



# Diverse and unified mechanisms of transcription initiation in bacteria

James Chen , Hande Boyaci and Elizabeth A. Campbell

**Abstract** | Transcription of DNA is a fundamental process in all cellular organisms. The enzyme responsible for transcription, RNA polymerase, is conserved in general architecture and catalytic function across the three domains of life. Diverse mechanisms are used among and within the different branches to regulate transcription initiation. Mechanistic studies of transcription initiation in bacteria are especially amenable because the promoter recognition and melting steps are much less complicated than in eukaryotes or archaea. Also, bacteria have critical roles in human health as pathogens and commensals, and the bacterial RNA polymerase is a proven target for antibiotics. Recent biophysical studies of RNA polymerases and their inhibition, as well as transcription initiation and transcription factors, have detailed the mechanisms of transcription initiation in phylogenetically diverse bacteria, inspiring this Review to examine unifying and diverse themes in this process.

## Initiation factors

Factors that coordinate with RNA polymerase to facilitate the start of the transcription cycle.

## Promoter melting

Process in which duplex promoter DNA is unwound during transcription initiation.

Transcription, the act of transcribing DNA into RNA, is one of the key processes in the central dogma of molecular biology and is thus executed in all cellular organisms<sup>1</sup> (FIG. 1). The enzyme responsible for all cellular transcription, the DNA-dependent RNA polymerase (RNAP), is composed of multiple subunits<sup>2</sup>. The subunit compositions of RNAPs from the three domains of life, Eukarya, Archaea and Bacteria, differ substantially; yet a core set of subunits share sequence similarity and are conserved in general architecture<sup>3</sup>. This concordance in structure reflects the universal functions of the enzymes. In general, cellular multisubunit RNAPs locate promoters, which are sequences that direct the enzyme to the beginning of genes<sup>4</sup>. After promoter binding, the enzyme melts (that is, opens and unwinds) the DNA, catalyses templated de novo polymerization of ribonucleotides (rNTPs), and then transitions from an initiation complex to an elongation complex until transcription of the gene is completed and the process terminated (FIG. 1a). These minimum steps, which are subject to complex regulation, are performed by all cellular DNA-dependent RNAPs.

Although the general transcription cycle and the overall structures of RNAPs are shared, details of the initiation, elongation and termination steps differ among and within the domains of life<sup>2</sup>. This is exemplified by the evolutionary diversity of initiation factors<sup>3,5,6</sup>. In bacteria, a primary step for the regulation of gene expression is transcription initiation.

In this Review, we first delineate the universal features and mechanisms of transcription initiation in bacteria, including a discussion of recent structural studies that have shed light on the intricate RNAP motions involved

in promoter melting. We then present newly described specialized mechanisms that regulate initiation, which resulted from recent research that has examined transcriptional systems beyond the historical model bacteria, *Escherichia coli* and *Bacillus subtilis*. Research of evolutionarily diverse clades, specifically Alphaproteobacteria, Gammaproteobacteria and Actinobacteria, has provided insights into similarities and differences in both the basal steps and the regulated steps of transcription initiation. These recent discoveries have made it abundantly clear that bacterial transcription mechanisms are diverse, which causes differences in susceptibility to antibiotics that inhibit transcription. We review the mechanisms of those RNAP inhibitors that have recently been evaluated across clades of bacteria. We hope to impart how studying RNAPs, and their regulators from evolutionarily divergent bacteria, informs on universal pathways and yet reveals a vast range of regulatory mechanisms. This Review also stresses the essentiality of research on transcriptional systems from pathogens for antibiotic development and optimization.

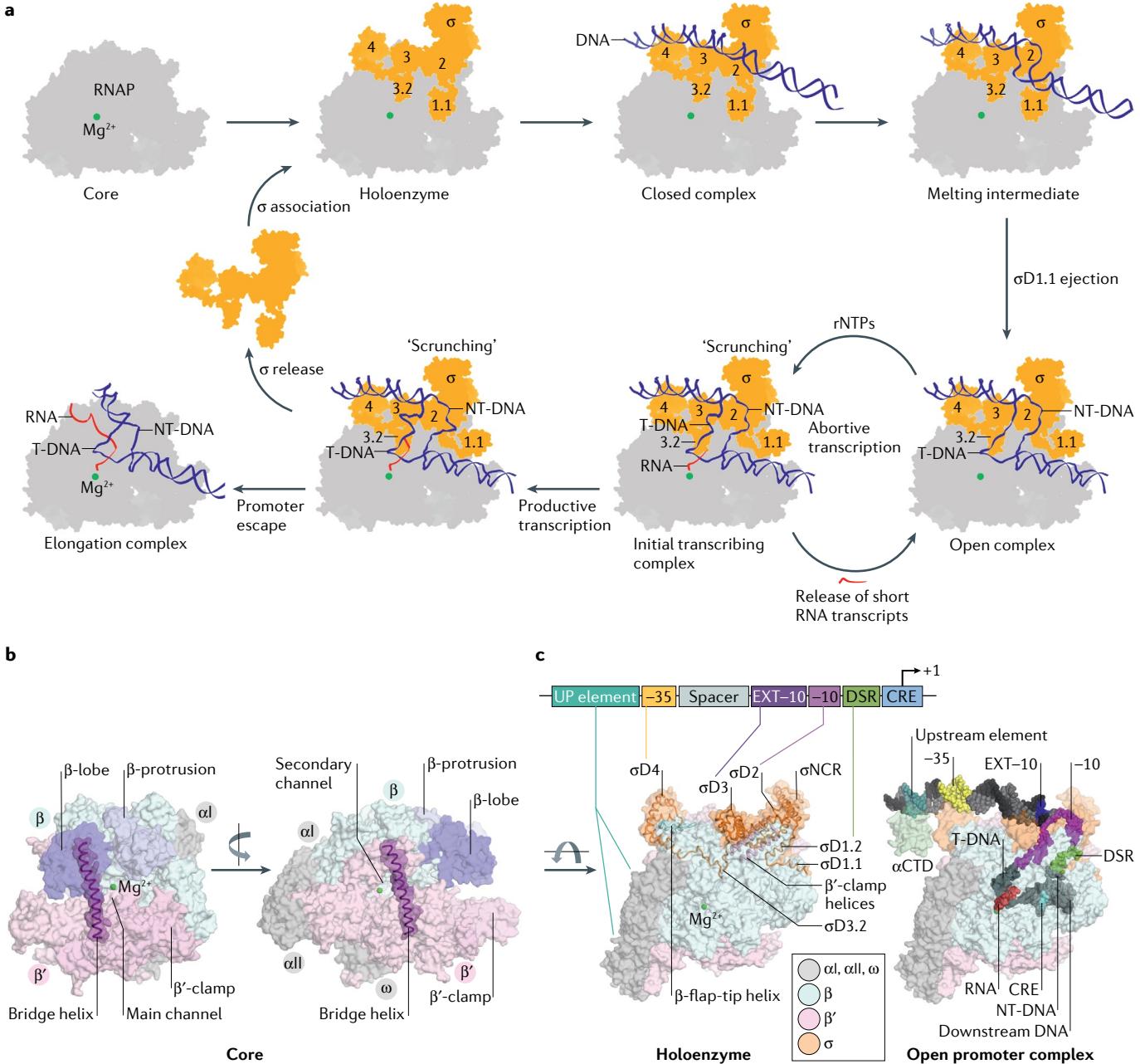
## Universal themes in bacterial transcription

**Architecture of RNAP, the initiation factor  $\sigma$  and promoters.** In bacteria, transcription is performed by a single RNAP. Bacterial RNAPs consist of an evolutionarily conserved catalytic core with the polypeptide subunit composition  $\alpha 2\beta\beta'\omega$ <sup>7</sup> (FIGS 1b,2). The core enzyme has several modules that are conserved in all cellular RNAPs, and some of the salient features are highlighted in FIG. 1b. The crystal structures of the bacterial, yeast and archaeal core RNAPs identified features that are characteristic of

Laboratory of Molecular Biophysics, The Rockefeller University, New York, NY, USA.

✉ e-mail: elizabeth.campbell@rockefeller.edu

<https://doi.org/10.1038/s41579-020-00450-2>



**Fig. 1 | Overview of the steps of bacterial transcription initiation and RNA polymerase. a** Schematic of bacterial transcription initiation. The cartoon shows the steps of bacterial transcription initiation, starting from core RNA polymerase (RNAP) to the elongation complex. The core RNAP (coloured in grey with the active site  $Mg^{2+}$  depicted in green) binds a promoter specificity factor,  $\sigma$  (orange), generating the holoenzyme. The holoenzyme recognizes and melts promoter DNA (blue) to form the open promoter complex. In the presence of ribonucleotides (rNTPs), the initiation complex scrunches and can synthesize many RNA transcripts (red) in a process called 'abortive transcription'. Eventually, the complex produces a transcript length that competes with  $\sigma$ , which in combination with scrunching causes  $\sigma$  to disassociate from the core RNAP, which leads to promoter escape. The enzyme then transitions to the elongation complex and completes transcription of the gene. **b** Architecture of the core RNAP, highlighting the active site  $Mg^{2+}$ , the main channel, the bridge helix (in cartoon tubes), the  $\beta$ -lobe and the  $\beta'$ -clamp and the  $\beta$ -protrusion and the  $\beta'$ -clamp. **c** Holoenzyme configuration and

recognition of housekeeping promoter. In the left panels, the holoenzyme is shown as a transparent molecular surface and coloured according to the colour key, with  $\sigma$ , the  $\beta$ -flap-tip helix and the clamp helices shown as cartoon tubes. The regions of the holoenzyme that bind the promoter elements are highlighted:  $\alpha$ -carboxy-terminal domains ( $\alpha$ CTDs; pale green) bind the upstream (UP) element (teal),  $\sigma$ -domain 4 ( $\sigma$ D4) binds the  $-35$  element (yellow),  $\sigma$ D3.2 is buried in the RNAP active site cleft,  $\sigma$ D3 interacts with the extended (EXT)- $-10$  element (magenta), the  $\sigma$ -non-conserved region ( $\sigma$ NCR) is located between  $\sigma$ D2 and  $\sigma$ D1.2,  $\sigma$ D1.2 interacts with the discriminator region (DSR; forest green) and the  $\beta$ -subunit of RNAP recognizes the core-recognition element (CRE; light blue). The transcription start site is shown as '+1', and the direction of transcription is depicted by the black arrow below the transcription start site. The structure of the open promoter complex primed for initiation showing the promoter DNA placed into RNAP with an RNA primer is shown in the right panel. NT-DNA, non-template DNA; T-DNA, template DNA.

all cellular RNAPs<sup>8–10</sup>, including the two mobile pincers that make up the active site cleft for DNA loading. The  $\beta'$ -clamp forms one pincer, while the  $\beta$ -lobe and the  $\beta$ -protrusion form the second pincer. The active site is marked by the required catalytic Mg<sup>2+</sup> and the bridge helix that is important for the addition of nucleotides. The bridge helix divides the active site cleft, thus creating a secondary smaller channel where incoming rNTPs are thought to enter<sup>8</sup>.

Although the catalytic core is capable of RNA polymerization, bacterial RNAPs require an initiation factor,  $\sigma$ , for promoter-specific DNA binding and unwinding<sup>11</sup> (FIG. 3). All bacteria possess a primary housekeeping  $\sigma$ -factor that controls the transcription of essential genes during normal growth conditions. Most transcription initiation events in bacteria involve RNAP bound to the primary  $\sigma$ -factor<sup>12,13</sup>. The domain architecture of primary  $\sigma$ -factors consists of three highly conserved domains ( $\sigma$ -domain 2 ( $\sigma$ D2),  $\sigma$ D3 and  $\sigma$ D4), a non-conserved region ( $\sigma$ NCR) and a poorly conserved amino-terminal domain ( $\sigma$ D1.1) that are connected by flexible linkers<sup>14</sup>. Binding of  $\sigma$  to core RNAP leads to the formation of the RNAP holoenzyme (via  $\sigma$ D4 interaction with the  $\beta$ -flap-tip helix and  $\sigma$ D2 interaction with the  $\beta'$ -clamp helices) in which the domains of  $\sigma$  are extended and spatially reorganized to engage with DNA<sup>15–17</sup> (illustrated in FIG. 1c, left). This Review focuses on transcription initiation by the  $\sigma^{70}$  family of  $\sigma$ -factors and will not discuss initiation by the  $\sigma^{54}$  family, which was elegantly and recently reviewed<sup>18</sup>. The  $\sigma^{54}$  family requires additional activator proteins and initiates transcription using unique mechanisms that are beyond the scope of this Review.  $\sigma$ -Factors guide the core RNAP to DNA promoters by providing specific contacts with the DNA<sup>14,16</sup>. Bacterial promoters that are recognized by the primary  $\sigma$ -factor are generally composed of two sequence motifs that are located upstream of the transcription start site: the –35 DNA element (consensus sequence TTGACA) and the –10 DNA element (consensus sequence TATAAT)<sup>19,20</sup>. Certain promoters also contain additional DNA elements such as the upstream element<sup>21</sup>, the extended –10 element<sup>22</sup>, the discriminator<sup>23</sup> and/or a core-recognition element<sup>24</sup>.

