acad. Sci. USA* **112**, 7171–7176 (2015).
45. Srivastava, D. B. *et al.* Structure and function of CarD, an essential mycobacterial transcription factor. *Proc. Natl Acad. Sci. USA* **110**, 12619–12624 (2013).
46. Bae, B. *et al.* CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *eLife* **4**, e08505 (2015).
47. Perederina, A. *et al.* Regulation through the secondary channel — structural framework for ppGpp–DksA synergism during transcription. *Cell* **118**, 297–309 (2004).
48. Paul, B. J. *et al.* DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**, 311–322 (2004).
49. Paul, B. J., Berkmen, M. B. & Gourse, R. L. DksA potentiates direct activation of amino acid promoters by ppGpp. *Proc. Natl Acad. Sci. USA* **102**, 7823–7828 (2005).
In this study, an *in vitro* assay is used to provide the proof that ppGpp, together with DksA, really can directly stimulate transcription initiation at a promoter. This paper provides a wonderful lesson in how the limitations of genetics can be atoned for by amazing biochemistry.

50. Lennon, C. W. *et al.* Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation. *Genes Dev.* **26**, 2634–2646 (2012).
51. Ross, W., Vrentas, C. E., Sanchez-Vazquez, P., Gaal, T. & Gourse, R. L. The magic spot: a ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. *Mol. Cell* **50**, 420–429 (2013).
52. Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**, 298–309 (2015).
53. Zenkin, N. & Yuzenkova, Y. New insights into the functions of transcription factors that bind the RNA polymerase secondary channel. *Biomolecules* **5**, 1195–1209 (2015).
54. Yuzenkova, Y., Roghanian, M. & Zenkin, N. Multiple active centers of multi-subunit RNA polymerases. *Transcription* **3**, 115–118 (2012).
55. Friedman, L. J. & Gelles, J. Multi-wavelength single-molecule fluorescence analysis of transcription mechanisms. *Methods* **86**, 27–36 (2015).
56. Zhang, Y. *et al.* DksA guards elongating RNA polymerase against ribosome-stalling-induced arrest. *Mol. Cell* **53**, 766–778 (2014).
57. Cavanagh, A. T. & Wassarman, K. M. 6S RNA, a global regulator of transcription in *Escherichia coli*. *Bacillus subtilis*, and beyond. *Annu. Rev. Microbiol.* **68**, 45–60 (2014).
58. Liu, B. *et al.* A bacteriophage transcription regulator inhibits bacterial transcription initiation by sigma-factor displacement. *Nucleic Acids Res.* **42**, 4294–4305 (2014).
59. Lambert, L. J., Wei, Y., Schirf, V., Demeler, B. & Werner, M. H. T4 AsiA blocks DNA recognition by remodeling σ70 region 4. *EMBO J.* **23**, 2952–2962 (2004).
60. Gregory, B. D. *et al.* A regulator that inhibits transcription by targeting an intersubunit interaction of the RNA polymerase holoenzyme. *Proc. Natl Acad. Sci. USA* **101**, 4554–4559 (2004).
61. Hinton, D. M. Transcriptional control in the prereplicative phase of T4 development. *Virol. J.* **7**, 289 (2010).
62. Griffith, K. L., Shah, I. M., Myers, T. E., O'Neill, M. C. & Wolf, R. E. Evidence for "pre-recruitment" as a new mechanism of transcription activation in *Escherichia coli*: the large excess of SoxS binding sites per cell relative to the number of SoxS molecules per cell. *Biochem. Biophys. Res. Commun.* **291**, 979–986 (2002).
63. Shah, I. M. & Wolf, R. E. Novel protein–protein interaction between *Escherichia coli* SoxS and the DNA binding determinant of the RNA polymerase α-subunit: SoxS functions as a co-sigma factor and redeploys RNA polymerase from UP-element-containing promoters to SoxS-dependent promoters during oxidative stress. *J. Mol. Biol.* **343**, 513–532 (2004).
64. Zuber, P. Management of oxidative stress in *Bacillus*. *Annu. Rev. Microbiol.* **63**, 575–597 (2009).
65. Newberry, K. J., Nakano, S., Zuber, P. & Brennan, R. G. Crystal structure of the *Bacillus subtilis* anti-alpha, global transcriptional regulator, Spx, in complex with the α C-terminal domain of RNA polymerase. *Proc. Natl Acad. Sci. USA* **102**, 15839–15844 (2005).
66. Lamour, V., Westblade, L. F., Campbell, E. A. & Darst, S. A. Crystal structure of the *in vivo*-assembled *Bacillus subtilis* Spx/RNA polymerase α subunit C-terminal domain complex. *J. Struct. Biol.* **168**, 352–356 (2009).
