Transcription activation in bacteria

Busby, Stephen J W

DOI:
10.1099/mic.0.000783

Document Version
Peer reviewed version

Citation for published version (Harvard):
Busby, SJW 2019, 'Transcription activation in bacteria: ancient and modern', *Microbiology*, vol. 165, no. 4, pp. 386-395. https://doi.org/10.1099/mic.0.000783

Marjory Stephenson Lecture 2017 (revised 26th January, 2019)

Transcription activation in bacteria: ancient and modern

Stephen J. W. Busby

School of Biosciences & Institute of Microbiology and Infection, University of Birmingham, Birmingham B15 2TT, U.K.

Abstract

Activation of the transcription of genes is central to many processes of adaptation and differentiation in bacteria. Here, I review the molecular mechanisms by which transcription factors can activate the initiation of specific transcripts at bacterial promoters. The story is presented in the context of Marjory Stephenson's pioneering work on enzymatic adaptation in bacteria, and sets the different mechanisms in the greater context of how transcription regulatory mechanisms evolved.

Keywords: Bacterial Transcription; Promoters; Activators; DNA-dependent RNA polymerase

Introduction

Most bacteria rely heavily on transcription activation to ensure the regulated expression of key genes. Here, I will give an overview of research in laboratories worldwide that has sought to understand the molecular mechanisms behind this process. Memorial lectures give the opportunity to present a story and connect it to the past. For transcription activation, and the Marjory Stephenson Lecture, this is an easy task. I shall try to show that, to reach a full understanding of transcription activation (or, in fact, any biological phenomenon), one needs to take a historical perspective.

Marjory Stephenson was a staff member in the Department of Biochemistry at the University of Cambridge for nearly 30 years, up to her death in 1948. Her major achievements concerned the identification of various microbial enzyme-mediated reactions in cell free extracts, and her findings laid the foundations for our subsequent understanding of microbial metabolism [1]. In addition, she noted that, for many of the enzymes that she studied, their activity depended critically on the growth conditions of the host microbe. The biological significance of this, which became known as 'enzymatic adaptation', was easy to appreciate, but understanding the molecular mechanism posed a tougher problem, and this became a major aim of Jacques Monod's research from 1940 onwards at the Pasteur Institute in Paris. It is worth noting that Monod's starting hypothesis for the *Escherichia coli* lactose (*lac*) operon was that the adaptation occurred at the level of the enzyme β-galactosidase, as its activity increases when cells are grown in a lactose-containing medium. However, Monod, working with his mentor Andre Lwoff, and colleague, Francois Jacob, came to realise that, in this case, adaptation took place at the level of the initiation of transcription. The tale of this Nobel Prize-winning work has been told many times [2], and includes the establishment of the operon model, where groups of bacterial genes are transcribed together, starting at a promoter, that can be regulated by a repressor that responds to an environmental signal. Thus, the

E. coli lac operon repressor (LacI) was the first of thousands of transcription repressors to be identified, and this was followed shortly after by the discovery that some promoters are regulated by an activating factor rather than a repressor. Our current understanding is that most strains of *E. coli* contain hundreds of transcription repressors and activators and, together, they make up an effective network supporting enzymatic adaptation [3]. At many promoters, regulation is due to the combined action of several transcription factors. Hence, at the promoter that drives the expression of the *E. coli lac* operon, transcript initiation is controlled by both a repressor (LacI) and an activator (the cyclic AMP receptor protein, CRP, also known as CAP or CGA). Both have been the subject of intensive research and much of what we understand about transcription activation and repression has emerged from their study [4, 5]. Here I focus on the mechanism of activation.

Transcript initiation in bacteria and its activation

Bacterial transcripts are synthesised from nucleoside triphosphate precursors by the multisubunit DNA-dependent RNA polymerase (RNAP). The crucial subunit in orchestrating transcript initiation at promoters is the sigma subunit that consists of independently folding domains (numbered 1-4) joined by flexible linkers, with each domain playing a different role [6]. Promoter recognition is the first step in transcript initiation and this involves the interaction of sigma Domain 3 with the promoter extended -10 element, and sigma Domain 4 with the promoter -35 element (Figure 1). At some promoters, further interactions between the C-terminal domains of the RNA polymerase alpha subunits (αCTD) and upstream sequences also contribute to the formation of the initial complex between RNAP and promoter DNA. This is known as the 'closed complex' because the two strands of the DNA duplex in the promoter remain base-paired. For transcript initiation to occur, the promoter DNA strands around the transcript start must be separated to position the template strand in the catalytic site of the RNAP. To assure this, Domain 2 of the RNAP sigma subunit drives local DNA strand separation by making specific interactions with the non-template strand of the promoter -10 element [7-9]. This results in the promoter DNA template strand being able to access the RNAP catalytic site, where it is positioned by a combination of other interactions, so that template-directed joining of the two initiating nucleoside triphosphates occurs. This RNAP-promoter complex, known as the 'open complex' is a prerequisite for promoter-driven transcript formation. Both open and closed complexes are a nexus for regulation involving both transcription factors and ligands, so, for example, many repressors (such as LacI) simply block RNAP access to key promoter elements. The crucial point is that most sigma factors are competent to enable transcript initiation by RNAP at promoters because, in the closed complex, sigma Domain 2 is positioned so that it can spontaneously drive open complex formation (the exception is with sigma factors related to Sigma-54, see later).

The situation, however, is more complex for promoters where transcript initiation depends on an activator protein. At these promoters, RNAP is unable to initiate transcript formation due to a shortcoming, usually either in the formation of the closed complex or the positioning of sigma Domain 2, so that a stable open complex is unable to form. Early experiments suggested that many activator-dependent promoters possessed elements whose base sequence corresponds poorly to the consensus, and this led to the contention that activators compensate for 'defects' in the base

sequences at target promoters. Hence activators might work by inducing a conformation change in promoter DNA to facilitate its interaction with RNAP, or by contacting RNAP directly resulting in its recruitment to or stabilisation at the target promoter [10]. This issue was first resolved from studies with *E. coli* CRP at the *lac* promoter.