During promoter binding, the upstream element (located 40–60 bases upstream of the transcription start site) is recognized by the carboxy-terminal domains of the RNAP  $\alpha$ -subunits<sup>21</sup>. The –35 element is bound to  $\sigma$ D4 (REF.<sup>14</sup>), and unwinding occurs at the start of the –10 element, which interacts with  $\sigma$ D2 (REF.<sup>25</sup>). If present, the extended –10 element (TG at positions –14 and –13) is recognized by  $\sigma$ D3. The open complex is stabilized by interactions in the discriminator region (between position –6 and position –3), which contacts  $\sigma$ D1.2 (REF.<sup>23</sup>). Further stabilizing contacts include the core-recognition element (located at position +2), which binds in a pocket of the  $\beta$ -subunit<sup>24</sup>. The combination of these DNA elements and their interactions with the holoenzyme modulate the strength of a promoter by contributing to the initial RNAP binding and subsequent steps that lead to the formation and stability of the open promoter complex. FIGURE 1c highlights key features of promoters and how they interact with

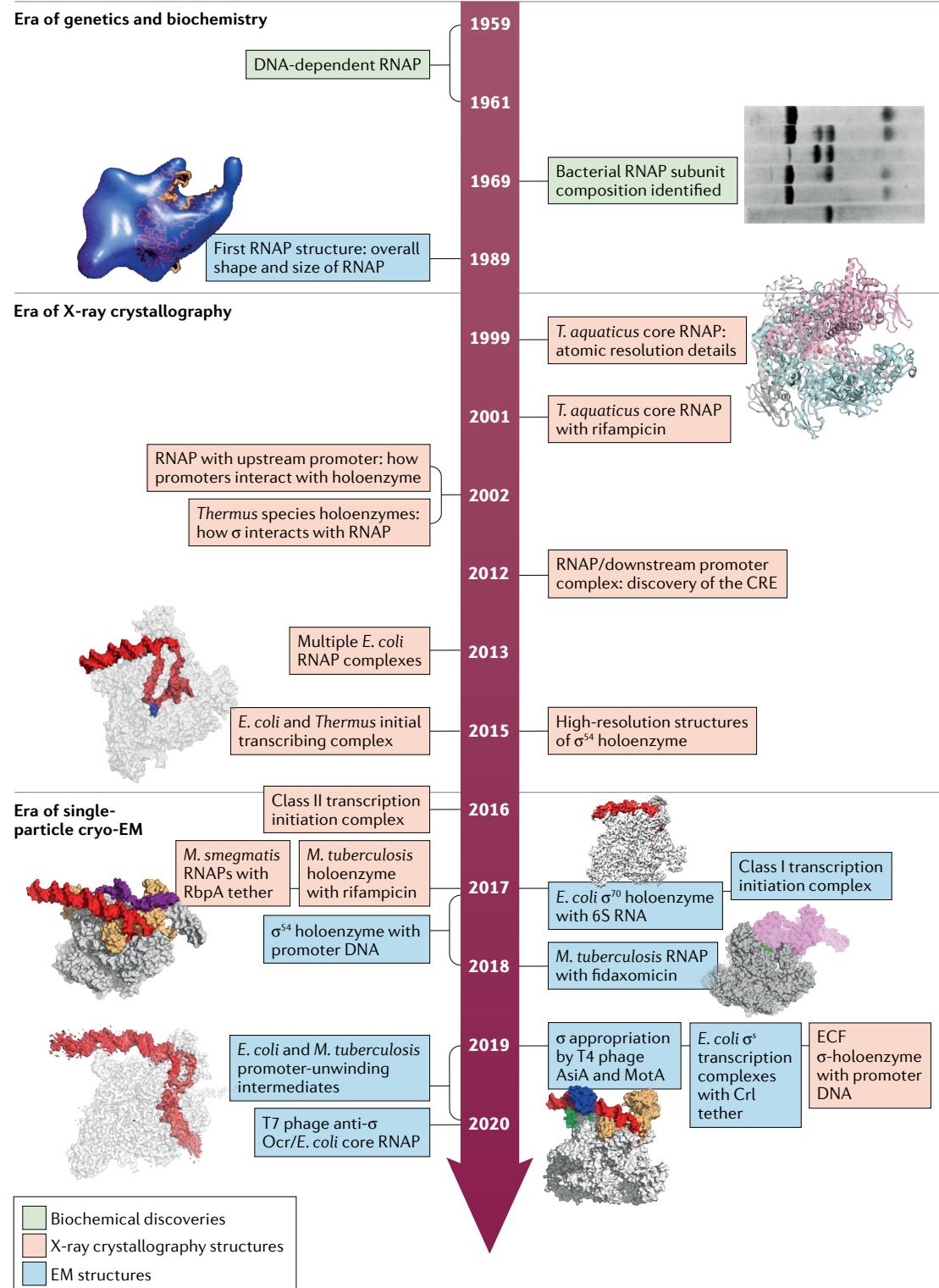
the RNAP holoenzyme in the final open complex. Most of these promoter elements and their interactions have been confirmed structurally and functionally for several clades of bacteria, including *Thermus* species, Gammaproteobacteria and Actinobacteria<sup>26–29</sup>. However, genomic studies suggest that the prevalence of each promoter element in different clades may differ considerably, such as the heavier reliance on extended –10 motifs compared with –35 elements in Gram-positive bacteria<sup>26,30,31</sup>.

**Promoter melting.** RNAP resembles a crab claw that opens and closes during transcription<sup>32,33</sup>. To efficiently initiate promoter melting, RNAP closes transiently to nucleate melting of the –10 element, then opens for DNA loading in the channel<sup>34</sup> (FIG. 4). After melting is completed, RNAP has to close to secure the DNA and ensure processive transcription.

Until recently, the details of how RNAP unwinds promoter DNA while forming the transcriptionally competent open complex were mostly unknown. Biochemical studies laid the foundation for our understanding of promoter melting, and included several observations. First, bacterial holoenzymes use binding free energy to isomerize into the fully melted, transcriptionally competent, open promoter complex (that is, the open complex)<sup>35,36</sup>. Second, this is a multistep process that is characterized by several RNAP–promoter intermediates that have been identified in multiple kinetic studies in *E. coli*<sup>36–39</sup> and more recently in *Mycobacterium tuberculosis*<sup>40,41</sup>. Finally, the promoter sequence determines the rates of each step that lead to the open complex, which results in differential kinetics between promoters and thus differing transcriptional outputs<sup>35,42</sup>.

Transcriptional initiation studies in mycobacteria, *Thermus* species and *B. subtilis* suggest that the kinetics of initiation and the nature of populated stable intermediates differ among the RNAPs from these bacteria<sup>40,41,43,44,45</sup>. For example, on the same promoter, complexes of *E. coli* RNAP were shown to be considerably more stable than complexes of *M. tuberculosis* RNAP<sup>40,41,45</sup>. Although the ratios of specific populated intermediates are expected to differ, the motions of RNAP and the path of DNA melting are expected to conform, given the conservation of the DNA cleft, the consensus architecture of RNAPs, and the structural uniformity of open complexes in different clades of bacteria<sup>26,28,29,46</sup>.

Recent advances in single-particle cryo-electron microscopy (cryo-EM) have led to high-resolution visualization of dynamic macromolecular complexes without the constraints of crystallization, which preferentially captures static conformations. RNAP is a remarkably flexible enzyme, with the pincers moving through a range of 15 Å during the transcription cycle<sup>32,46,47</sup>. Cryo-EM enables the classification and extraction of poorly populated structural states not previously amenable to crystallization and has recently been applied to capture a series of stepwise promoter melting intermediates. These approaches include obtaining structures of RNAPs from different organisms and RNAP complexes with promoters that have slow unwinding kinetics or



form reversible open complexes. Additional approaches include the use of inhibitors and transcription factors that regulate the formation of the open complex to trap melting intermediates.

These structures revealed the clamp dynamics in RNAPs, domain-specific and lineage-specific insertion movements during promoter melting, and the path of DNA during unwinding<sup>46–48</sup>. FIGURE 4 presents the data

from two structural studies to provide a composite view of RNAP motions that lead to promoter melting. Early promoter melting intermediates of *E. coli* RNAP were captured with a reversible ribosomal protein promoter using the transcription factor TraR, which increases the population of the intermediates at equilibrium<sup>48</sup>. These structures span the initial recognition of the duplex promoter in a ‘closed’ complex to the final open complex.

◀ Fig. 2 | Towards structural elucidations of RNA polymerase initiation complexes.

In 1969, the composition of the core subunits of RNA polymerase (RNAP) was elucidated<sup>7</sup>. It was not until 1989 that the first structure of RNAP was visualized by 3D reconstructions of *Escherichia coli* holoenzyme RNAP taken from electron microscopy (EM) images of negatively stained 2D crystals<sup>143</sup>. This structure revealed the shape, size and active site cleft of RNAP. Ten years later, the atomic resolution structure of *Thermus aquaticus* RNAP was solved by X-ray crystallography<sup>8</sup>. Thirty years of biochemical and genetic studies of transcription in bacteria could now be contextualized structurally. The high-resolution structure showed the details of the catalytic site, including the active site Mg<sup>2+</sup>, and identified multiple motifs that were soon identified to be structurally conserved in the eukaryotic RNAPs<sup>6,144</sup>. This milestone in transcriptional studies revolutionized the future of the field and set the stage for an impressive number of structures of RNAP complexes to follow. X-ray crystallography of these macromolecules reigned for almost two decades as the method for structural determination of transcriptional complexes. Because it was assumed that the thermophilic enzymes, owing to their intrinsic stability, were more amenable to crystallography than their mesophilic homologues, the transcription field limited structural studies of RNAPs to the genetically intractable species *T. aquaticus* and *Thermus thermophilus*. This approach changed in 2013, when three groups independently crystallized *E. coli* RNAP in various complexes<sup>145–147</sup>. The third group of bacteria whose structure of RNAP has been elucidated is mycobacteria. In 2017, *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* transcription initiation complexes with antibiotics and the RbpA tether were determined by X-ray crystallography<sup>26,40,114</sup>. In the early 2010s, electron microscopy re-emerged as a tool to obtain structures of large macromolecular complexes owing to advances in electron detectors and software that made high-resolution imaging possible. Given its mass of more than 400 kDa, RNAP was the perfect specimen for the resolution revolution that was occurring in the field of single-particle cryo-electron microscopy (cryo-EM). In 2016, high-resolution (lower than 4 Å) single-particle cryo-EM structures of yeast RNAP transcription initiation complexes were solved<sup>148</sup>. These structures were rapidly followed by the cryo-EM structures of various bacterial RNAP complexes<sup>109,149,150</sup>. This method was then applied to observe *M. tuberculosis* RNAP in a range of molecular conformations that the enzyme goes through during promoter unwinding, including the clear capture of a promoter-unwinding intermediate<sup>46,47</sup>. That work was then followed by additional initiation pathway intermediates with the *E. coli* RNAP<sup>48</sup>. Single-particle cryo-EM has now made it possible to determine structures of transient steps of transcription initiation. In full circle, the electron microscope is now the leading instrument for determining structures of RNAP macromolecular complexes. The timeline highlights some of the breakthroughs in the structural biology of RNAP, with a focus on the initiation complexes. CRE, core-recognition element; ECF, extracytoplasmic function. Image for 1969 adapted from REF.<sup>11</sup>, Springer Nature Limited. Image for 1989 adapted from REF.<sup>143</sup>, Springer Nature Limited.

**Abortive transcription**

Process where RNA polymerase cycles between the initial transcribing complex and the open complex, leading to the release of short RNAs.

**Scrunching**

Process during transcription initiation in which the RNA polymerase pulls the downstream DNA into the active site, causing an expansion of the transcription bubble inside the RNA polymerase cleft.