67. Mangel, W. F. & Chamberlin, M. J. Studies of ribonucleic acid chain initiation by *Escherichia coli* ribonucleic acid polymerase bound to T7 deoxyribonucleic acid. I. An assay for the rate and extent of ribonucleic acid chain initiation. *J. Biol. Chem.* **249**, 2995–3001 (1974).
68. Murray, H. D., Schneider, D. A. & Gourse, R. L. Control of rRNA expression by small molecules is dynamic and nonredundant. *Mol. Cell* **12**, 125–134 (2003).
69. Schneider, D. A., Gaal, T. & Gourse, R. L. NTP-sensing by rRNA promoters in *Escherichia coli* is direct. *Proc. Natl Acad. Sci. USA* **99**, 8602–8607 (2002).
70. Krasny, L. & Gourse, R. L. An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. *EMBO J.* **23**, 4473–4483 (2004).
71. Liu, K., Bittner, A. N. & Wang, J. D. Diversity in (p)ppGpp metabolism and effectors. *Curr. Opin. Microbiol.* **24**, 72–79 (2015).
This study provides a very clear account of how ppGpp (and pppGpp) has different roles in different bacteria. A warning that what is true for *E. coli* may not be true for other bacteria, let alone elephants.
72. Turnbough, C. L. & Switzer, R. L. Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. *Microbiol. Mol. Biol. Rev.* **72**, 266–300 (2008).
73. Martinez-Antonio, A. & Collado-Vides, J. Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol.* **6**, 482–489 (2003).
74. Ishihama, A. Prokaryotic genome regulation: a revolutionary paradigm. *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* **88**, 485–508 (2012).
75. Cho, B. K., Palsson, B. & Zengler, K. Deciphering the regulatory codes in bacterial genomes. *Biotechnol. J.* **6**, 1052–1063 (2011).
76. Salgado, H. *et al.* Extracting regulatory networks of *Escherichia coli* from RegulonDB. *Methods Mol. Biol.* **804**, 179–195 (2012).
77. Dillon, S. C. & Dorman, C. J. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat. Rev. Microbiol.* **8**, 185–195 (2010).
78. Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* **181**, 6361–6370 (1999).
79. Azam, T. A. & Ishihama, A. Twelve species of the nucleoid-associated protein from *Escherichia coli*. Sequence recognition specificity and DNA binding affinity. *J. Biol. Chem.* **274**, 33105–33113 (1999).
80. Luijsterburg, M. S., Noom, M. C., Wuite, G. J. & Dame, R. T. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J. Struct. Biol.* **156**, 262–272 (2006).
81. Browning, D. F., Grainger, D. C. & Busby, S. J. Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr. Opin. Microbiol.* **13**, 773–780 (2010).
82. Grainger, D. C., Goldberg, M. D., Lee, D. J. & Busby, S. J. Selective repression by Fis and H-NS at the *Escherichia coli dps* promoter. *Mol. Microbiol.* **68**, 1366–1377 (2008).
83. Sobetzko, P., Glinkowska, M., Travers, A. & Muskhelishvili, G. DNA thermodynamic stability and supercoil dynamics determine the gene expression program during the bacterial growth cycle. *Mol. Biosyst.* **9**, 1643–1651 (2013).
84. Sobetzko, P., Travers, A. & Muskhelishvili, G. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *Proc. Natl Acad. Sci. USA* **109**, E42–E50 (2012).
85. Dorman, C. J. Co-operative roles for DNA supercoiling and nucleoid-associated proteins in the regulation of bacterial transcription. *Biochem. Soc. Trans.* **41**, 542–547 (2013).
86. Zhang, W. & Baseman, J. B. Transcriptional regulation of MG_149, an osmoinducible lipoprotein gene from *Mycoplasma genitalium*. *Mol. Microbiol.* **81**, 327–339 (2011).
87. Neumann, S. & Quinones, A. Discoordinate gene expression of *gyrA* and *gyrB* in response to DNA gyrase inhibition in *Escherichia coli*. *J. Basic Microbiol.* **37**, 53–69 (1997).
88. Lal, A. *et al.* Genome scale patterns of supercoiling in a bacterial chromosome. *Nat. Commun.* **7**, 11055 (2016).
89. Bryant, J. A., Sellars, L. E., Busby, S. J. & Lee, D. J. Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Res.* **42**, 11383–11392 (2014).
90. Swint-Kruse, L. & Matthews, K. S. Allostery in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.* **12**, 129–137 (2009).