Early studies had shown that CRP binding to target sites was dependent on triggering by the ligand cyclic AMP (cAMP), and that the cAMP-CRP complex was necessary and sufficient for RNAP to initiate transcription at the *E. coli lac* operon promoter [11]. A high resolution structure of DNA-bound cAMP-CRP showed that binding induced a sharp bend at target DNA, suggesting that activation might be driven by induced DNA distortion [12]. However, the discovery of mutations in the *crp* gene, which resulted in single amino acid substitutions in CRP that stopped activation at the *lac* promoter, without affecting CRP expression, CRP folding, cAMP binding or DNA binding and bending, argued that DNA distortion was not sufficient for activation. The observation that these substitutions clustered at a single locus, a β-turn on the CRP surface, just adjacent to the DNA-binding helix-turn-helix, identified a surface determinant (known as Activating Region 1, AR1) that was later shown directly to interact with a cognate determinant in αCTD [13]. The cAMP-CRP complex is functional as a dimer, but, for transcription activation at the *E. coli lac* promoter, AR1 function is required only in the 'downstream' CRP subunit (Figure 2). Structural analyses of the interactions of cAMP-CRP and αCTD at the *lac* promoter argue that the CRP-RNAP interaction appears to act as 'Velcro' and no major conformation change in either partner occurs [14-16]. Hence, the AR1-αCTD interaction serves to recruit RNA polymerase to the *lac* promoter, thereby positioning the RNAP sigma factor such that it is suitably placed to drive open complex formation.

Class I and Class II activation

At the *E. coli lac* promoter, the DNA site for CRP is centred between base pairs -61 and 62 upstream from the transcript start, and thus AR1 of the downstream subunit of CRP interacts with αCTD that is 'sandwiched' between CRP and Domain 4 of the RNAP sigma subunit [16, 17]. A similar mechanism was proposed for other CRP-dependent promoters where the DNA site for CRP was located further upstream [13]. At these promoters, it is proposed that the flexible linker joining the RNAP alpha subunit's N-terminal (αNTD) to αCTD facilitates the same AR1-αCTD interaction. However, at many CRP-dependent promoters, the DNA site for CRP is closer to the promoter and is adjacent to the promoter -35 element, precluding a mechanism whereby the downstream subunit of the CRP dimer could interact with αCTD. Genetic and biochemical analysis showed that, at these promoters, two alternative activating regions in the downstream CRP subunit, known as Activating Region 2 (AR2) and Activating Region 3 (AR3), make direct interactions with αNTD and sigma Domain 4 respectively [13, 15]. Moreover, at many such promoters αCTD binds upstream of the CRP dimer and interacts with the upstream subunit via AR1 (Figure 2).

The study of different CRP-dependent promoters reveals two distinct modes of activation, often referred to as Class I and Class II [15-18]. At promoters where CRP binds upstream of the different promoter elements, activation involves interaction between AR1 in the downstream submit of the CRP dimer and αCTD (Class I). At promoters where bound CRP abuts the promoter -35 element,

activation can involve AR2 and/or AR3 in the downstream subunit of the CRP dimer and AR1 in the upstream subunit (Class II). Studies with other systems showed that these modes apply to many other transcription factors, including other members of the CRP family of transcription factors, and members of the AraC family. Other transcription factors function either by a Class I mechanism or by a Class II mechanism, but, unlike CRP, are unable to do both [19]. For any factor, its potential to activate promoters with different architectures appears to depend on its activating regions.

It is especially interesting to compare *E. coli* CRP with *E. coli* FNR, which is responsible for transcription activation in response to oxygen starvation and shares sequence and structural features with CRP [20]. Although equivalents of AR1, AR2 and AR3 can be identified in FNR, AR3 is the dominant activating region, unlike for CRP where AR1 and AR2 dominate, with AR3 playing but a minor role. Thus, in *E. coli*, most FNR-dependent promoters have a Class II architecture.

The discovery of the simple Class I and Class II mechanisms of activation provide an explanation for the organisation of many more complex promoters where activation is co-dependent on two different transcription factors [19, 21]. Hence, in many cases, one factor functions by a Class I mechanism whilst the other uses the Class II option with the molecular Velcro contributed by each activator being pooled to assure RNAP recruitment. At other promoters, co-activation is assured by two Class I activators (Figure 2). Hence, activating regions appear to have evolved as determinants on the surface of transcription factors in order to equip these factors to manage activation by RNAP recruitment at promoters with particular architectures.

Activation by anti-repression

An alternative mechanism of transcription activation was discovered from analysis of the *E. coli nir* operon promoter that regulates the expression of an NADH-dependent nitrite reductase [22]. Transcription of this operon is triggered by two signals, oxygen starvation, mediated by FNR, and the presence of nitrate or nitrite ions, mediated by NarL, a typical Response-Regulator family transcription factor. FNR was shown to bind to a single target site, abutting the promoter -35 element, and found to activate the *nir* operon promoter by a Class II mechanism, whilst NarL was shown to bind to a single target site upstream. Although, *nir* operon expression is totally dependent on NarL, the deletion of DNA sequences upstream of the DNA site for NarL relieved the requirement for NarL and resulted in a simple Class II FNR-dependent promoter. This led to the model (Figure 3) in which upstream-bound nucleoid-associated proteins (NAPs, in this case IHF and Fis) suppress FNR-dependent activation and the role of NarL is to counteract this suppression [23]. Our studies identified two DNA sites for IHF and showed that NarL displaces IHF from the promoter-proximal site (IHF-I). Similar activation by displacement of NAPs, or other repressors, has now been found at many locations, for example, in reversing gene silencing by H-NS [24].