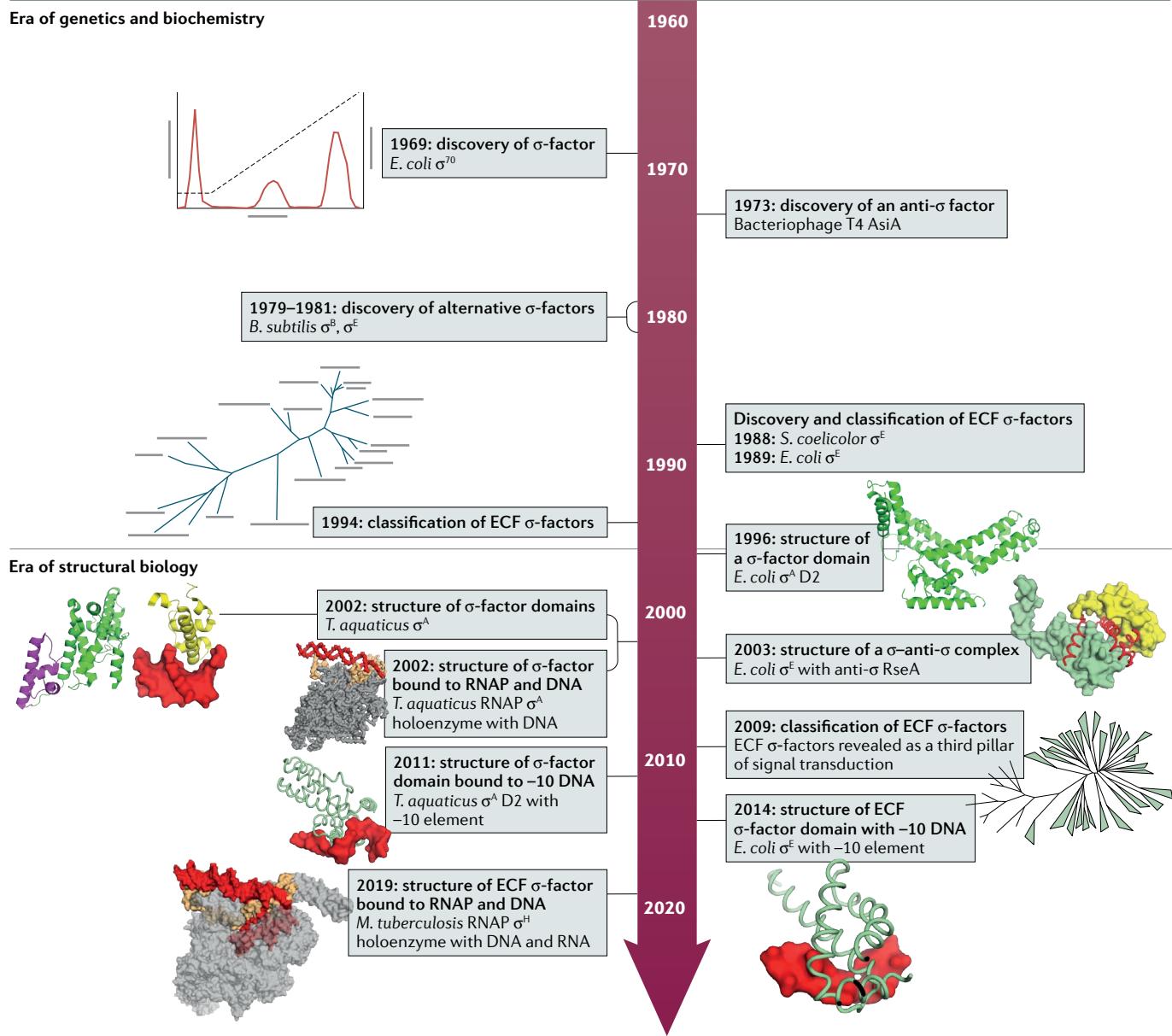
In these structures, the duplex promoter DNA is initially bound as a closed complex. Then the -10 element is nucleated by β'-clamp closure, followed by bubble propagation leading to an early partially melted intermediate (four or five bases unwound), while σD1.1 still occupies the downstream RNAP channel. A late melting intermediate with a partially melted (8 of 13 bases) promoter was observed with the *M. tuberculosis* RNAP, which forms an unstable and reversible open complex on its ribosomal RNA operon promoter<sup>46</sup>. This intermediate was also observed kinetically on the same promoter<sup>40</sup>. This structure indicated that unwinding of the downstream part of the transcription bubble (the unwound section of the promoter) occurs within the central RNAP cleft after σD1.1 has been ejected from the downstream RNAP channel. The same study also showed that the antibiotic corallopyronin freezes the enzyme in this intermediate conformation by inhibiting the clamp movements. This suggests that clamp opening is required for late but not early melting.

**After promoter melting.** Once the promoter is fully melted, the DNA template strand is loaded into the RNAP active site, positioning the transcription start site

for templating RNA catalysis<sup>28,29</sup>. The enzyme then goes through a series of steps, illustrated in FIG. 1a. Briefly, the transcriptionally competent open complex begins de novo RNA synthesis by catalysing phosphodiester bonds between incoming rNTPs to form the initial transcribing complex. Some complexes, depending on the promoter, undergo a non-productive cycle, called ‘abortive transcription’, where the enzyme cycles between the initial transcribing complex and the open complex, releasing short RNAs<sup>49–51</sup>. The ‘decision’ to enter or leave the non-productive cycle to advance into productive transcription depends on the promoter sequence, interactions with the σ-factor and the initial transcribing sequence<sup>52,53</sup>. The increased length of the RNA–DNA hybrid induces promoter escape by causing scrunching in the transcription bubble generated by the transcribing RNAP, which remains static while unwinding and pulling the downstream DNA into the active site cleft<sup>54–56</sup> while the growing RNA chain clashes with σD3.2 (REFS<sup>16,57–61</sup>). Recent biochemical data have added to the nuances of promoter escape, including how pausing after synthesis of a 6-mer RNA affects the branching of productive versus non-productive transcription during initiation<sup>62</sup>. Promoter escape is not well characterized in other bacterial systems. However, recent work comparing RNAPs from *E. coli* and *M. tuberculosis* showed that *M. tuberculosis* RNAP escapes more readily than *E. coli* RNAP on the same promoter. This study showed that two essential housekeeping transcription factors not found in *E. coli*, CarD and RbpA, slow escape kinetics, which suggests that promoter escape is a prime step of regulation in this pathogen<sup>63</sup>. In the productive transcription step, the RNAP escapes into the elongating stage, and the σ-factor can dissociate from the holoenzyme and bind another core RNAP. The current view is that σ-release is not obligatory but rather is stochastic<sup>64</sup>. Both σD3.2 and σD4 must be displaced from their original interactions with RNAP as they are in the path of the elongating RNA. However, σ can be retained throughout much of the elongation cycle<sup>65–68</sup>, most likely via interaction of σD2 and the clamp helices. This retention has been suggested to be a mechanism to regulate elongation and pausing<sup>64,69,70</sup>. Most of the σ-cycle paradigm was formulated from studies in *E. coli*, and it remains to be seen whether these principles apply to other bacteria. We note that the features of σ and RNAP that mediate this cycle are conserved in other bacteria. Indeed, a recent structural study with *Thermus thermophilus* and *M. tuberculosis* holoenzymes describes a similar physical mechanism of σ displacement as described for *E. coli*<sup>61</sup>. FIGURE 1a encapsulates the pathway from σ association to core RNAP to the elongating step.

**Redistributing RNAP holoenzyme populations**

**Alternative σ-factors.** In addition to the primary σ-factor, which is essential for viability during normal growth conditions, almost all bacteria possess additional σ-factors known as alternative σ-factors. These factors recognize DNA sequence motifs distinct from the primary housekeeping σ-factor. Alternative σ-factors enable specific regulon expression in response to stress, cell density, developmental transitions and nutritional



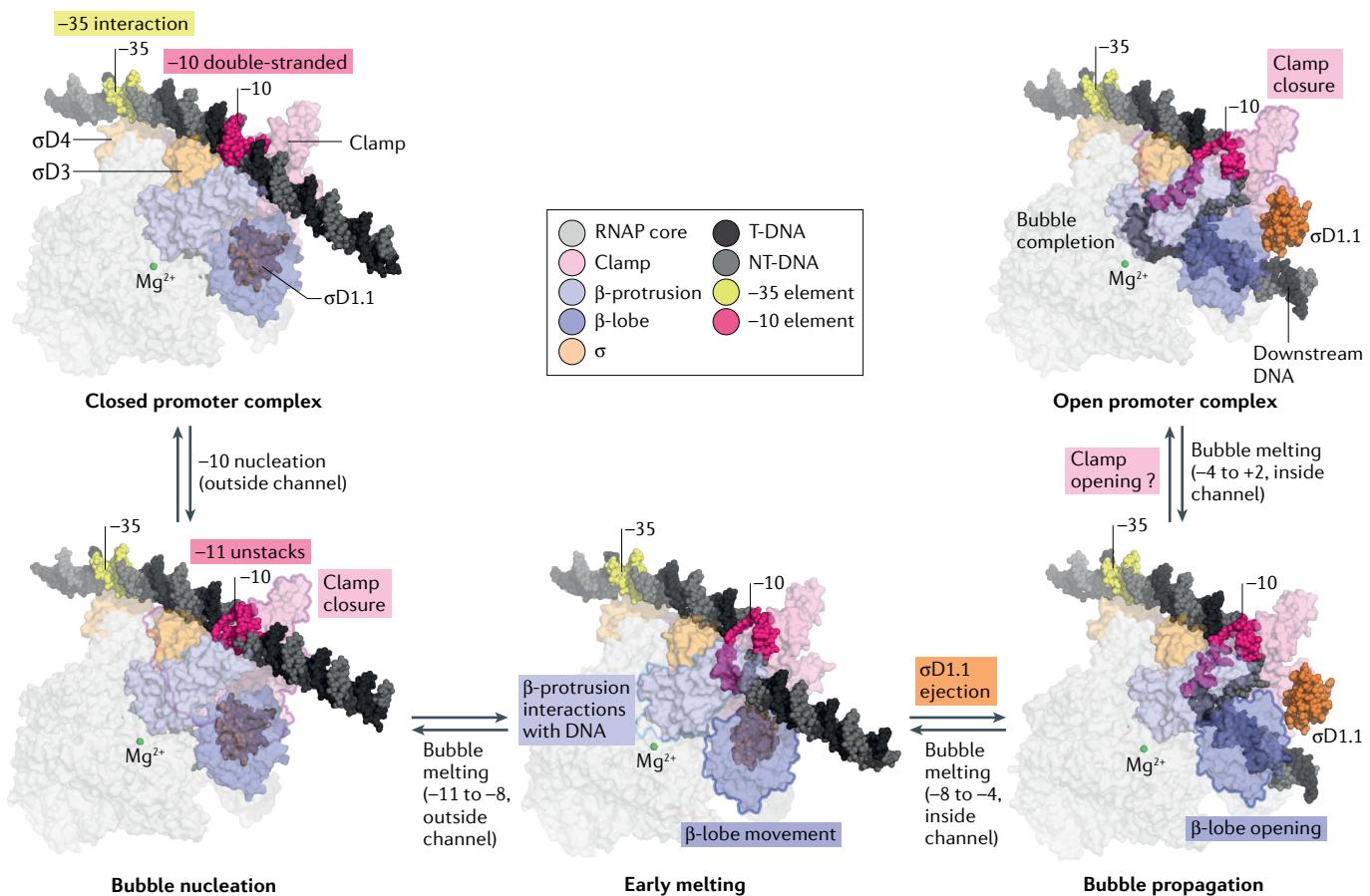
**Fig. 3 | 50 years of studying  $\sigma$ .** The year 2019 marked the 50th anniversary of the discovery of  $\sigma$ -factors. In 1969, the seminal discovery that an extraneous factor was required for the core enzyme to initiate transcription was presented<sup>11</sup>. Shortly afterwards, the factors that inhibit  $\sigma$ -factors (anti- $\sigma$ -factors) were discovered in bacteriophage T4 (REF.<sup>151</sup>). In 1979, the diversity of  $\sigma$ -factors was realized by the discovery that *Bacillus subtilis* RNA polymerase (RNAP) required alternative  $\sigma$ -factors under different environmental conditions<sup>152,153</sup>. Between 1988 and 1994, the most abundant and divergent group of alternative  $\sigma$ -factors was discovered and classified as the extracytoplasmic function (ECF)  $\sigma$ -factors<sup>154–156</sup>. Structural studies put into context the extensive genetic and biochemical groundwork, giving details of how  $\sigma$ -factors preserve their overall structures yet maintain promoter recognition specificity. This structural era started with the crystal structure of *Escherichia coli*  $\sigma$ -domain 2 ( $\sigma$ D2) in 1996 (REF.<sup>157</sup>), followed by structures of the individual domains of *Thermus aquaticus* housekeeping  $\sigma$ D4 with the 35 element<sup>14</sup> and  $\sigma$ D2 with the -10 element<sup>25</sup>, revealing the basis of promoter recognition. The first structures of  $\sigma$ -anti- $\sigma$  complexes appeared in the early years of the first decade of the twenty-first century and explained the basis of their inhibitory activity<sup>80,81</sup>. In 2009, a comprehensive genomic analysis of ECF  $\sigma$ -factors proposed them as the third pillar of bacterial signal transduction<sup>75</sup>. Studies of the promoter specificity of ECF  $\sigma$ -factors and their negative regulation by anti- $\sigma$ -factors have been applied to synthetic biology<sup>158</sup>. The crystal structure of  $\sigma$ D2 of *E. coli*  $\sigma^E$  in complex with a -10 promoter element then enabled the comparison of promoter recognition by this divergent class of  $\sigma$ -factors with that of the housekeeping  $\sigma$ -factors<sup>159</sup>. The  $\sigma$ -factors have also been structurally characterized in complex with RNAP since 2002 (REF.<sup>16</sup>), both by X-ray crystallography and more recently by cryo-electron microscopy (cryo-EM), culminating with crystal structures of the ECF  $\sigma$ -holoenzymes with DNA<sup>61,160,161</sup>. *M. tuberculosis*, *Mycobacterium tuberculosis*; *S. coelicolor*, *Streptomyces coelicolor*. Image for 1969 adapted from REF.<sup>11</sup>, Springer Nature Limited. Image for 1994 cladogram adapted with permission from REF.<sup>156</sup>, PNAS. Copyright (1994) National Academy of Sciences, USA. Image for 2009 adapted with permission from REF.<sup>75</sup>, Wiley.

cues<sup>71–74</sup>. The collection of alternative σ-factors can range in number from zero to hundreds among different species. The alternative σ-factors are classified into four general groups on the basis of the presence of ancillary regions. The group IV σ-factors, also called ‘extracytoplasmic function σ-factors’ (ECF σ-factors), contain only the minimal domains, σD2 and σD4, that are sufficient to bind to core RNAP and promoter elements. This group, which represents the most abundant and diverse σ-factors, has been extensively characterized functionally and phylogenetically and is referred to as the third pillar of signal transduction in bacteria (the other two being one- and two-component systems)<sup>75</sup>. ECF σ-factors are involved in a range of functions, including response to oxidative stress, metal homeostasis, virulence, periplasmic and envelope stress, and sporulation.