91. Semsey, S., Tolstorukov, M. Y., Virnik, K., Zhurkin, V. B. & Adhya, S. DNA trajectory in the Gal repressosome. *Genes Dev.* **18**, 1898–1907 (2004).
92. Butala, M. *et al.* Double locking of an *Escherichia coli* promoter by two repressors prevents premature colicin expression and cell lysis. *Mol. Microbiol.* **86**, 129–139 (2012).
93. Kamensek, S. *et al.* Silencing of DNase colicin E8 gene expression by a complex nucleoprotein assembly ensures timely colicin induction. *PLoS Genet.* **11**, e1005354 (2015).
94. Valentin-Hansen, P., Sogaard-Andersen, L. & Pedersen, H. A flexible partnership: the CytR anti-activator and the cAMP–CRP activator protein, comrades in transcription control. *Mol. Microbiol.* **20**, 461–466 (1996).
95. Monsalve, M., Mencia, M., Salas, M. & Rojo, F. Protein p4 represses phage Φ29 A2c promoter by interacting with the α subunit of *Bacillus subtilis* RNA polymerase. *Proc. Natl Acad. Sci. USA* **93**, 8913–8918 (1996).
96. Lee, D. J., Minchin, S. D. & Busby, S. J. Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**, 125–152 (2012).
97. Benoff, B. *et al.* Structural basis of transcription activation: the CAP–αCTD–DNA complex. *Science* **297**, 1562–1566 (2002).
98. Gaston, K., Bell, A., Kolb, A., Buc, H. & Busby, S. Stringent spacing requirements for transcription activation by CRP. *Cell* **62**, 733–743 (1990).
99. Zhou, Y., Kolb, A., Busby, S. J. & Wang, Y. P. Spacing requirements for class I transcription activation in bacteria are set by promoter elements. *Nucleic Acids Res.* **42**, 9209–9216 (2014).
100. Dove, S. L., Huang, F. W. & Hochschild, A. Mechanism for a transcriptional activator that works at the isomerization step. *Proc. Natl Acad. Sci. USA* **97**, 13215–13220 (2000).
101. Jain, D., Nickels, B. E., Sun, L., Hochschild, A. & Darst, S. A. Structure of a ternary transcription activation complex. *Mol. Cell* **13**, 45–53 (2004).
102. Niu, W., Kim, Y., Tau, G., Heyduk, T. & Ebright, R. H. Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. *Cell* **87**, 1123–1134 (1996).
103. Feng, Y., Zhang, Y. & Ebright, R. H. Structural basis of transcription activation. *Science* **352**, 1330–1333 (2016).
This paper describes the first full molecular picture of class II transcription activation at a target promoter, and contrasts with previous reports that have described components of the process.
104. Brown, N. L., Stoyanov, J. V., Kidd, S. P. & Hobman, J. L. The MerR family of transcriptional regulators. *FEMS Microbiol. Rev.* **27**, 145–163 (2003).
105. Philips, S. J. *et al.* Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science* **349**, 877–881 (2015).
Working with CueR, which senses copper and silver ions, this paper presents high-resolution structures of its complexes at a target sequence when in either activator or repressor mode. The results complement previous work detailed in reference 106 showing that DNA distortion by such factors is uneven.
106. Heldwein, E. E. & Brennan, R. G. Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**, 378–382 (2001).
107. Bustamante, V. H., Santana, F. J., Calva, E. & Puente, J. L. Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. *Mol. Microbiol.* **39**, 664–678 (2001).
108. Sperandio, V. *et al.* Activation of enteropathogenic *Escherichia coli* (EPEC) *LEE2* and *LEE3* operons by Ler. *Mol. Microbiol.* **38**, 781–793 (2000).
109. Browning, D. F., Cole, J. A. & Busby, S. J. Transcription activation by remodelling of a nucleoprotein assembly: the role of NarL at the FNR-dependent *Escherichia coli nir* promoter. *Mol. Microbiol.* **53**, 203–215 (2004).
110. Browning, D. F., Cole, J. A. & Busby, S. J. Regulation by nucleoid-associated proteins at the *Escherichia coli nir* operon promoter. *J. Bacteriol.* **190**, 7258–7267 (2008).
111. Tyson, K. L., Cole, J. A. & Busby, S. J. Nitrite and nitrate regulation at the promoters of two *Escherichia coli* operons encoding nitrite reductase: identification of common target heptamers for both NarP- and NarL-dependent regulation. *Mol. Microbiol.* **13**, 1045–1055 (1994).
112. Bush, M. & Dixon, R. The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. *Microbiol. Mol. Biol. Rev.* **76**, 497–529 (2012).