NAPs are found in all bacteria and play an essential role in compacting bacterial DNA by bending and bridging mechanisms [25]. Just as for eukaryotic nucleosomes, it is easy to see how they have an overall negative effect on transcription. As bacteria evolved, and genome sizes grew, a balance must have developed between the conflicting needs of compacting genomic DNA and transcribing

it into RNA. Since both DNA compaction and DNA transcription are essential requirements for any bacterium (in contrast to transcription regulation that is optional), one can imagine an early bacterial world without much regulation, and the scene was thus set for 'ancient' transcription activation by certain proteins evolving to induce transcription by locally disrupting NAP-mediated DNA folding.

Detailed analysis of the action of IHF at the *E. coli nir* operon promoter has shown that, whilst IHF binding to the promoter-proximal IHF-I site is inhibitory, IHF binding to the promoter-distal IHF-II site stimulates promoter activity [26, 27]. This unexpected observation shows that, although most NAP binding is inhibitory for transcription, binding of certain NAPs at certain locations can be activatory at specific promoters.

Transcription activation in bacteria: the big picture

The crucial observation that it is Domain 2 of sigma factors that orchestrates the local opening of DNA at transcript start sites, which enables transcript initiation by RNAP, means that the overriding aim of any mechanism to positively promote transcript initiation at a promoter will be to position Domain 2 correctly with respect to a potential -10 element. Activators have evolved to facilitate this either by simply pushing NAPs, or other repressors, out of the way in order to free up promoter elements that can then be served by the RNAP, or by interacting with RNAP so as to hold it in place at the target promoter such that sigma Domain 2 is correctly positioned. Studies with CRP uncovered the Class I and Class II mechanisms for doing this, and a further way is seen at promoters dependent on MerR family activators where binding of the activator induces conformational changes in the promoter DNA that adjust the juxtaposition of the promoter 35 and -10 elements, again to assure correct positioning of sigma Domain 2 [28]. None of these mechanisms involve major conformational changes in either the activator or the RNAP as they are focussed on delivery of RNAP to the promoter [14-18]. However, these mechanisms are insufficient for activation at promoters that are served by RNAP that carries a sigma factor belonging to the Sigma-54 family. Recall that the vast majority of sigma factors are related to the *E. coli* housekeeping sigma factor, Sigma-70, and are equipped with a functional Domain 2, but most bacteria also carry a sigma factor related to *E. coli* Sigma-54, which is completely unrelated to Sigma-70, with a different domain structure [29]. Whilst RNAP carrying Sigma-54, or any related sigma factor, is able to recognise and bind at target promoters, it is unable independently to orchestrate open complex formation and transcript initiation (in contrast to RNAP containing a sigma factor related to Sigma-70). It has long been known that promoters served by RNAP containing Sigma-54 require special activators known as Enhancer Binding Proteins (EBPs). These activators contain a domain that couples ATP hydrolysis to drive interactions with RNAP-bound Sigma-54 at target promoters which result in open complex formation. Structural studies have now confirmed that Sigma-54 and Sigma-70 have little in common, shown that certain Sigma-54 domains actively impede open complex formation, and suggested that ATP-driven interactions between EBPs and RNAP at target promoters remodel the Sigma-54 thereby enabling transcript initiation by RNAP [30-33]. This 'active' mechanism for transcription activation contrasts with the 'passive' recruitment mechanisms at promoters served

by RNAP carrying a sigma factor related to Sigma-70 (Figure 4). However, at some Sigma-54-dependent promoters, in addition to an EBP, other activators may be required to assist recruitment of RNAP to the closed complex.

The discovery that some NAPs are competent for Class I or Class II activation at certain promoters argues that many (if not most) transcription factors may have evolved from NAPs and this argument is bolstered by the observation that many transcription factors display properties associated with NAPs [34]. A key finding, which resulted from a study of CRP binding across the *E. coli* K-12 chromosome using chromatin immunoprecipitation (ChIP), was that most binding targets for CRP are not associated with Class I or Class II transcriptional regulation [35]. Taken together with the high abundance of CRP in many *E. coli* strains, and CRP-induced DNA bending, it is easy to speculate that the original role of CRP was as a NAP, and its activation function at certain target promoters evolved later with the evolution of its activating regions. In another example, of the 20 documented binding targets for RutR in *E. coli* K-12, only one is associated with transcriptional regulation [36]. Hence, it is plausible that an 'ancient' era of transcription activation simply by disrupting NAP-mediated repression, was followed by a 'modern' era of NAP-derived transcription factors with activating regions (Figure 5).

Postscript: the rise and rise of bacterial Sigma-70-like factors

Whilst RNAP in each of the three kingdoms of life uses essentially the same mechanism to synthesise DNA-templated RNA, different strategies are used in each kingdom to manage site-specific initiation [37]. The seductive simplicity of the Operon Model belies that fact that these strategies have evolved, presumably from a starting scenario where *de novo* RNA synthesis initiated at random locations. The domain structure of sigma factors related to Sigma-70 tells us that the 'winning' strategy in bacteria involved both the recognition of specific sequence elements (using Domains 3 and 4) and local unwinding and sequestration of the non-template strand (using Domain 2). Remarkably, Sigma-54 uses different determinants both to recognise sequence elements at target promoters (Figure 4), and to drive local unwinding, and recent structural data suggest that the mechanism of unwinding involves separation of the two strands of the DNA duplex by a long helix acting a 'crowbar, rather than by sequestration of the non-template DNA strand [33]. In any case, the existence of alternative sigma factors related to Sigma-54 argues that evolution could well have tried many different ways to engineer site-specific transcript initiation, and maybe some of these are yet to be discovered. One can speculate that the persistence of Sigma-54-like factors in the face of the dominance of the Sigma-70 family is because of their near absolute dependence on an EBP activator, and such tight regulation must have conferred the advantage necessary for their survival. One of the major plusses of site-specific transcript initiation for any gene or set of genes is that it sets the scene for the evolution of gene specific regulation, involving either activators or repressors. Perhaps the most striking lesson to be learned from considering the action of different bacterial transcription factors is how distinct mechanisms evolved in response to the different steps of the transcription cycle, and according to the sigma factor involved.