Thus, the availability of specific holoenzymes can drastically reprogramme gene expression and thus provides a powerful way to express discrete regulons under

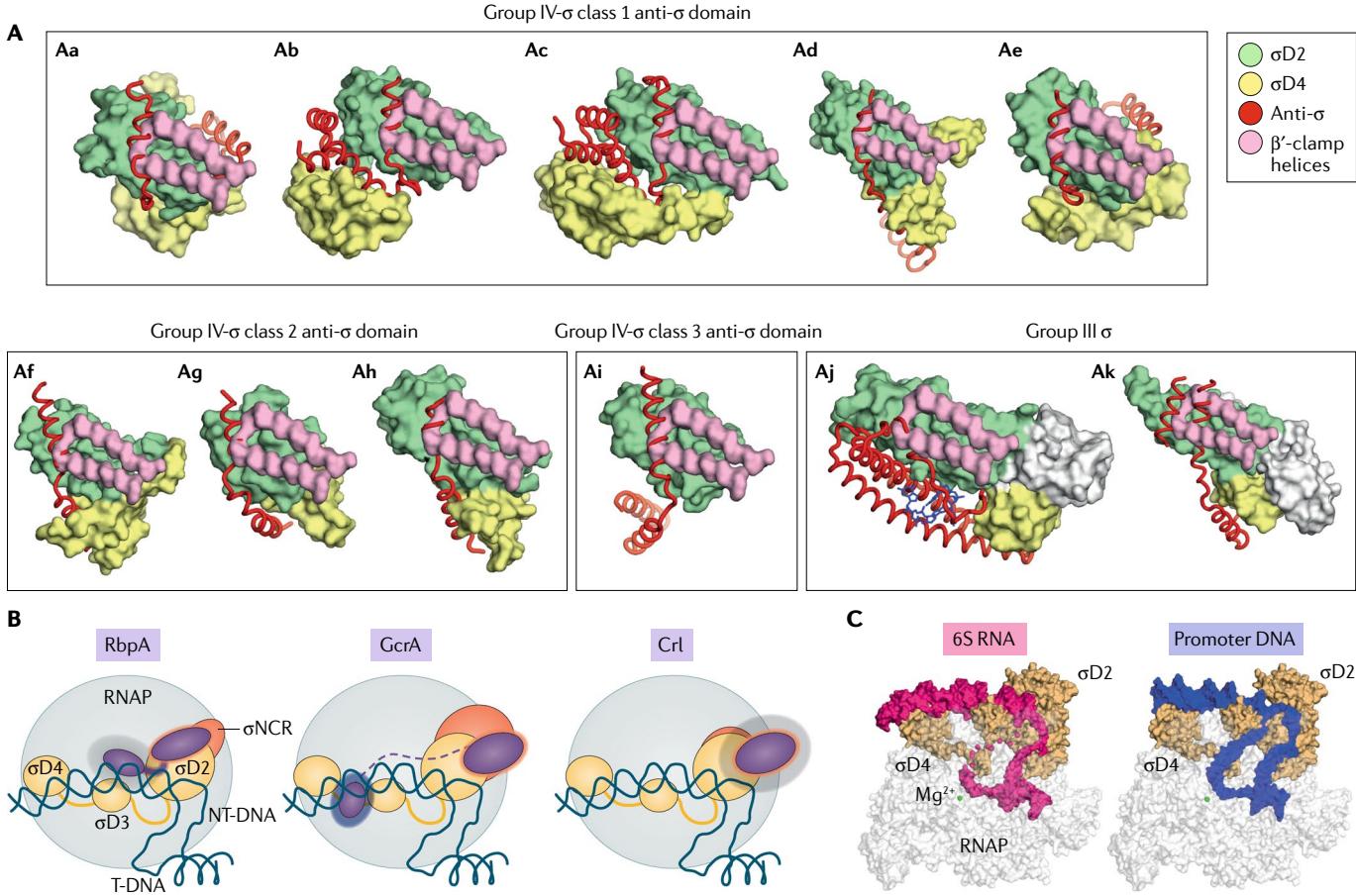
specific conditions. In this section, we discuss classic and recently described modes of regulating specific holoenzyme populations as a means for fine-tuning gene expression. We review three examples of how bacteria use different means to prevent, enhance or redistribute specific holoenzyme populations.

**What is new with anti-σ-factors?** Anti-σ-factors associate with their cognate σ-factor and function by preventing the σ-factor from interacting with core RNAP, thus inhibiting the assembly of the holoenzyme<sup>76</sup>. The genes encoding the anti-σ-factors are usually found directly downstream of those that encode the σ-factor they regulate. More than 30% of the ECF σ-factors are regulated by anti-σ-factors<sup>77–79</sup>. Multiple structural studies of different groups of σ-factors with their cognate anti-σ-factors have revealed the molecular basis of how anti-σ-factors inhibit holoenzyme formation<sup>78,80–88</sup>. The structures showed that anti-σ-factors bind at least one



**Fig. 4 | Structure-based model of open promoter complex formation highlighting RNA polymerase motions that lead to promoter melting.** Structural models of RNA polymerase (RNAP) were generated from cryo-electron microscopy structures of promoter melting intermediates from *Escherichia coli* (Protein Data Bank (PDB) identifiers (IDs) 6PSQ, 6PSS, 6PST and 6OUL)<sup>48</sup> and *Mycobacterium tuberculosis* (PDB ID 6EE8)<sup>46</sup> with lineage-specific inserts removed for clarity. Proteins are shown as a transparent molecular surface, whereas σ-domain 1.1 (σD1.1) and promoter DNA are shown as spheres. The Mg<sup>2+</sup> (shown as a green sphere) marks the active site. The pathway is composed of structures delineating bubble propagation that leads to the open promoter complex. The model consists

of the following steps: the duplex promoter DNA is initially bound as a closed complex; then the −10 element unwinding is nucleated by clamp (made up of σD2 and the β'-clamp module) closure leading to unstacking of the −11 element; the nucleated bubble propagates to an early partially melted intermediate (four or five bases unwound), while σD1.1 still occupies the downstream RNAP channel; the β-lobe opens, allowing the transcription bubble to propagate further (eight bases unwound), and σD1.1 is ejected from the channel; the bubble completes melting on clamp opening (needs to be verified experimentally); the completed transcription bubble is stabilized in the RNAP cleft by the clamp closure. NT-DNA, non-template DNA; T-DNA, template DNA.



**Fig. 5 | Factors that regulate the pool of holoenzymes.** **A** | Despite divergent sequence, structure and phylogenetic origins, anti- $\sigma$ -factors (red) prevent binding of the alternative  $\sigma$ -factors by occluding the major core-binding interface ( $\sigma$ -domain 2 ( $\sigma$ D2); green) from the core  $\beta'$ -clamp helices (pink). The  $\beta'$ -clamp helices were modelled by superimposing  $\sigma$ D2 from structures of the  $\sigma$ -anti- $\sigma$  complexes onto the  $\sigma^A$  holoenzyme structure (Protein Data Bank (PDB) identifier (ID) 19LU)<sup>16</sup>. The structures shown are as follows: *Rhodobacter sphaeroides*  $\sigma^E$ -ChrR (PDB ID 2Q1Z; panel **Aa**)<sup>17</sup>, *Escherichia coli*  $\sigma^F$ -RseA (PDB ID 1OR7; panel **Ab**)<sup>18</sup>, *Pseudomonas aeruginosa*  $\sigma^H$ -MucA (PDB ID 6IN7; panel **Ac**)<sup>16</sup>, *Bacillus subtilis*  $\sigma^W$ -RsiW (PDB ID 5WUR; panel **Ad**)<sup>16</sup>, *Mycobacterium tuberculosis*  $\sigma^K$ -RskA (PDB ID 4NQW; panel **Ae**)<sup>19</sup>, *Bartonella quintana*  $\sigma^F$ -NepR (PDB ID 5UXX; panel **Af**),

*Caulobacter vibrioides*  $\sigma$  mimic PhyR-NepR (PDB ID 3T0Y; panel **Ag**)<sup>20</sup>, *Cupriavidus metallidurans*  $\sigma^{CnrH}$ -CnrY (PDB ID 4CXF; panel **Ah**)<sup>21</sup>, *Streptomyces venezuelae*  $\sigma^{BldN}$ -RsbN (PDB ID 6DXO; panel **Ai**)<sup>22</sup>, *S. venezuelae*  $\sigma^{WhiG}$ -RsiG (PDB ID 6PFJ; panel **Aj**)<sup>23</sup> and *Aquifex aeolicus*  $\sigma^{FlgA}$ -FlgM (PDB ID 1SC5; panel **Ak**)<sup>24</sup>. **B** | Schematics of the three founding members of the  $\sigma$ -tethers (RbpA, GcrA and Crl; purple), tethering  $\sigma$ D2 and  $\sigma$ -non-conserved region ( $\sigma$ NCR) to the core enzyme and or promoter DNA. **C** | 6S RNA mimics the promoter architecture of an open complex, sequestering the housekeeping  $\sigma$ -holoenzyme. Shown are the cryo-electron microscopy structure of 6S RNA-holoenzyme (PDB ID 5VT0)<sup>25</sup> and the crystal structure of  $\sigma^{70}$  holoenzyme with promoter DNA (PDB ID 4YLN)<sup>26</sup>. NT-DNA non-template DNA; RNAP, RNA polymerase; T-DNA, template DNA.

RNAP-binding surface of their cognate  $\sigma$  and physically block the association with the core RNAP. In addition, the anti- $\sigma$ -factor often contorts the  $\sigma$ -factor in variable configurations such that additional core-binding interfaces are inaccessible.

Although this general inhibitory mechanism is conserved, anti- $\sigma$  factors can be categorized into diverse classes<sup>74,79,89</sup>. Many are multidomain proteins with an N-terminal anti- $\sigma$  domain fused to one of a wide range of regulatory domains, which is often subject to signal transduction or proteolysis. These regulatory domains often contain transmembrane regions (the majority), metal-binding motifs, short disordered tails, or serine and threonine kinase motifs<sup>77,78</sup>. Three anti- $\sigma$  domain classes have been identified for the ECF  $\sigma$ -factors<sup>77,78,83</sup>. Even so, the anti- $\sigma$  domains are not well conserved neither in sequence nor in structure and thus interact with their cognate  $\sigma$ -factor in very different ways (FIG. 5A).

However, one common theme emerges when aligning  $\sigma$ D2, which is the most conserved domain of  $\sigma$ -factors, from the published structures of  $\sigma$ -anti- $\sigma$  complexes from distantly related bacteria. Each of the anti- $\sigma$ -factors uses a single helix to bind to the same surface of its cognate  $\sigma$ D2, which would physically prevent the interaction with the RNAP  $\beta'$ -clamp helices (FIG. 5A). This finding suggests that this interface of  $\sigma$ , although conserved in function and essential for the interaction with core RNAP, is variable enough to be targeted by specific anti- $\sigma$ -factors. The interaction between  $\sigma$ D2 and the  $\beta'$ -clamp helices of RNAP is the primary determinant for holoenzyme assembly and stability and would thus serve as an interface for drug targeting. More recently, a new type of anti- $\sigma$  (anti- $\sigma^{RsiG}$ ), which requires cyclic di-GMP to bind to the  $\sigma$ -factor ( $\sigma^{WhiG}$ ), similarly occludes  $\sigma$ D2 binding to the clamp helices<sup>84</sup>. The analysis of a catalogue of  $\sigma$ -anti- $\sigma$  structures revealed that although the

sequence, structure and function of the anti- $\sigma$ -factors differ widely across bacteria, a common feature of inhibition seems to have been selected. We note that two exceptions to this theme are the phage anti- $\sigma$  AsiA and the bacterial anti- $\sigma$  Rsd, both of which inhibit the housekeeping  $\sigma$ -factor in *E. coli* by binding  $\sigma$ D4 (REFS<sup>89–91</sup>).

**$\sigma$ -tethers — a new class of  $\sigma$  regulators.** In contrast to anti- $\sigma$ -factors, which prevent the binding of  $\sigma$  to core RNAP, recent studies have revealed a new group of factors that promote the association of  $\sigma$  with core RNAP and/or DNA. These factors have been characterized in diverse species of bacteria, and we introduce the term ‘ $\sigma$ -tethers’ to describe the mechanism of this emerging group of RNAP regulators (FIG. 5B). The first example of a  $\sigma$ -tether is RbpA, an essential housekeeping  $\sigma$ -binding factor that activates transcription in Actinobacteria<sup>92–94</sup>. Structural studies of RbpA in complex with the *M. tuberculosis* holoenzyme and promoter DNA<sup>40,46</sup> illustrated the following: RbpA interacts with the non-conserved region and  $\sigma$ D2 of the housekeeping  $\sigma$  via its  $\sigma$ -interacting domain; RbpA also makes contacts with the core RNAP via a second domain called the ‘core-binding domain’; the linker between these two domains of RbpA contacts the phosphate backbone of promoter DNA and the N-terminal tail contacts the template DNA in the active site. This tethered configuration of RbpA is consistent with its described roles in stabilizing the interactions between  $\sigma$  and the core<sup>92,95</sup> and between the holoenzyme and DNA<sup>96</sup> (FIG. 5B, left panel).

The second example of a  $\sigma$ -tether is GcrA (conserved in Alphaproteobacteria), which has been characterized in *Caulobacter crescentus*. GcrA is a transcription factor that binds to the housekeeping  $\sigma$ -factor and activates methylated promoters<sup>97</sup>, a function critical for coupling DNA replication with cell division during cell cycle progression<sup>98</sup>. Two X-ray crystal structures — one of the  $\sigma$ -interacting domain of GcrA with  $\sigma$ D2 and the non-conserved insert of  $\sigma$ D2, and the other of the DNA-binding domain of GcrA on methylated DNA — revealed that this protein, like RbpA, also functions as a tether between  $\sigma$ D2 and promoter DNA<sup>99</sup> (FIG. 5B, middle panel).