113. Buck, M. *et al.* A second paradigm for gene activation in bacteria. *Biochem. Soc. Trans.* **34**, 1067–1071 (2006).
114. Casadesus, J. & Low, D. Epigenetic gene regulation in the bacterial world. *Microbiol. Mol. Biol. Rev.* **70**, 830–856 (2006).
115. van der Woude, M. W. & Henderson, I. R. Regulation and function of Ag43 (flu). *Annu. Rev. Microbiol.* **62**, 153–169 (2008).

116. Sanchez-Romero, M. A., Cota, I. & Casadesus, J. DNA methylation in bacteria: from the methyl group to the methylome. *Curr. Opin. Microbiol.* **25**, 9–16 (2015).
117. van der Woude, M. W. & Baumler, A. J. Phase and antigenic variation in bacteria. *Clin. Microbiol. Rev.* **17**, 581–611 (2004).
118. van der Woude, M. W. Phase variation: how to create and coordinate population diversity. *Curr. Opin. Microbiol.* **14**, 205–211 (2011).
119. Cerdeno-Tarraga, A. M. *et al.* Extensive DNA inversions in the *B. fragilis* genome control variable gene expression. *Science* **307**, 1463–1465 (2005).
120. Moxon, R., Bayliss, C. & Hood, D. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu. Rev. Genet.* **40**, 307–333 (2006).
121. Jacob, F. Evolution and tinkering. *Science* **196**, 1161–1166 (1977).
122. Wade, J. T. & Grainger, D. C. Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat. Rev. Microbiol.* **12**, 647–653 (2014).
123. Grohmann, D. & Werner, F. Recent advances in the understanding of archaeal transcription. *Curr. Opin. Microbiol.* **14**, 328–334 (2011).
124. Visweswariah, S. S. & Busby, S. J. Evolution of bacterial transcription factors: how proteins take on new tasks, but do not always stop doing the old ones. *Trends Microbiol.* **23**, 463–467 (2015).
125. Grainger, D. C., Hurd, D., Harrison, M., Holdstock, J. & Busby, S. J. Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc. Natl Acad. Sci. USA* **102**, 17693–17698 (2005).
126. Shimada, T., Ishihama, A., Busby, S. J. & Grainger, D. C. The *Escherichia coli* RutR transcription factor binds at targets within genes as well as intergenic regions. *Nucleic Acids Res.* **36**, 3950–3955 (2008).
127. Aiyar, S. E. *et al.* Architecture of Fis-activated transcription complexes at the *Escherichia coli rrnB* P1 and *rrnE* P1 promoters. *J. Mol. Biol.* **316**, 501–516 (2002).
128. Rossiter, A. E. *et al.* Expression of different bacterial cytotoxins is controlled by two global transcription factors, CRP and Fis, that co-operate in a shared-recruitment mechanism. *Biochem. J.* **466**, 323–335 (2015).
129. Lisser, S. & Margalit, H. Compilation of *E. coli* mRNA promoter sequences. *Nucleic Acids Res.* **21**, 1507–1516 (1993).
130. Hook-Barnard, I. G. & Hinton, D. M. Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul. Syst. Bio* **1**, 275–293 (2007).
131. Ross, W., Ernst, A. & Gourse, R. L. Fine structure of *E. coli* RNA polymerase–promoter interactions: α subunit binding to the UP element minor groove. *Genes Dev.* **15**, 491–506 (2001).
132. Sclavi, B. *et al.* Real-time characterization of intermediates in the pathway to open complex formation by *Escherichia coli* RNA polymerase at the T7A1 promoter. *Proc. Natl Acad. Sci. USA* **102**, 4706–4711 (2005).
133. Davis, C. A., Bingman, C. A., Landick, R., Record, M. T. & Saecker, R. M. Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA* **104**, 7833–7838 (2007).
134. Buckle, M., Pemberton, I. K., Jacquet, M. A. & Buc, H. The kinetics of sigma subunit directed promoter recognition by *E. coli* RNA polymerase. *J. Mol. Biol.* **285**, 955–964 (1999).
135. Stracy, M. *et al.* Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid. *Proc. Natl Acad. Sci. USA* **112**, E4390–E4399 (2015).
This study shows that live microscopy can be used to observe individual molecules of RNA polymerase in *E. coli*. This is a prerequisite for future direct observation of responses to regulatory triggers.
136. Patrick, M., Dennis, P. P., Ehrenberg, M. & Bremer, H. Free RNA polymerase in *E. coli*. *Biochimie* **119**, 80–91 (2015).
137. Hsu, L. M. Promoter escape by *Escherichia coli* RNA polymerase. *EcoSal Plus* http://dx.doi.org/10.1128/ecosalplus.4.5.2.2 (2008).