Work in the author's laboratory has been funded principally by the UK BBSRC, the Wellcome Trust and the Leverhulme Trust. The data, models and ideas discussed here arise from so many colleagues, both in Birmingham and world-wide, that it is impossible to list all of them here. However, I am especially grateful to Doug Browning for drawing the figures and to Dave Grainger for checking the text. The author declares no conflicts of interest.

1. **Robertson M.** Marjory Stephenson, 1885-1948. *Biographical memoirs of Fellows of the Royal Society* 1949; DOI: 10.1098/rsbm.1949.0013.
2. **Ullmann A.** *Origins of Molecular Biology: A tribute to Jacques Monod.* Herndon, VA, USA: ASM Press; 2004.
3. **Santos-Zavaleta A, Sánchez-Pérez M, Salgado H, Velázquez-Ramírez DA, Gama-Castro S *et al*.** A unified resource for transcriptional regulation in *Escherichia coli* K-12 incorporating high-throughput-generated binding data into RegulonDB version 10.0. *BMC Biol* 2018; 16:91. doi:10.1186/s12915-018-0555-y
4. **Muller-Hill B.** *The* lac *Operon*. D-10785 Berlin: Walter de Gruyter & Co; 1996.
5. **Kolb A, Busby S, Buc H, Garges S, Adhya, S.** Transcriptional regulation by cyclic AMP and its receptor protein. *Annu Rev Biochem* 1993; 62:749-795.
6. **Feklistov A, Sharon BD, Darst SA, Gross CA.** Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu Rev Microbiol* 2014; 68:357-376.
7. **Feklistov A, Darst SA.** Structural basis for promoter -10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* 2011; 147:1257-1269.
8. **Zhang Y, Feng Y, Chatterjee S, Tuske S, Ho MX *et al*.** Structural basis of transcription initiation. *Science* 2012; 338:1076-1080.
9. **Zuo Y, Steitz TA.** Crystal structure of the *E. coli* transcription initiation complexes with a complete bubble. *Mol Cell* 2015; 58:534-540.
10. **Rhodius V, Busby S.** Positive activation of gene expression. *Curr Op Microbiol* 1998; 1:152-159.
11. **de Crombrugghe B, Busby S, Buc H.** Cyclic AMP receptor protein: role in transcription activation. *Science* 1984; 224:831-838.
12. **Schultz SC, Shields GC, Steitz TA.** Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 1991; 253:1001-1007.
13. **Busby S, Ebright RH.** Transcription activation by catabolite activator protein (CAP). *J Mol Biol* 1999; 293:199-213.
14. **Benoff B, Yang H, Lawson CL, Parkinson G, Liu J *et al*.** Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* 2002; 297:1562-1566.
15. **Lawson CL, Swigon D, Murakami KS, Darst SA, Berman HM *et al*.** Catabolite activator protein: DNA binding and transcription activation. *Curr Opin Struct Biol* 2004; 14:10-20.
16. **Liu B, Hong C, Huang RK, Yu Z, Steitz TA.** Structural basis of bacterial transcription activation. *Science* 2017; 358:947-951.

17. **Hudson BP, Quispe J, Lara-Gonzalez S, Kim Y, Berman HM *et al*.** Three-dimensional EM structure of an intact activator-dependent transcription initiation complex. *Proc Natl Acad Sci USA* 2009; 106:19830-19835.
18. **Feng Y, Zhang Y, Ebright RH.** Structural basis of transcription activation. *Science* 2016; 352:1330-1333.
19. **Lee DJ, Minchin SD, Busby SJ.** Activating transcription in bacteria. *Annu Rev Microbiol* 2012; 66:125-152.
20. **Browning D, Lee D, Green J, Busby, S.** Secrets of bacterial transcription initiation taught by the *Escherichia coli* FNR protein. In: *Signals, Switches, Regulons & Cascades: Control of Bacterial Gene Expression. SGM Symposium* 2002; 61:127-142.
21. **Barnard A, Wolfe A, Busby, S.** Regulation at complex bacterial promoters: how bacteria use different promoter organisations to produce different regulatory outcomes. *Curr Op Microbiol* 2004*;* 7**:**102-108.
22. **Wu H-C, Tyson K, Cole J, Busby, S.** Regulation of the *E. coli nir* operon by two transcription factors: a new mechanism to account for co-dependence on two activators. *Mol Microbiol* 1998: 27:493-505.
23. **Browning D, Cole J, Busby, S.** Suppression of FNR-dependent transcription activation at the *Escherichia coli nir* promoter by Fis, IHF and H-NS: modulation of transcription initiation by a complex nucleo-protein assembly. *Mol Microbiol* 2000; 37:1258-1269.
24. **Browning D, Grainger D, Busby S.** Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Op Microbiol* 2010; 13:773-780.
25. **Dillon SC, Dorman CJ.** Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* 2010; 8:185-195.
26. **Browning DF, Cole, JA, Busby, SJ.** Transcription activation by remodeling of a nucleoprotein assembly: the role of NarL at the FNR-dependent *Escherichia coli nir* promoter. *Mol Microbiol* 2004; 53:203-215.
27. **Browning DF, Cole JA, Busby, SJ.** Regulation by nucleoid-associated proteins at the *Escherichia coli nir* operon promoter. *J Bacteriol* 2008; 190:7258-7267.
28. **Browning D, Busby S.** Local and global regulation of transcription initiation in bacteria. *Nature Rev Microbiol* 2016; 14:638-650.
29. **Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD**. The bacterial enhancer-dependent sigma(54) transcription factor. *J Bacteriol* 2000; 182:4129-4136.
30. **Yang Y, Darbari VC, Zhang N, Lu D, Glyde R *et al*.** Structures of the RNA polymerase-σ54 reveal new and conserved regulatory strategies. *Science* 2015; 349:882–885.
31. **Campbell EA, Kamath S, Rajashankar KR, Wu M, Darst SA.** Crystal structure of *Aquifex aeolicus* σ(N) bound to promoter DNA and the structure of σ(N)-holoenzyme. *Proc Natl Acad Sci U S A* 2017; 114:1805-1814.
32. **Glyde R, Ye F, Darbari VC, Zhang N, Buck M *et al*.** Structures of RNA polymerase closed and intermediate complexes reveal mechanisms of DNA opening and transcription initiation. *Mol Cell* 2017; 67:106-116.