The third example of a  $\sigma$ -tether is the transcriptional regulator Crl, which has been implicated in facilitating holoenzyme assembly in *E. coli*, and was shown to bind to the stationary phase  $\sigma$ -factor,  $\sigma^s$ . Crl is critical for  $\sigma^s$  to compete with the housekeeping  $\sigma$ -factor during the transition from exponential to stationary phase<sup>100,101</sup>. Crl activates genes involved in the stationary phase in *E. coli*, which also have roles in virulence and infection in many pathogenic Gammaproteobacteria such as *Vibrio cholera*, *Salmonella enterica* and *Yersinia pestis*. The cryo-EM structures of Crl in complex with a  $\sigma^s$  holoenzyme revealed that Crl also tethers  $\sigma^s$  to RNAP via the clamp toe, which is a domain on the  $\beta'$ -subunit, and evidence also suggests that Crl transiently contacts DNA during promoter unwinding<sup>102,103</sup> (FIG. 5B, right).

**Sugar puckers**  
Ring conformation of the deoxyribose or ribose sugar in DNA or RNA, respectively.

**Stringent response**  
Bacterial response to stresses in the environment such as amino acid starvation.

**6S RNA — a promoter mimic that modulates the pool of holoenzyme.** Bacterial 6S RNA is a non-coding RNA that binds and regulates *E. coli* RNAP activity<sup>104</sup>. Although

6S RNAs are widely distributed in bacteria<sup>105,106</sup>, they have been best characterized in *E. coli*. *E. coli* 6S RNA uses a unique mechanism to control the availability of specific holoenzymes and thus regulates gene expression (reviewed in REF.<sup>107</sup>). The 6S RNA binds to  $\sigma^{70}$  holoenzymes but not holoenzymes containing alternative  $\sigma$ -factors and inhibits the ability of the enzyme to bind promoter DNA. This sequestration of  $\sigma^{70}$  holoenzymes has a role in the transcriptional reprogramming during the transitions between exponential and stationary growth phases. The preferential binding of 6S RNA to  $\sigma^{70}$  holoenzyme promotes the increased activity of  $\sigma^s$ . Although not conserved in sequence, the secondary structure of 6S RNAs contains a prominent central bubble that is reminiscent of the promoter DNA bubble in the transcription open promoter complex<sup>105,106</sup> and functions as a transcription template during outgrowth from the stationary phase<sup>108</sup>. Through this mechanism, 6S RNA plays an important part in long-term survival during and escape from the stationary phase. Some bacteria have two copies of the 6S RNA, and, although the roles are less characterized, expression profiles suggest that 6S RNAs may be necessary for virulence and other cellular functions in different bacteria<sup>107</sup>.

A cryo-EM structure supported by biochemical assays of *E. coli*  $\sigma^{70}$  holoenzyme with 6S RNA showed that two amino acids largely determine the specificity of 6S RNA for  $\sigma^{70}$  holoenzymes over  $\sigma^s$  holoenzymes in *E. coli*. The structure also revealed that although 6S RNA maintained A-form C3'-endo sugar pucks, it adopts a widened major groove, giving the RNA an unusual overall architecture that mimics B-form promoter DNA<sup>109</sup>. This surprising observation explained how a non-coding RNA functions as a B-form DNA mimetic to regulate transcription by a DNA-dependent RNAP (FIG. 5C).

### Lineage-specific regulation

**Lineage-specific domains in core RNAP.** The catalytic core RNAP is conserved in sequence, structure and fundamental function in all cellular organisms<sup>110,111</sup>. In bacteria, the  $\beta$  and  $\beta'$  subunits are highly conserved. However, these shared regions are often separated by spacers of non-conserved lineage-specific inserts that range from 50 to 500 amino acids in size. These inserts are typically independently folded, located on the surface of the RNAP and highly mobile. FIGURE 6a highlights inserts in the RNAPs of the three phylogenetically distant clades, *Thermus* (*Thermus aquaticus*), Proteobacteria (*E. coli*) and Actinobacteria (*M. tuberculosis*), whose structures have been determined.

The specific functions of these inserts are mostly unknown. On the basis of the structural models of RNAP from *Thermus* spp., it has been proposed that the  $\beta'i2$  insert (FIG. 6a, left panel) stabilizes the binding of  $\sigma$  to the RNAP<sup>17</sup>. In *E. coli*, deletions of the  $\beta i4$  and  $\beta'i6$  inserts in RNAP (FIG. 6a, middle panel) lead to temperature sensitivity and affect cell viability, respectively<sup>112</sup>. Recent work demonstrated that these inserts have essential roles in the transcriptional regulation by the transcription factor TraR and likely DksA (see later), which implicates these factors in regulation of the stringent response<sup>113</sup>. Structures of RNAPs from mycobacteria revealed that

<sup>a</sup>

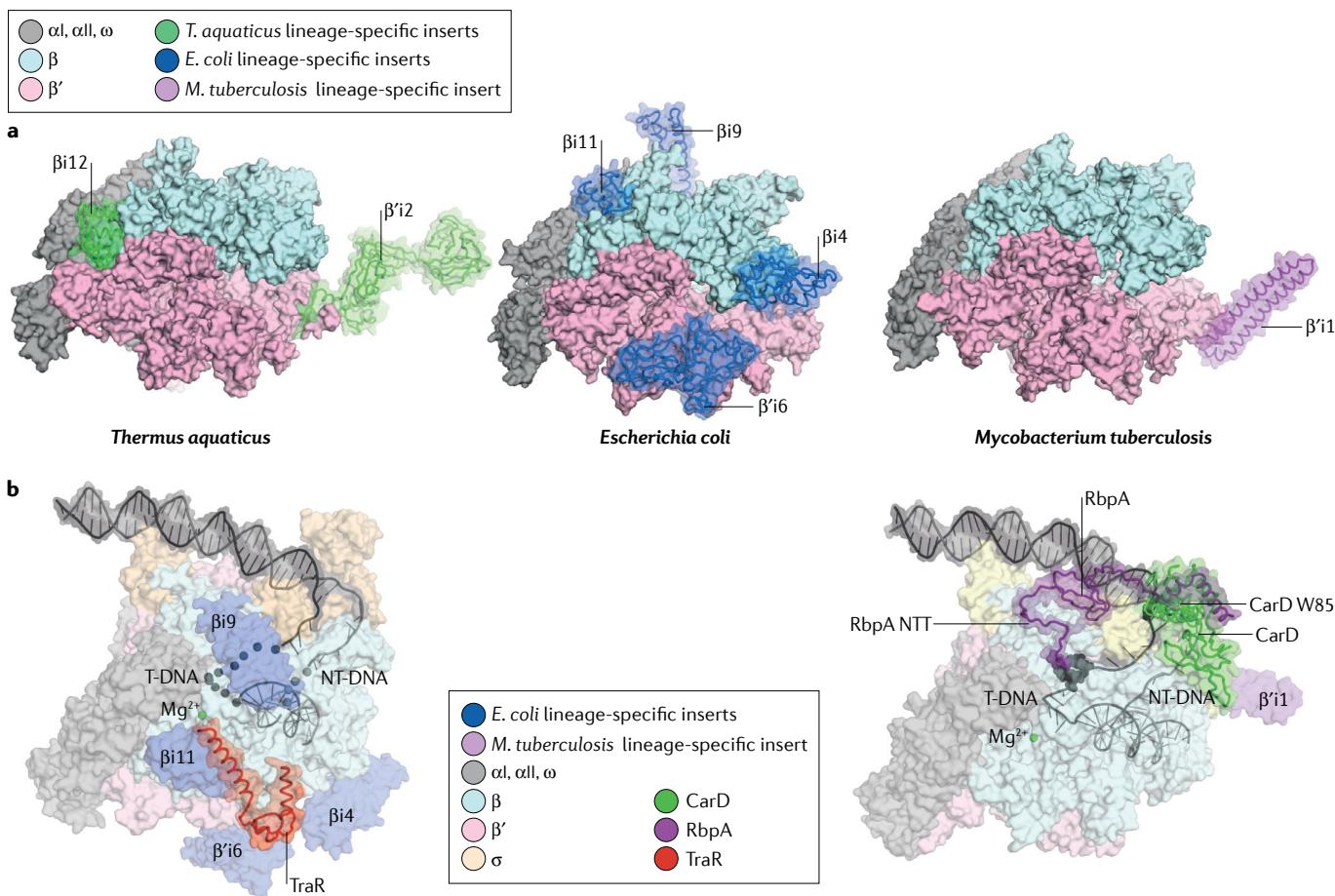
The housekeeping  $\sigma$ -factor in bacteria that controls the expression of most genes during growth.

the  $\beta'i1$  insert (FIG. 6a, right panel) emerges from the tip of the RNAP clamp module, is coupled to clamp conformations and interacts with the N terminus of  $\sigma^A$ , which increases the stability of the open complex<sup>17,114</sup>. Besides TraR and DksA, it is unknown whether additional transcription factors interact with these lineage-specific inserts, leaving their roles in regulating transcription a compelling research area.

**Lineage-specific transcription factors.** Classic transcription factors bind directly to DNA upstream of the –35 element to either activate transcription by recruiting RNAP to DNA or inhibit transcription by blocking RNAP binding to DNA<sup>115</sup>. However, several transcription factors do not directly bind DNA and use unique mechanisms to regulate transcription. Recently, the mechanisms for some of these factors have been revealed by a combination of approaches, including transcriptomics,

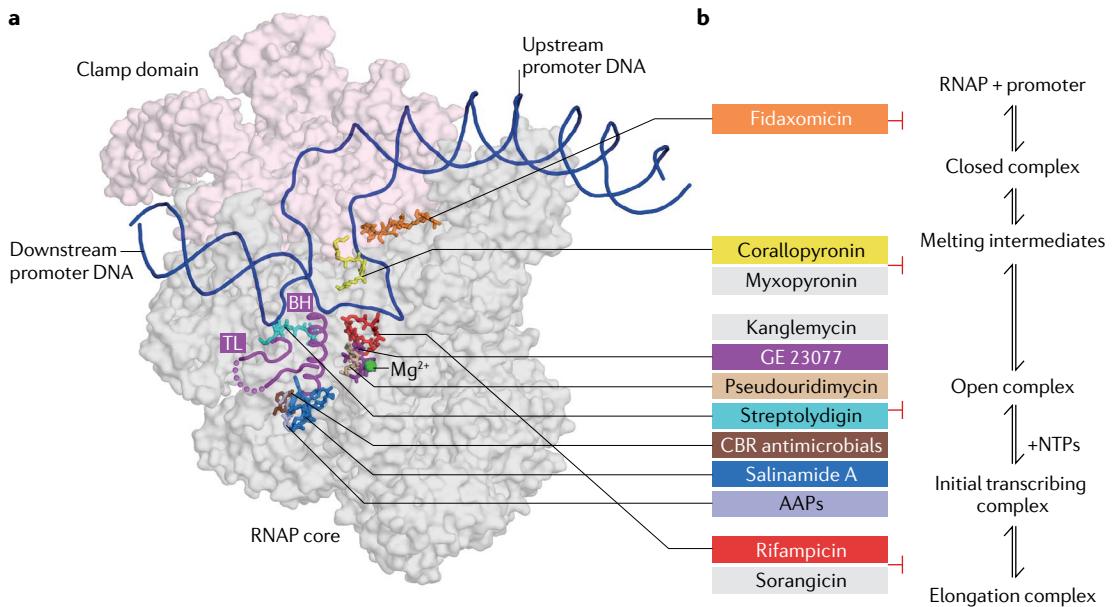
bioinformatics, biochemistry and structural biology. These ‘unconventional’ factors include SutA found in *Pseudomonas aeruginosa*<sup>116</sup>, DksA-ppGpp<sup>117,118</sup> and TraR<sup>113,119</sup> found in Proteobacteria, RbpA<sup>26,93,94</sup> found in Actinobacteria and the more widespread CarD found in one-third of the bacterial genomes sequenced so far<sup>120</sup>. These unconventional factors bind directly to RNAP to exert regulatory effects. FIGURE 6b highlights two recent RNAP structures from *E. coli* and *M. tuberculosis* that depict how these factors regulate transcription initiation.

In *E. coli*, TraR binds the secondary channel of RNAP and interacts directly with the  $\beta'i4$  and  $\beta'i6$  lineage-specific inserts (FIG. 6b, left panel), which causes conformational changes in the RNAP that lead to differential regulation of certain promoters<sup>48</sup>. In *M. tuberculosis*, CarD binds the  $\beta$ -protrusion and wedges the upstream DNA fork junction, thereby stabilizing the transcription bubble<sup>121</sup>.