138. Skancke, J., Bar, N., Kuiper, M. & Hsu, L. M. Sequence-dependent promoter escape efficiency is strongly influenced by bias for the pretranslocated state during initial transcription. *Biochemistry* **54**, 4267–4275 (2015).
139. Bauer, D. L. V., Duchi, D. & Kapanidis, A. N. *E. coli* RNA polymerase pauses during initial transcription. *Biophys. J.* **110** (Suppl. 1), 21a (2016).
140. Reppas, N. B., Wade, J. T., Church, G. M. & Struhl, K. The transition between transcriptional initiation and elongation in *E. coli* is highly variable and often rate limiting. *Mol. Cell* **24**, 747–757 (2006).
141. Sendy, B., Lee, D. J., Busby, S. J. & Bryant, J. A. RNA polymerase supply and flux through the lac operon in *Escherichia coli*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* (in the press).
142. Jacob, F. *La Statue Intérieure* (in French) (Éditions Odile Jacob, 1987).
143. Müller-Hill, B. *The lac Operon. A Short History of a Genetic Paradigm* (Walter de Gruyter, 1996).
144. Schwartz, M. in *Origins of Molecular Biology. A Tribute to Jacques Monod* (eds Lwoff, A. & Ullmann, A.) 207–216 (Academic Press, 1979).
145. Cases, I. & de Lorenzo, V. The black cat/white cat principle of signal integration in bacterial promoters. *EMBO J.* **20**, 1–11 (2001).
146. Pul, Ü. & Wagner, R. in *Bacterial Chromatin* (eds Dame, R. T. & Dorman, C. J.) 149–173 (Springer, 2010).
147. Dorman, C. J. H-NS: a universal regulator for a dynamic genome. *Nat. Rev. Microbiol.* **2**, 391–400 (2004).
148. Stoebel, D. M., Free, A. & Dorman, C. J. Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154**, 2533–2545 (2008).
149. Ohniwa, R. L. *et al.* Dynamic state of DNA topology is essential for genome condensation in bacteria. *EMBO J.* **25**, 5591–5602 (2006).
150. Micka, B. & Marahiel, M. A. The DNA-binding protein HBsu is essential for normal growth and development in *Bacillus subtilis*. *Biochimie* **74**, 641–650 (1992).
151. Grove, A. Functional evolution of bacterial histone-like HU proteins. *Curr. Issues Mol. Biol.* **13**, 1–12 (2011).
152. Ball, C. A., Osuna, R., Ferguson, K. C. & Johnson, R. C. Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J. Bacteriol.* **174**, 8043–8056 (1992).
153. Dame, R. T., Noom, M. C. & Wuite, G. J. Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**, 387–390 (2006).
154. Dame, R. T., Wyman, C. & Goosen, N. H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.* **28**, 3504–3510 (2000).
155. Maurer, S., Fritz, J. & Muskhelishvili, G. A systematic *in vitro* study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. *J. Mol. Biol.* **387**, 1261–1276 (2009).
156. Grainger, D. C., Hurd, D., Goldberg, M. D. & Busby, S. J. Association of nucleoid proteins with coding and non-coding segments of the *Escherichia coli* genome. *Nucleic Acids Res.* **34**, 4642–4652 (2006).
157. Kahramanoglou, C. *et al.* Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Res.* **39**, 2073–2091 (2011).
158. Cho, B. K., Knight, E. M., Barrett, C. L. & Palsson, B. O. Genome-wide analysis of Fis binding in *Escherichia coli* indicates a causative role for A-/AT-tracts. *Genome Res.* **18**, 900–910 (2008).
159. Singh, S. S. *et al.* Widespread suppression of intragenic transcription initiation by H-NS. *Genes Dev.* **28**, 214–219 (2014).
160. Vora, T., Hottes, A. K. & Tavazoie, S. Protein occupancy landscape of a bacterial genome. *Mol. Cell* **35**, 247–253 (2009).
A novel application of genomics to monitor the protein landscape in different parts of a bacterial chromosome, with direct observation of tracts of the genome where transcription is silenced.
161. Haugen, S. P., Ross, W. & Gourse, R. L. Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat. Rev. Microbiol.* **6**, 507–519 (2008).

Acknowledgements

The authors were supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC; grant BB/J006076/1) and by the Industrial Biotechnology Catalyst programme (funded by Innovate UK, the BBSRC and the UK Engineering and Physical Sciences Research Council (EPSRC)) to support the translation, development and commercialisation of innovative industrial biotechnology processes.

Competing interests statement

The authors declare no competing interests.