33. **Glyde R, Ye F, Jovanovic M, Kotta-Loizou I, Buck M *et al*.** Structures of bacterial RNA polymerase complexes reveal the mechanism of DNA loading and transcription initiation. *Mol Cell* 2018. 70:1111-1120.
34. **Visweswariah SS, Busby SJ.** Evolution of bacterial transcription factors: how proteins take on new tasks, but do not always stop doing the old ones. *Trends in Microbiology* 2015; doi: 10.1016/j.tim.2015.04.009.
35. **Grainger D, Hurd D, Harrison M, Holdstock J, Busby, S**. Studies of the distribution of *Escherichia coli* cAMP receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci USA* 2005; 102:17693-17698.
36. **Shimada T, Ishihama A, Busby S, Grainger D.** The *Escherichia coli* RutR transcription factor binds at targets within genes as well as intergenic regions. *Nucl Acids Res* 2008; 36: 3950-3955.
37. **Fouqueau T, Blombach F, Werner F.** Evolutionary Origins of two-barrel RNA Polymerases and site-specific transcription initiation. *Annu Rev Microbiol* 2017; 71:331-348.

Figure Legends

Figure 1: Bacterial promoter elements and their recognition during transcript initiation

a: Sketch showing the organization of different promoter elements at a model bacterial promoter with respect to the transcript start that is denoted by '+1'. Elements are denoted by coloured rectangles with their consensus sequence motifs indicated. 'Dis' and 'CRE' denote the discriminator and core-recognition elements respectively. Note that bacterial promoters usually contain a selection of these elements but rarely contain all of them [28].

b: Sketch illustrating the key interactions between different domains of the RNAP sigma and alpha subunits and different promoter elements in the initial "closed" complex at a promoter. Note that the DNA duplex remains double-stranded and that the contacts result in Sigma Domain 2 being positioned adjacent to the promoter -10 element. Sigma Domain 1 prevents DNA template access to the RNAP active site.

c: Sketch illustrating the key interactions between different RNAP determinants and promoter elements in the transcriptionally-competent "open" complex at a promoter. The template strand is shown in yellow, and the non-template strand is in blue. Note that the DNA duplex around position +1 is unwound, with Sigma Domain 2 making specific base contacts with the single-stranded non-template strand of the promoter -10 element, thereby permitting the single-stranded template strand to enter the RNAP active site, which has become accessible due to movement of Sigma Domain 1.

Figure 2: Different architectures at bacterial activator-dependent promoters

Each panel illustrates the architecture of an Activator-RNAP-promoter DNA complex responsible for transcript initiation. RNAP is drawn as an oval with the 4 domains of the sigma subunit indicated by 4 segments of a smaller oval (shaded light blue). The two RNAP alpha subunit N-terminal and C-terminal domains, each joined by a linker, shown as a black line, are shown separately (αNTD, pale blue circles, and αCTD, pale blue ovals with yellow centres).

a: in Class I activation, the activator (shown by yellow circles to indicate a dimer) binds upstream and interacts (red dot) with αCTD to recruit RNAP to the promoter.

b: in Class II activation, the activator (shown by blue circles to indicate a dimer) binds adjacent to the promoter -35 element and interacts with determinants in RNAP (lilac dot) to recruit RNAP to the promoter.

c: a Class I activator (yellow) and a Class II activator can work together to recruit RNAP to the promoter.

d: two Class I activators (yellow and pink) can work together to recruit RNAP to the promoter. Each make Class I interactions (red dots).

Figure 3: NarL-dependent activation by anti-repression

The sketch illustrates the complex interactions at the *E. coli nir* operon promoter that result in co-dependence on FNR (dimer shown as dark blue circles, binding to its DNA target shown as dark blue horizontal arrows) and NarL (dimer shown as yellow circles, binding to its DNA target shown as yellow arrows). The *nir* operon transcript start is denoted by +1 and upstream binding site locations are denoted as –n, where n is the centre of the site.

a: the *nir* promoter is dependent on activation by FNR that acts via a Class II mechanism (see Figure 2b). Activation by FNR is suppressed by upstream bound nucleoid associated proteins, Fis (bound to a target at position -142) and IHF (bound to the IHF-I target at position -88). IHF bound at a second upstream target (IHF-II at position -115) stimulates FNR-dependent activation.

b: NarL binding (in response to the presence of nitrite or nitrate ions) displaces IHF from the IHF-I site, thereby relieving the suppression of FNR-dependent activation of the *nir* promoter.