**Fig. 6 | Diversity in bacterial RNAPs and their regulation.** **a** | Structures of *Thermus aquaticus* RNA polymerase (RNAP; Protein Data Bank (PDB) identifier (ID) 4XLR)<sup>121</sup>, *Escherichia coli* RNAP (PDB ID 4LK1)<sup>146</sup> and *Mycobacterium tuberculosis* RNAP (PDB ID 6C05)<sup>47</sup>, highlighting the disposition of lineage-specific inserts. RNAPs are shown as molecular surfaces. Lineage-specific inserts are shown as transparent molecular surfaces with cartoon tubes superimposed (*T. aquaticus* lineage-specific inserts, green; *E. coli* lineage-specific inserts, blue; *M. tuberculosis* lineage-specific insert, magenta). **b** | Examples of unconventional transcription factors that regulate transcription initiation. Proteins are shown as a transparent molecular surface. Promoter DNA is shown as

cartoons superimposed on a transparent molecular surface. The active site  $Mg^{2+}$  is shown as a green sphere. *E. coli* RNAP bound to promoter DNA and TraR (PDB ID 6PSV)<sup>48</sup> (left panel). TraR (red) is shown as cartoon tubes. Highlighted is the interaction between TraR and the *E. coli* RNAP lineage-specific inserts  $\beta i4$  and  $\beta i6$  (coloured blue). *M. tuberculosis* open promoter complex (PDB ID 6EDT)<sup>46</sup> bound to RbpA (purple) and CarD (green) (right panel). Highlighted is the CarD W85 that ‘wedges’ the upstream fork junction, and the RbpA amino-terminal tail (NTT) that interacts with the template DNA (T-DNA) that is loaded in the RNAP active site.  $\beta'i1$  is also shown and is coloured magenta. NT-DNA, non-template DNA.



**Fig. 7 | Inhibitors of RNA polymerase.** **a** | The positions of various antibiotics that bind RNA polymerase (RNAP) are indicated. The structures of sorangicin and kanglemycin are not shown as they overlap with the structure of rifampicin. The structure of myxopyronin is not shown as it overlaps with the structure of corallopyronin. The bridge helix (BH) and trigger loop (TL) are drawn as cartoons. The active site Mg<sup>2+</sup> is shown as a green sphere for reference. **b** | The minimal mechanism for transcription initiation is shown, and the steps inhibited by each antibiotic are indicated. AAPs, N<sup>o</sup>-aryloyl-N-arylphenylalaninamides; NTPs, nucleotides.

(FIG. 6b, right panel). RbpA interacts with σD2 and tethers σ to the core RNAP<sup>40</sup>. A recent cryo-EM structure of an *M. tuberculosis* open promoter complex revealed that the N-terminal tail of RbpA interacts with the template DNA of the promoter DNA near the active site<sup>46</sup> (FIG. 6b, right panel), which has been shown to influence promoter escape kinetics<sup>63</sup>.

#### Inhibiting transcription initiation

Bacterial RNAP is a proven target for antibiotics. Two clinically used antibiotics target RNAP: rifampicin is used for the first-line treatment of tuberculosis caused by *M. tuberculosis*; fidaxomicin is used to combat intestinal infections caused by *Clostridioides difficile*<sup>122</sup>. Many other RNAP inhibitors have been characterized, and these have been useful as probes of the fundamental mechanisms of transcription. By binding to different regions of RNAP, they block transcription at different steps, thus providing insights into the roles of these sites in the transcription cycle. In this section, we summarize recent work on the mechanisms of RNAP inhibitors that affect transcription initiation. FIGURE 7 illustrates the binding sites of these inhibitors and the steps during the transcription cycle that they affect.

**Promoter melting.** The required movements of opening and closing of RNAP during promoter unwinding (FIG. 4) present opportunities for the regulation of RNAP activity. Two classes of inhibitors have been found to block RNAP closing (fidaxomicin) and opening (corallopyronin and myxopyronin).

Fidaxomicin kills *M. tuberculosis*<sup>123</sup>, but, because of its poor systemic absorption, it cannot be used to clear the infection in the lungs. Cryo-EM structures

of *M. tuberculosis* RNAP in complex with fidaxomicin revealed that the antibiotic binds in a groove between the clamp domain and the rest of the RNAP, acting as a ‘doorstop’ to prevent clamp closure for the initial recognition and melting of the –10 element<sup>17,124</sup>. Remarkably, the essential Actinobacteria-specific general transcription factor RbpA interacts with fidaxomicin and increases the in vivo and in vitro sensitivity to the drug. Thus, fidaxomicin inhibits *M. tuberculosis* RNAP much better than *E. coli* RNAP, partly due to the presence of RbpA<sup>47</sup>. These studies emphasize the importance of studying the native RNAP complexes in drug design<sup>17</sup>.

In contrast to fidaxomicin, myxopyronin (an RNAP inhibitor that is structurally related to corallopyronin) closes the RNAP clamp and traps a partially melted promoter bubble<sup>32,125,126</sup>. Cryo-EM studies of *M. tuberculosis* RNAP holoenzyme with corallopyronin and a de novo melted promoter resulted in an intermediate with an 8-nucleotide transcription bubble instead of the full 13-nucleotide bubble<sup>46</sup>. A control experiment without corallopyronin resulted in two structures: an open promoter complex and the same intermediate captured with corallopyronin<sup>46</sup>. This finding suggests that corallopyronin stabilizes a natural intermediate of transcription initiation and inhibits the completion of the open promoter complex.

**Substrate binding, catalysis and productive transcription complexes.** Kanglemycins are rifamycin congeners that have been identified through genome mining and metagenomic analysis of soil microbiomes and inhibit rifampicin-resistant bacteria<sup>127</sup>. Kanglemycins share the same chemical structure as rifampicin but contain additional functional groups, including an acid moiety and a

sugar moiety. Crystal structures and biochemical assays of mycobacteria and *T. thermophilus* RNAPs in complex with kanglemycins show that these inhibitors bind in the same pocket as rifampicin but inhibit transcription at an earlier step. Whereas rifampicin inhibits the transition from the initial transcribing complex to the elongating complex, kanglemycins prevent binding of the initiating rNTP through clashes with the kanglemycin acid moiety<sup>127,128</sup>. This clash was confirmed biochemically as ribonucleosides with monophosphates, but not triphosphates, were capable of initiation<sup>127</sup>. Importantly, the structures showed that the sugar moiety of kanglemycin, which is absent in rifampicin, makes additional contacts with the RNAP. This observation explained the potency of kanglemycins with regard to rifampicin-resistant RNAPs and informs future approaches to combat drug-resistant mycobacteria.

Pseudouridimycin is another inhibitor of nucleotide substrate binding that is active against Gram-positive bacteria and Gram-negative bacteria and does not exhibit cross-resistance to current clinical antibiotics<sup>129</sup>. It is a nucleoside analogue inhibitor that mimics uridine triphosphate. Similarly, another broad-spectrum antibiotic, GE23077, also occupies the initiating nucleotide-binding site and prevents rNTP substrate loading<sup>130</sup>. Linking this inhibitor with rifamycin covalently enhanced its potency against rifampicin-resistant bacteria<sup>130</sup>.

Streptolydigin inhibits transcription initiation and elongation by preventing phosphodiester bond catalysis in bacterial RNAPs but not eukaryotic RNAPs<sup>131–133</sup>. The crystal structures of streptolydigin in complex with *T. thermophilus* RNAP showed that it binds near the trigger loop and bridge helix. These two dynamic structural elements have roles in the nucleotide addition cycle by changing their conformations<sup>134,135</sup>. Hence, streptolydigin was proposed to inhibit transcription by both locking the bridge helix in the straight conformation and unfolding the trigger loop<sup>134,135</sup>.

The CBR antimicrobials, which were discovered through small-molecule library screening, are another class of bacterial RNAP inhibitors that inhibit the catalytic function of RNAP<sup>136</sup>. X-ray crystallography studies with *E. coli* RNAP established that CBR compounds bind near the trigger loop and bridge helix in a region distal to the binding site of streptolydigin and other antibiotics, and hence CBR compounds are not cross-resistant with most other RNAP inhibitors<sup>137,138</sup>. Through these interactions, the CBR antimicrobials inhibit nucleotide addition and pyrophosphorolysis. The CBR antimicrobials are narrow-spectrum antibiotics, inhibiting only those RNAPs from Gram-negative bacteria. A recent screen revealed a new class of compounds, *N*<sup>a</sup>-aroyl-*N*-arylphenylalaninamides, that function similarly to CBR antimicrobials but selectively inhibits *M. tuberculosis* RNAP. Structural studies of these compounds in complex with *M. tuberculosis* RNAP revealed that they also overlap with the CBR antimicrobial-binding site and are likely to exert their function by the same mechanism<sup>114</sup>.

Moreover, salinamides have been proposed to inhibit RNAP by a similar mechanism to the CBR antimicrobials<sup>39</sup>. Whereas CBR antimicrobials inhibit Gram-negative

bacteria and the *N*<sup>a</sup>-aroyl-*N*-arylphenylalaninamides inhibit mycobacteria, salinamides inhibit transcription in both Gram-positive bacteria and Gram-negative bacteria<sup>114,139</sup>.

Rifampicin is one of the most potent broad-spectrum antibiotics used against tuberculosis since 1968. However, resistant strains have emerged, which limits its efficacy and clinical use. Rifampicin sterically blocks RNA synthesis beyond two or three nucleotides by clashing with the elongating transcript<sup>140,141</sup>. Another chemically unrelated RNAP inhibitor, sorangicin, shares the same binding site and mechanism of inhibition as rifampicin<sup>142</sup>. Sorangicin inhibits some, but not all, rifampicin-resistant RNAPs, hinting at differences in interactions with different mutants. This ability of sorangicin to inhibit some rifampicin-resistant RNAPs was attributed in part to its conformational flexibility relative to rifampicin.

The antibiotics reviewed above target various steps of transcription initiation (FIG. 7) and inhibit RNAP by a number of mechanisms. These inhibitors differ not only in their specific mechanisms but also in their selectivity against different bacterial RNAPs. Remarkably, in some cases, this selectivity is provided by a transcription factor, rather than RNAP itself. Characterization of the inhibitors provides insight into their mode of action and helps to advance our understanding of the fundamental principles of bacterial transcription.

## Conclusions

Structural biology, combined with genetic, single-molecule and biochemical studies, has revealed new regulatory mechanisms that extend the early models of transcriptional regulation. This Review has discussed studies that revealed both universal and lineage-specific mechanisms in the regulation of bacterial transcription initiation. These mechanisms rely on regulatory elements, including regions of the RNAP, transcription factors and non-coding RNAs. Until very recently, studies of in vitro bacterial transcription relied on *E. coli*. However, partly owing to cryo-EM, it is now possible to structurally characterize different bacterial RNAPs. The study of transcription in different bacterial lineages not only informs us of clade-specific transcriptional regulation but also contributes to our fundamental understanding of RNAP and its motions as it unwinds DNA. The differences between groups of bacteria are also essential considerations as we seek out new antimicrobials in the age of antibiotic-resistant pathogens, highlighting the importance of studying transcription outside the classic systems and organisms. In vitro characterization of inhibitors of specific pathogenic bacterial RNAPs will provide a platform for the design of diverse inhibitors that specifically target drug-resistant pathogens. In summary, the research reviewed here has expanded our grasp of transcriptional regulation in bacteria and revealed the ever-expanding range of modulatory mechanisms. This Review aims to convey that we have only just begun to sample the range of diverse systems in bacterial transcription initiation, and we expect similar findings in other steps of transcription.

### Trigger loop

Mobile structural element in the RNA polymerase active centre that plays key roles in the nucleotide addition cycle.

### Pyrophosphorolysis

The reverse polymerization reaction in which a pyrophosphate can cleave a phosphodiester bond at the 3' end of an RNA.