Figure 4: Contrasting mechanisms of activation at bacterial promoters dependent on sigma-70 and sigma-54

RNAP carrying sigma-70 or sigma-54 is illustrated as in Figure 2.

a: Activator-dependent recruitment of RNAP to a target promoter (as in Figure 2a) leads to open complex formation as in Figure 1c.

b: RNAP carrying sigma-54 forms a closed complex due to binding of promoter -12 and -24 elements by different sigma determinants. Isomerisation to the open complex requires the action of an Enhancer Binding Protein (EBP, red ovals, bound to its target Upstream Activating sequence, UAS, red rectangle). EBP-catalysed ATP hydrolysis drives the reorganisation of sigma-54 for open complex formation. At many EBP-dependent promoters, EBP action is facilitated by DNA bending induced by the binding of nucleoid asscociated protein, IHF (blue striped circles).

Figure 5: Proposed evolution of bacterial transcription factors from nucleoid associated proteins.

The figure (adapted from [34]) illustrates a hypothetical timeline of evolution:

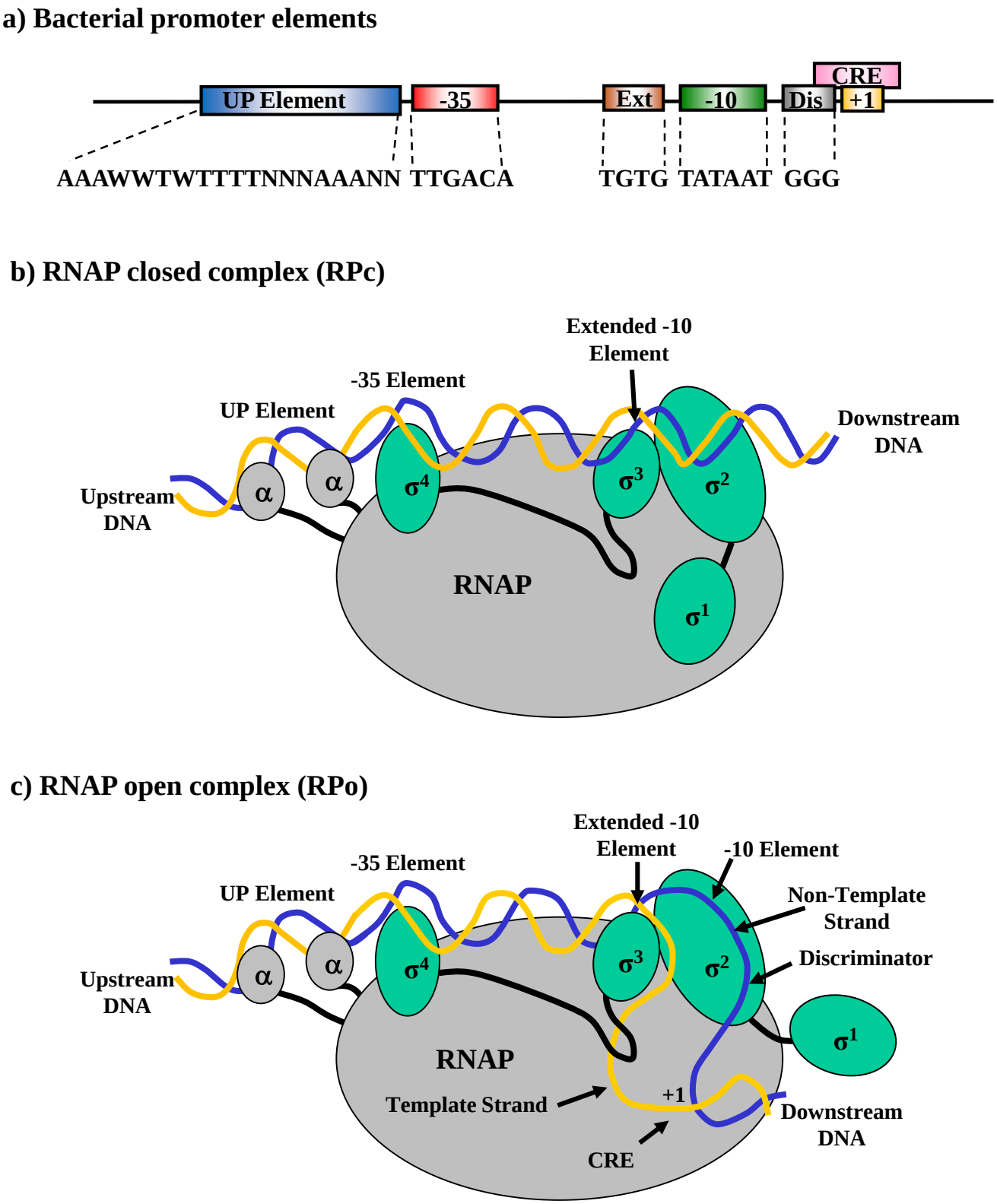
a: Nucleoid associated proteins (maroon and green ovals) organize the bacterial chromosome by binding ubiquitously and compete with RNAP.

b: Some nucleoid associated proteins (maroon ovals) acquire tight binding affinity for specific DNA targets (grey ovals illustrate dimers binding to grey boxes that depict operator sequences). This results in repression of certain promoters.

c: Some of the proteins bound at specific targets (pink dimers) acquire signal-responsive sensor domains (also shown in pink) that permits certain promoters to be derepressed by a signal.

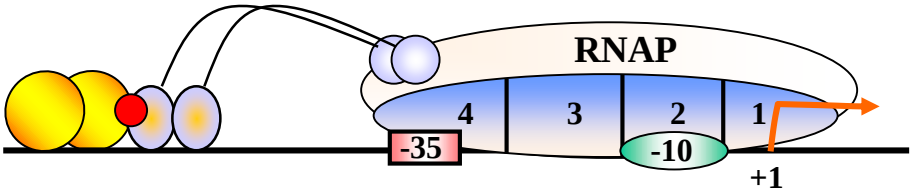
d: Both ligand-sensitive and ligand-insensitive factors acquire surface exposed determinants (red dots) that can recruit RNA polymerase to promoters.