1. Crick, F. Central dogma of molecular biology. *Nature* **227**, 561–563 (1970).
2. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **9**, 85–98 (2011).
3. Decker, K. B. & Hinton, D. M. Transcription regulation at the core: similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annu. Rev. Microbiol.* **67**, 113–139 (2013).
4. Pribnow, D. Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proc. Natl Acad. Sci. USA* **72**, 784–788 (1975).
5. Blombach, F., Smollett, K. L., Grohmann, D. & Werner, F. Molecular mechanisms of transcription initiation — structure, function, and evolution of TFE/TFIIE-like factors and open complex formation. *J. Mol. Biol.* **428**, 2592–2606 (2016).
6. Cramer, P. Structure and function of RNA polymerase II. *Adv. Protein Chem.* **67**, 1–42 (2004).
7. Burgess, R. R. Separation and characterization of the subunits of ribonucleic acid polymerase. *J. Biol. Chem.* **244**, 6168–6176 (1969).
8. Zhang, G. et al. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**, 811–824 (1999). **First crystal structure of a DNA-dependent RNAP that provided high-resolution details of the core RNAP and defined many of the RNAP modules.**
9. Cramer, P. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640–649 (2000).
10. Hirata, A., Klein, B. J. & Murakami, K. S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851–854 (2008).
11. Burgess, R. R., Travers, A. A., Dunn, J. J. & Bautz, E. K. F. Factor stimulating transcription by RNA polymerase. *Nature* **221**, 43–46 (1969). **Demonstrated that a core RNAP required a σ-factor for transcription.**
12. Gruber, T. M. & Gross, C. A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441–466 (2003).
13. 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).
14. Campbell, E. A. et al. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell* **9**, 527–539 (2002).
15. Callaci, S. & Heyduk, T. Conformation and DNA binding properties of a single-stranded DNA binding region of σ<sup>70</sup> subunit from *Escherichia coli* RNA polymerase are modulated by an interaction with the core enzyme. *Biochemistry* **37**, 3312–3320 (1998).
16. Murakami, K. S., Masuda, S. & Darst, S. A. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* **296**, 1280–1284 (2002).
17. Vassilyev, D. G. et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**, 712–719 (2002). **High-resolution structure of RNAP holoenzyme showing the housekeeping σ-factor domain architecture in context with RNAP.**
18. Danson, A. E., Jovanovic, M., Buck, M. & Zhang, X. Mechanisms of σS-dependent transcription initiation and regulation. *J. Mol. Biol.* **431**, 3960–3974 (2019).
19. Hawley, D. K. & McClure, W. R. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**, 2237–2255 (1983).
20. Simpson, R. B. The molecular topography of RNA polymerase-promoter interaction. *Cell* **18**, 277–285 (1979).
21. Ross, W. et al. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* **262**, 1407–1413 (1993). **Describes the discovery of upstream elements.**
22. Barne, K. A., Bown, J. A., Busby, S. J. & Minchin, S. D. Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J.* **16**, 4034–4040 (1997).
23. Haugen, S. P. et al. rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase. *Cell* **125**, 1069–1082 (2006).
24. Zhang, Y. et al. Structural basis of transcription initiation. *Science* **338**, 1076–1080 (2012).
25. Feklistov, A. & Darst, S. A. Structural basis for promoter –10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**, 1257–1269 (2011). **First high-resolution view describing the interaction of the σ-factor with the –10 element.**
26. Hubin, E. A., Lilic, M., Darst, S. A. & Campbell, E. A. Structural insights into the mycobacterial transcription initiation complex from analysis of X-ray crystal structures. *Nat. Commun.* **8**, 16072 (2017).
27. Murakami, K. S. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285–1290 (2002).
28. 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). **High-resolution crystal structure of a transcription initiation complex with a complete bubble and nascent RNA.**
29. Zuo, Y. & Steitz, T. A. Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol. Cell* **58**, 534–540 (2015). **First crystal structure at mid-range resolution of a transcription initiation complex with a complete bubble and nascent RNA.**
30. Cortes, T. et al. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in mycobacterium tuberculosis. *Cell Rep.* **5**, 1121–1131 (2013).
31. Shell, S. S. et al. Leaderless transcripts and small proteins are common features of the mycobacterial translational landscape. *PLoS Genet.* **11**, e1005641 (2015).
32. Chakraborty, A. et al. Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**, 591–595 (2012). **Used single-molecule fluorescence energy transfer experiments to detect the motions of the RNAP clamp.**
33. Duchi, D., Mazumder, A., Malinen, A. M., Ebright, R. H. & Kapanidis, A. N. The RNA polymerase clamp interconverts dynamically among three states and is stabilized in a partly closed state by ppGpp. *Nucleic Acids Res.* **46**, 7284–7295 (2018).
34. Feklistov, A. et al. RNA polymerase motions during promoter melting. *Science* **356**, 863–866 (2017).
35. Ruff, E. F. et al. *E. coli* RNA polymerase determinants of open complex lifetime and structure. *J. Mol. Biol.* **427**, 2435–2450 (2015).
36. Saecker, R. M., Record, M. T. & deHaseth, P. L. Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *J. Mol. Biol.* **412**, 754–771 (2011).
37. Buc, H. & McClure, W. R. Kinetics of open complex formation between *Escherichia coli* RNA polymerase and the lac UV5 promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry* **24**, 2712–2723 (1985). **Early evidence that transcription initiation occurs in multiple steps.**
38. Roe, J.-H., Burgess, R. R. & Record, M. T. Temperature dependence of the rate constants of the *Escherichia coli* RNA polymerase-λPR promoter interaction: assignment of the kinetic steps corresponding to protein conformational change and DNA opening. *J. Mol. Biol.* **184**, 441–453 (1985). **Early kinetic probing of the transcription rate constants that set the stage for future kinetic studies of promoter melting by RNAP.**
39. Ruff, E. F., Record, M. T. & Artsimovitch, I. Initial events in bacterial transcription initiation. *Biomolecules* **5**, 1035–1062 (2015).
40. Hubin, E. A. et al. Structure and function of the mycobacterial transcription initiation complex with the essential regulator RpbA. *eLife* **6**, e22520 (2017). **One of the first structures of mycobacterial RNAP with detailed quantitative analysis comparing the kinetics of open promoter formation by mycobacteria and *E. coli* RNAPs and the effects of transcription factors on this process.**
41. Rammohan, J. et al. CarD stabilizes mycobacterial open complexes via a two-tiered kinetic mechanism. *Nucleic Acids Res.* **43**, 3272–3285 (2015).
42. 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).
43. Whipple, F. W. & Sonenshein, A. L. Mechanism of initiation of transcription by *Bacillus subtilis* RNA polymerase at several promoters. *J. Mol. Biol.* **223**, 399–414 (1992).
44. Schroeder, L. A. & deHaseth, P. L. Mechanistic differences in promoter DNA melting by *Thermus aquaticus* and *Escherichia coli* RNA polymerases. *J. Biol. Chem.* **280**, 17422–17429 (2005).
45. Davis, E., Chen, J., Leon, K., Darst, S. A. & Campbell, E. A. Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. *Nucleic Acids Res.* **43**, 433–445 (2015).
46. Boyaci, H., Chen, J., Jansen, R., Darst, S. A. & Campbell, E. A. Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature* **565**, 382–385 (2019). **First structural snapshot of an RNAP promoter melting intermediate.**
47. Boyaci, H. et al. Fidaxomicin jams *Mycobacterium tuberculosis* RNA polymerase motions needed for initiation via RpbA contacts. *eLife* **7**, e34823 (2018).
48. Chen, J. et al. Stepwise promoter melting by bacterial RNA polymerase. *Mol. Cell* **78**, 275–288.e6 (2020). **Captured structural snapshots of promoter melting intermediates that span the RNAP closed complex to the final transcriptionally competent open promoter complex.**
49. McClure, W. R., Cech, C. L. & Johnston, D. E. A steady state assay for the RNA polymerase initiation reaction. *J. Biol. Chem.* **253**, 8941–8948 (1978).
50. Carpousis, A. J. & Gralla, J. D. Cycling of ribonucleic acid polymerase to produce oligonucleotides during initiation in vitro at the lac UV5 promoter. *Biochemistry* **19**, 3245–3253 (1980).
51. Goldman, S. R., Ebright, R. H. & Nickels, B. E. Direct detection of abortive RNA transcripts in vivo. *Science* **324**, 927 (2009).
52. Hsu, L. M., Vo, N. V., Kane, C. M. & Chamberlin, M. J. In vitro studies of transcript initiation by *Escherichia coli* RNA polymerase. 1. RNA chain initiation, abortive initiation, and promoter escape at three bacteriophage promoters. *Biochemistry* **42**, 3777–3786 (2003).
53. Hsu, L. M. Monitoring abortive initiation. *Methods* **47**, 25–36 (2009).
54. Kapanidis, A. N. et al. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**, 1144–1147 (2006).
55. Revyakin, A., Liu, C., Ebright, R. H. & Strick, T. R. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. *Science* **314**, 1139–1143 (2006). **Single-molecule studies revealing the role of scrunching in abortive initiation and promoter escape.**
56. Winkelmann, J. T. et al. Crosslink mapping at amino acid-base resolution reveals the path of scrunched DNA in initial transcribing complexes. *Mol. Cell* **59**, 768–780 (2015).
57. Duchi, D. et al. RNA polymerase pausing during initial transcription. *Mol. Cell* **63**, 939–950 (2016).
58. Kulbachinskii, A. & Mustaev, A. Region 3.2 of the σ subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *J. Biol. Chem.* **281**, 18273–18276 (2006).
59. Pupov, D., Kuzin, I., Bass, I. & Kulbachinskii, A. Distinct functions of the RNA polymerase σ subunit region 3.2 in RNA priming and promoter escape. *Nucleic Acids Res.* **42**, 4494–4504 (2014).
60. Samanta, S. & Martin, C. T. Insights into the mechanism of initial transcription in *Escherichia coli* RNA polymerase. *J. Biol. Chem.* **288**, 31993–32003 (2013).
61. Li, L., Molodtsov, V., Lin, W., Ebright, R. H. & Zhang, Y. RNA extension drives a stepwise displacement of an initiation-factor structural module in initial transcription. *Proc. Natl. Acad. Sci. USA* **117**, 5801–5809 (2020).
62. Dulin, D. et al. Pausing controls branching between productive and non-productive pathways during initial transcription in bacteria. *Nat. Commun.* **9**, 1478 (2018).
63. Jensen, D., Manzano, A. R., Rammohan, J., Stallings, C. L. & Galburd, E. A. CarD and RpbA modify the kinetics of initial transcription and slow promoter escape of the *Mycobacterium tuberculosis* RNA polymerase. *Nucleic Acids Res.* **47**, 6685–6698 (2019).
64. Mooney, R. A., Darst, S. A. & Landick, R. Sigma and RNA polymerase: an on-again, off-again relationship? *Mol. Cell* **20**, 335–345 (2005).
65. Harden, T. T. et al. Bacterial RNA polymerase can retain σ70 throughout transcription. *Proc. Natl. Acad. Sci. USA* **113**, 602–607 (2016).

66. Mukhopadhyay, J. et al. Translocation of  $\sigma$ 70 with RNA polymerase during transcription: fluorescence resonance energy transfer assay for movement relative to DNA. *Cell* **106**, 453–463 (2001).
67. Bar-Nahum, G. & Nudler, E. Isolation and characterization of sigma(70)-retaining transcription elongation complexes from Escherichia coli. *Cell* **106**, 443–451 (2001).
68. Kapanidis, A. N. et al. Retention of transcription initiation factor sigma70 in transcription elongation: single-molecule analysis. *Mol. Cell* **20**, 347–356 (2005).
69. Mooney, R. A. & Landick, R. Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. *Genes Dev.* **17**, 2839–2851 (2003).
70. Ring, B. Z., Yarnell, W. S. & Roberts, J. W. Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell* **86**, 485–493 (1996).
71. Davis, M. C., Kesthely, C. A., Franklin, E. A. & MacLellan, S. R. The essential activities of the bacterial sigma factor. *Can. J. Microbiol.* **63**, 89–99 (2017).
72. Gross, C. A. et al. The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.* **63**, 141–155 (1998).
73. Losick, R. & Pero, J. Cascades of sigma factors. *Cell* **25**, 582–584 (1981).
74. Paget, M. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biochemicals* **5**, 1245–1265 (2015).
75. Staroń, A. et al. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF)  $\sigma$  factor protein family: ECF  $\sigma$  factor classification. *Mol. Microbiol.* **74**, 557–581 (2009). **A comparative genomic analysis revealing the range and diversity of the ECF  $\sigma$ -factor in bacteria.**
76. Hughes, K. T. & Mathee, K. The anti-sigma factors. *Annu. Rev. Microbiol.* **52**, 231–286 (1998).
77. Campbell, E. A. et al. A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Mol. Cell* **27**, 793–805 (2007).
78. Maillard, A. P. et al. The crystal structure of the anti- $\sigma$  factor CnrY in complex with the  $\sigma$  factor CnrH shows a new structural class of anti- $\sigma$  factors targeting extracytoplasmic function  $\sigma$  factors. *J. Mol. Biol.* **426**, 2313–2327 (2014).
79. Sineva, E., Savkina, M. & Ades, S. E. Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors. *Curr. Opin. Microbiol.* **36**, 128–137 (2017).
80. Campbell, E. A. et al. Crystal structure of the Bacillus stearothermophilus anti-sigma factor SpollAB with the sporulation sigma factor sigmaF. *Cell* **108**, 795–807 (2002).
81. Campbell, E. A. et al. Crystal structure of Escherichia coli sigmaE with the cytoplasmic domain of its anti-sigma RseA. *Mol. Cell* **11**, 1067–1078 (2003).
82. Sorenson, M. K., Ray, S. S. & Darst, S. A. Crystal structure of the flagellar  $\sigma$ /anti- $\sigma$  complex  $\sigma$ 28/FliM reveals an intact  $\sigma$  factor in an inactive conformation. *Mol. Cell* **14**, 127–138 (2004).
83. Schumacher, M. A. et al. The crystal structure of the RsbN– $\sigma$ BldN complex from Streptomyces venezuelae defines a new structural class of anti- $\sigma$  factor. *Nucleic Acids Res.* **46**, 7405–7417 (2018).
84. Gallagher, K. A. et al. c-di-GMP arms an anti- $\sigma$  to control progression of multicellular differentiation in streptomycetes. *Mol. Cell* **77**, 586–599.e6 (2020).
85. Shukla, J., Gupta, R., Thakur, K. G., Gokhale, R. & Gopal, B. Structural basis for the redox sensitivity of the Mycobacterium tuberculosis SigK-RskA  $\sigma$ -anti- $\sigma$  complex. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 1026–1036 (2014).
86. Devkota, S. R., Kwon, E., Ha, S. C., Chang, H. W. & Kim, D. Y. Structural insights into the regulation of *Bacillus subtilis* SigW activity by anti-sigma RsiW. *PLoS ONE* **12**, e0174284 (2017).
87. Herrou, J., Rotksoff, G., Luo, Y., Roux, B. & Crosson, S. Structural basis of a protein partner switch that regulates the general stress response of  $\alpha$ -proteobacteria. *Proc. Natl. Acad. Sci. USA* **109**, 7973–7973 (2012).
88. Campagne, S. et al. Structural basis for sigma factor mimicry in the general stress response of Alphaproteobacteria. *Proc. Natl. Acad. Sci. USA* **109**, E1405–E1414 (2012).
89. 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).
90. Patikoglou, G. A. et al. Crystal structure of the Escherichia coli regulator of sigma70, Rsd, in complex with sigma70 domain 4. *J. Mol. Biol.* **372**, 649–659 (2007).
91. Shi, J. et al. Structural basis of  $\sigma$  appropriation. *Nucleic Acids Res.* **47**, 9423–9432 (2019).
92. Hu, Y., Morichaud, Z., Chen, S., Leonetti, J.-P. & Brodin, K. Mycobacterium tuberculosis RbpA protein is a new type of transcriptional activator that stabilizes the  $\sigma^70$ -containing RNA polymerase holoenzyme. *Nucleic Acids Res.* **40**, 6547–6557 (2012).
93. Bortoluzzi, A. et al. *Mycobacterium tuberculosis* RNA polymerase-binding protein A (RbpA) and its interactions with sigma factors. *J. Biol. Chem.* **288**, 14438–14450 (2013).
94. Tabib-Salazar, A. et al. The actinobacterial transcription factor RbpA binds to the principal sigma subunit of RNA polymerase. *Nucleic Acids Res.* **41**, 5679–5691 (2013).
95. Vishwakarma, R. K. et al. Single-molecule analysis reveals the mechanism of transcription activation in *M. tuberculosis*. *Sci. Adv.* **4**, eaao5498 (2018).
96. 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).
97. Haakonsen, D. L., Yuan, A. H. & Laub, M. T. The bacterial cell cycle regulator CcrA is a  $\sigma$ 70 cofactor that drives gene expression from a subset of methylated promoters. *Genes Dev.* **29**, 2272–2286 (2015).
98. Holtzendorff, J. et al. Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science* **304**, 983–987 (2004).
99. Wu, X. et al. Structural insights into the unique mechanism of transcription activation by Caulobacter crescentus CcrA. *Nucleic Acids Res.* **46**, 3245–3256 (2018).
100. Gaal, T., Mandel, M. J., Silhavy, T. J. & Course, R. L. Crl facilitates RNA polymerase holoenzyme formation. *J. Bacteriol.* **188**, 7966–7970 (2006).
101. Banta, A. B. et al. Key features of  $\sigma^S$  required for specific recognition by Crl, a transcription factor promoting assembly of RNA polymerase holoenzyme. *Proc. Natl. Acad. Sci. USA* **110**, 15955–15960 (2013).
102. Cartagena, A. J. et al. Structural basis for transcription activation by Crl through tethering of  $\sigma$ S and RNA polymerase. *Proc. Natl. Acad. Sci. USA* **116**, 18923–18927 (2019).
103. Xu, J. et al. Crl activates transcription by stabilizing active conformation of the master stress transcription initiation factor. *eLife* **8**, e50928 (2019).
104. Wasserman, K. M. & Storz, G. 6S RNA regulates E. coli RNA polymerase activity. *Cell* **101**, 613–623 (2000).
105. Trottochaud, A. E. & Wasserman, K. M. A highly conserved 6S RNA structure is required for regulation of transcription. *Nat. Struct. Mol. Biol.* **12**, 313–319 (2005).
106. Barrick, J. E. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *RNA* **11**, 774–784 (2005).
107. Wasserman, K. M. In *Regulating with RNA in Bacteria and Archaea* Ch. 20 (eds Storz, G. & Papenfort, K.) 355–367 (ASM Press, 2019).
108. Wasserman, K. M. & Saecker, R. M. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. *Science* **314**, 1601–1603 (2006). **Discovery that 6S RNA serves as a functional template for transcription initiation by RNAP.**
109. Chen, J. et al. 6S RNA mimics B-form DNA to regulate Escherichia coli RNA polymerase. *Mol. Cell* **68**, 388–397.e6 (2017).
110. Lane, W. J. & Darst, S. A. Molecular evolution of multisubunit RNA polymerases: sequence analysis. *J. Mol. Biol.* **395**, 671–685 (2010).
111. Lane, W. J. & Darst, S. A. Molecular evolution of multisubunit RNA polymerases: structural analysis. *J. Mol. Biol.* **395**, 686–704 (2010).
112. Artsimovitch, I., Svetlov, V., Murakami, K. S. & Landick, R. Co-overexpression of *Escherichia coli* RNA polymerase subunits allows isolation and analysis of mutant enzymes lacking lineage-specific sequence insertions. *J. Biol. Chem.* **278**, 12344–12355 (2003).
113. Chen, J. et al. E. coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. *eLife* **8**, e49375 (2019).
114. Lin, W. et al. Structural basis of *Mycobacterium tuberculosis* transcription and transcription inhibition. *Mol. Cell* **66**, 169–179.e8 (2017).
115. Browning, D. F. & Busby, S. J. W. Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **14**, 638 (2016).
116. Bergkessel, M. et al. The dormancy-specific regulator, SutA, is intrinsically disordered and modulates transcription initiation in *Pseudomonas aeruginosa*. *Mol. Microbiol.* **112**, 992–1009 (2019).
117. Molodtsov, V. et al. Allosteric effector ppGpp potentiates the inhibition of transcript initiation by DksA. *Mol. Cell* **69**, 828–839.e5 (2018).
118. Rutherford, S. T., Villers, C. L., Lee, J.-H., Ross, W. & Course, R. L. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes Dev.* **23**, 236–248 (2009).
119. Copalkrishnan, S., Ross, W., Chen, A. Y. & Course, R. L. TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp. *Proc. Natl. Acad. Sci. USA* **114**, E5539–E5548 (2017).
120. Srivastava, D. B. et al. Structure and function of CarD, an essential mycobacterial transcription factor. *Proc. Natl. Acad. Sci. USA* **110**, 12619–12624 (2013).
121. Bae, B. et al. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *eLife* **4**, e08505 (2015).
122. Venugopal, A. A. & Johnson, S. Fidaxomicin: a novel macrocyclic antibiotic approved for treatment of *Clostridium difficile* infection. *Clin. Infect. Dis.* **54**, 568–574 (2012).
123. Srivastava, A. et al. New target for inhibition of bacterial RNA polymerase: ‘switch region’. *Curr. Opin. Microbiol.* **14**, 532–543 (2011).
124. Lin, W. et al. Structural basis of transcription inhibition by fidaxomicin (lipiarmycin A3). *Mol. Cell* **70**, 60–70.e16 (2018).
125. Belogurov, G. A. et al. Transcription inactivation through local refolding of the RNA polymerase structure. *Nature* **457**, 352–355 (2009).
126. Mukhopadhyay, J. et al. The RNA polymerase “switch region” is a target for inhibitors. *Cell* **135**, 295–307 (2008).
127. Peek, J. et al. Rifamycin congeners kanglemycins are active against rifampicin-resistant bacteria via a distinct mechanism. *Nat. Commun.* **9**, 4147 (2018).
128. Mosaei, H. et al. Mode of action of kanglemycin A, an ansamycin natural product that is active against rifampicin-resistant *Mycobacterium tuberculosis*. *Mol. Cell* **72**, 263–274.e5 (2018).
129. Maffoli, S. I. et al. Antibacterial nucleoside-analog inhibitor of bacterial RNA polymerase. *Cell* **169**, 1240–1248.e23 (2017).
130. Zhang, Y. et al. GE23077 binds to the RNA polymerase ‘i’ and ‘i-1’ sites and prevents the binding of initiating nucleotides. *eLife* **3**, e02450 (2014).
131. McClure, W. R. On the mechanism of streptolydigin inhibition of Escherichia coli RNA polymerase. *J. Biol. Chem.* **255**, 1610–1616 (1980).
132. Cassani, G., Burgess, R. R., Goodman, H. M. & Gold, L. Inhibition of RNA polymerase by streptolydigin. *Nat. New Biol.* **230**, 197–200 (1971).
133. Siddhikol, C., Erbstösser, J. W. & Weisblum, B. Mode of action of streptolydigin. *J. Bacteriol.* **99**, 151–155 (1969).
134. Temriakov, D. et al. Structural basis of transcription inhibition by antibiotic streptolydigin. *Mol. Cell* **19**, 655–666 (2005).
135. Tuske, S. et al. Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell* **122**, 541–552 (2005).
136. Artsimovitch, I., Chu, C., Lynch, A. S. & Landick, R. A new class of bacteria RNA polymerase inhibitor affects nucleotide addition. *Science* **302**, 650–654 (2003).
137. Bae, B. et al. CBR antimicrobials inhibit RNA polymerase via at least two bridge-helix cap-mediated effects on nucleotide addition. *Proc. Natl. Acad. Sci. USA* **112**, E4178 (2015).
138. Feng, Y. et al. Structural basis of transcription inhibition by CBR hydroxamides and CBR pyrazoles. *Structure* **23**, 1470–1481 (2015).
139. Degen, D. et al. Transcription inhibition by the depsipeptide antibiotic salinamide A. *eLife* **3**, e02451 (2014).
140. McClure, W. R. & Cech, C. L. On the mechanism of rifampicin inhibition of RNA synthesis. *J. Biol. Chem.* **253**, 8949–8956 (1978).
141. Campbell, E. A. E. A. et al. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* **104**, 901–912 (2001). **Structure of RNAP bound to the clinically used antibiotic rifampicin revealing the basis of inhibition.**

142. Campbell, E. A. et al. Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* **24**, 674–682 (2005).
143. Darst, S. A., Kubalek, E. W. & Kornberg, R. D. Three-dimensional structure of Escherichia coli RNA polymerase holoenzyme determined by electron crystallography. *Nature* **340**, 730–732 (1989).
144. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1876–1882 (2001).
145. Murakami, K. S. X-ray crystal structure of Escherichia coli RNA polymerase  $\sigma$ 70 holoenzyme. *J. Biol. Chem.* **288**, 9126–9134 (2013).
146. Bae, B. et al. Phage T7 Gp2 inhibition of Escherichia coli RNA polymerase involves misappropriation of 70 domain 1.1. *Proc. Natl Acad. Sci. USA* **110**, 19772–19777 (2013).
147. Zuo, Y., Wang, Y. & Steitz, T. A. The mechanism of E. coli RNA polymerase regulation by ppGpp is suggested by the structure of their complex. *Mol. Cell* **50**, 430–436 (2013).
148. Plaschka, C. et al. Transcription initiation complex structures elucidate DNA opening. *Nature* **533**, 353–358 (2016).
149. Kang, J. Y. et al. Structural basis of transcription arrest by coliphage HK022 Nun in an Escherichia coli RNA polymerase elongation complex. *eLife* **6**, e25478 (2017).
150. Liu, B., Hong, C., Huang, R. K., Yu, Z. & Steitz, T. A. Structural basis of bacterial transcription activation. *Science* **358**, 947–951 (2017).
151. Stevens, A. An inhibitor of host sigma-stimulated core enzyme activity that purified with DNA-dependent RNA polymerase of E. coli following T4 phage infection. *Biochem. Biophys. Res. Commun.* **54**, 488–493 (1973).
152. Wiggs, J. L., Gilman, M. Z. & Chamberlin, M. J. Heterogeneity of RNA polymerase in *Bacillus subtilis*: evidence for an additional sigma factor in vegetative cells. *Proc. Natl Acad. Sci. USA* **78**, 2762–2766 (1981).
153. Haldenwang, W. G. & Losick, R. A modified RNA polymerase transcribes a cloned gene under sporulation control in *Bacillus subtilis*. *Nature* **282**, 256–260 (1979).
154. Buttner, M. J., Smith, A. M. & Bibb, M. J. At least three different RNA polymerase holoenzymes direct transcription of the agarase gene (*dagA*) of *Streptomyces coelicolor* A3(2). *Cell* **52**, 599–607 (1988).
155. Erickson, J. W. & Gross, C. A. Identification of the sigma E subunit of Escherichia coli RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression. *Genes Dev.* **3**, 1462–1471 (1989).
156. Lonetto, M. A., Brown, K. L., Rudd, K. E. & Buttner, M. J. Analysis of the *Streptomyces coelicolor* *sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase  $\sigma$  factors involved in the regulation of extracytoplasmic functions. *Proc. Natl Acad. Sci. USA* **91**, 7573–7577 (1994).
157. Malhotra, A., Severinova, E. & Darst, S. A. Crystal structure of a  $\sigma$ 70 subunit fragment from E. coli RNA polymerase. *Cell* **87**, 127–136 (1996).
158. Rhodius, V. A. et al. Design of orthogonal genetic switches based on a crosstalk map of  $\sigma$ s, anti- $\sigma$ s, and promoters. *Mol. Syst. Biol.* **9**, 702 (2013).
159. Campagne, S., Marsh, M. E., Capitani, G., Vorholt, J. A. & Allain, F. H. T. Structural basis for -10 promoter element melting by environmentally induced sigma factors. *Nat. Struct. Mol. Biol.* **21**, 269–276 (2014).
160. Li, L., Fang, C., Zhuang, N., Wang, T. & Zhang, Y. Structural basis for transcription initiation by bacterial ECF  $\sigma$  factors. *Nat. Commun.* **10**, 1153 (2019).
161. Lin, W. et al. Structural basis of ECF- $\sigma$ -factor-dependent transcription initiation. *Nat. Commun.* **10**, 710 (2019).
162. Li, S. et al. Structural basis for the recognition of MuCA by MucB and AlgU in *Pseudomonas aeruginosa*. *FEBS J.* **286**, 4982–4994 (2019).

**Acknowledgements**

The authors thank R. Landick, S. Darst and R. Froom for helpful discussions and copyediting. They apologize to colleagues whose work could not be cited owing to the scope and space limits of the Review. The authors are grateful for support from NIH grant 2-R01 GM114450 (E.A.C.) and the Charles H. Revson Foundation award CEN5650030 (H.B.).

**Author contributions**

J.C., H.B. and E.A.C. researched data for article, discussed the content and wrote the article.

**Competing interests**

The authors declare no competing interests.

**Publisher's note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2020