MS Review, Figure 1.

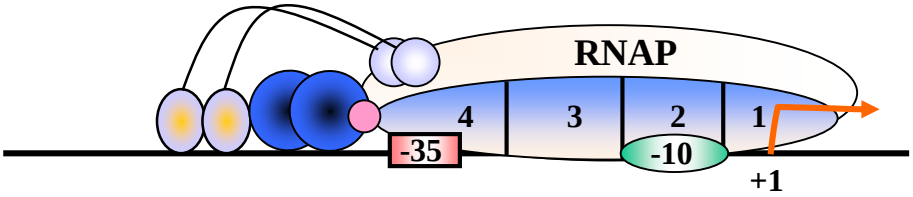


MS Review, Figure 2.

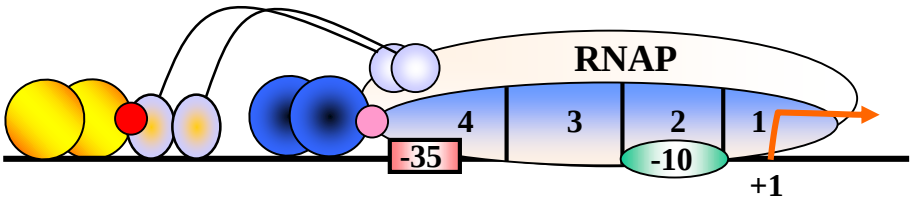
a) Class I activation.



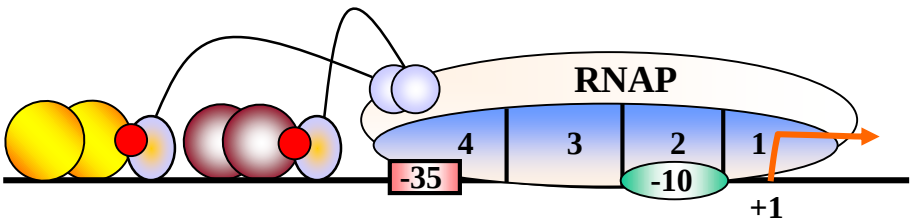
b) Class II activation.



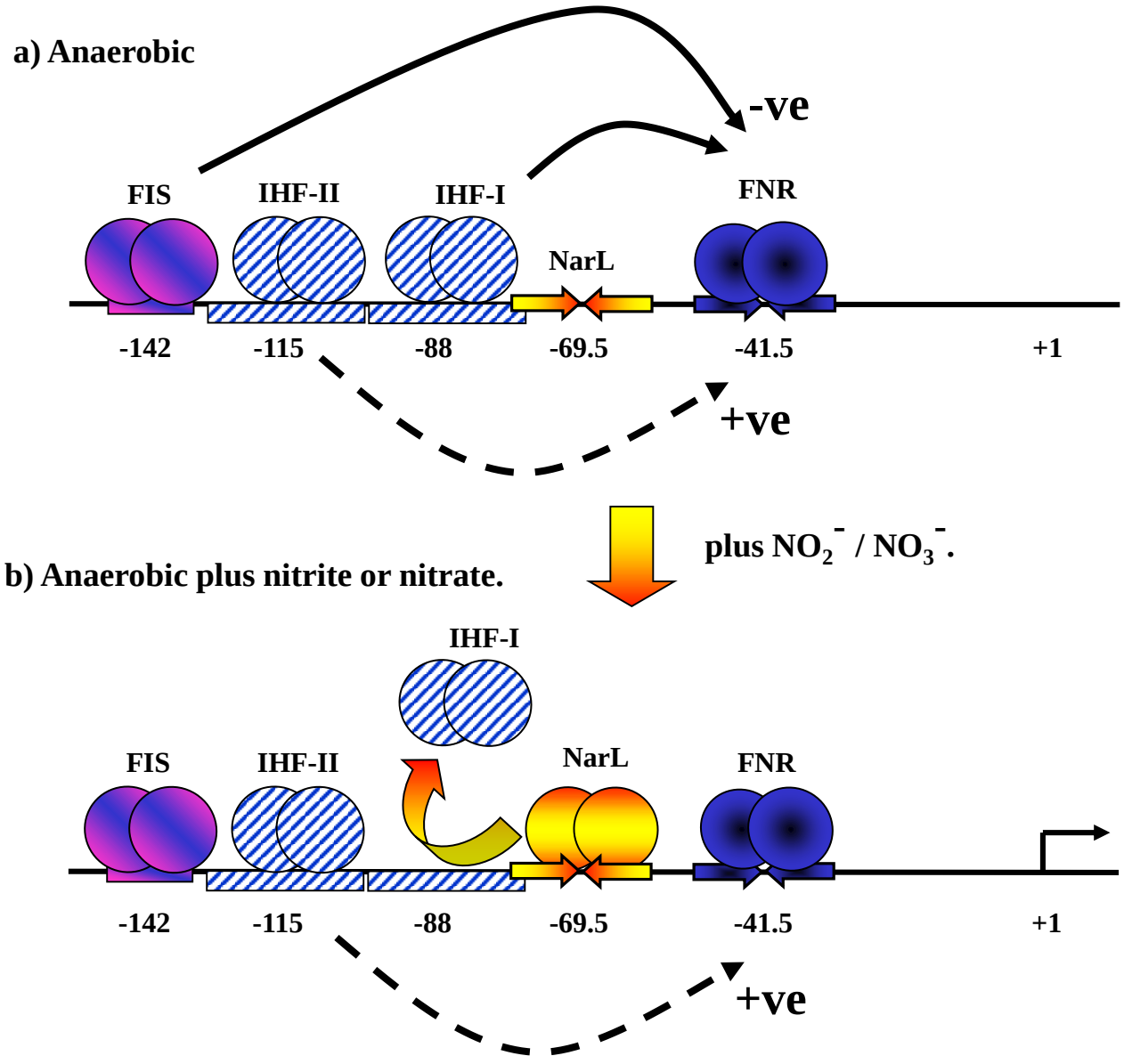
c) Class I/ Class II activation.



d) Class I/ Class I activation.

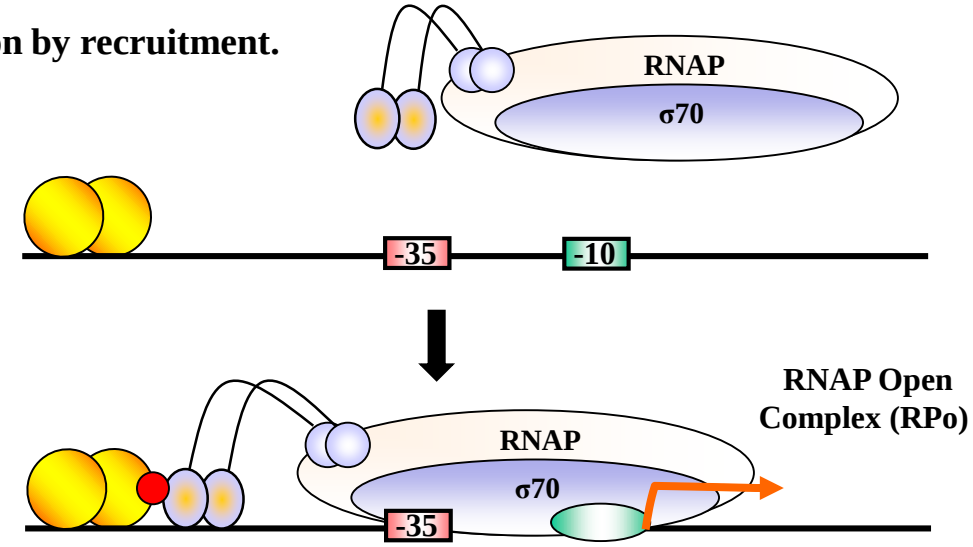


MS Review, Figure 3.

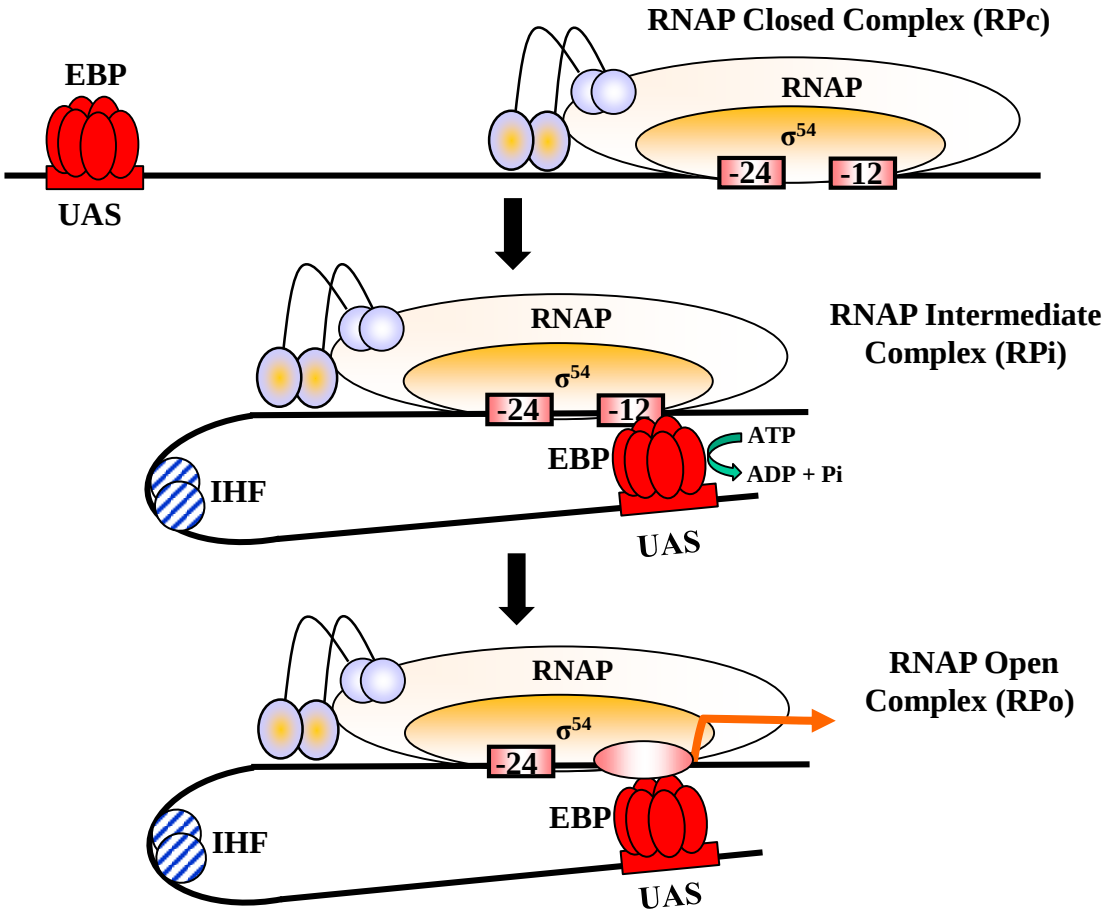


MS Review, Figure 4.

a) Activation by recruitment.



b) Sigma-54 dependent transcription.



MS Review, Figure 5.