Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

# Alternative transcriptional regulation in genome-reduced bacteria

Samuel Miravet-Verde[1,2,4], Verónica Lloréns-Rico[1,2,4] and Luis Serrano[1,2,3]

Transcription is a core process of bacterial physiology, and as such it must be tightly controlled, so that bacterial cells maintain steady levels of each RNA molecule in homeostasis and modify them in response to perturbations. The major regulators of transcription in bacteria (and in eukaryotes) are transcription factors. However, in genome-reduced bacteria, the limited number of these proteins is insufficient to explain the variety of responses shown upon changes in their environment. Thus, alternative regulators may play a central role in orchestrating RNA levels in these microorganisms. These alternative mechanisms rely on intrinsic features within DNA and RNA molecules, suggesting they are ancestral mechanisms shared among bacteria that could have an increased relevance on transcriptional regulation in minimal cells. In this review, we summarize the alternative elements that can regulate transcript abundance in genome-reduced bacteria and how they contribute to the RNA homeostasis at different levels.

**Addresses**
[1] EMBL/CRG Systems Biology Research Unit, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, 08003 Barcelona, Spain
[2] Universitat Pompeu Fabra (UPF), Dr. Aiguader 88, 08003 Barcelona, Spain
[3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluis Companys 23, 08010 Barcelona, Spain

Corresponding author: Serrano, Luis (luis.serrano@crg.eu)
[4] These authors contributed equally to this work.

**Current Opinion in Microbiology** 2017, **39**:89–95

This review comes from a themed issue on **Bacterial systems biology**

Edited by **Christoph Dehio** and **Dirk Bumann**

https://doi.org/10.1016/j.mib.2017.10.022



## Introduction

Genome-reduced bacteria are of remarkable interest as model organisms to study basic aspects of bacterial physiology. Because of their inherent simplicity, they are attractive for systems biology studies, whose results can be generalized to larger, more complex bacteria. These organisms have encountered defined niches to colonize as endosymbionts or pathogens, and have adapted to their environments by eliminating genes that are not required for their development. For instance, they have usually lost metabolic pathways to synthesize elements present in their natural environment [1]. Also, this niche adaptation has affected how gene expression is regulated in these organisms. Transcription factors (TFs), which have been traditionally considered the major drivers of transcriptional regulation, are scarce in bacteria with small genomes. In bacterial models like *Escherichia coli* or *Bacillus subtilis*, TFs represent 5–6% their total number of genes. This number is reduced by half (2.5% on average) in the Mollicutes class, a bacterial group including multiple minimal bacteria, most of them Mycoplasmas [2•]. A comparative analysis of 50 Mollicutes genomes identified 1–5 global regulators and up to 15 TFs in the Mycoplasmas with larger genome sizes [2•]. However, to the best of our knowledge, none of the putative global regulators has been characterized with the exception of the housekeeping sigma factor. Known transcription factors, including an additional sigma factor [3], only regulate a handful of genes [4].

Despite the tiny repertoire of TFs, these bacteria have not lost the ability to respond to a variety of external perturbations [4]. Therefore, it is possible that novel TFs remain undiscovered given the percentage of genes with unknown functions in these organisms, or that non-TF proteins with moonlighting functions act as TFs. Alternatively, different forms of regulating gene expression must exist, and may prevail, in these organisms. These alternative regulatory elements are probably not unique to genome-reduced bacteria, but they become more important as the process of genome reduction removes TFs to minimize the DNA content in these organisms. These alternative mechanisms of gene regulation are probably ancestral, as they are based in the chromosome structure and/or the intrinsic DNA or RNA sequences and not in proteins. The regulation they confer could have a smaller dynamical range and is more subtle than that by transcription factors, which makes it hard to observe in more complex bacteria. In this review, we focus on these other regulatory elements, from genome-wide to transcript-specific.

## Genome structure and DNA topology

First high-resolution 3D structure of a bacterial chromosome, obtained for *Caulobacter crescentus*, showed 23 interacting regions ranging from 30 to 400 kb bounded by

highly transcribed genes, known as chromosomal interaction domains (CIDs) [5]. Lately, ~20 CIDs were defined in *Bacillus subtilis* with a size between 50 and 300 kb [6]. Disposition of these elements is regulated by DNA supercoiling, which is controlled by topoisomerases [7] and nucleoid-associated proteins (NAPs) [8] (Figure 1a). *B. subtilis* presents four DNA topoisomerases: two ATP-independent (I and III) and two ATP-dependent (II, known as DNA gyrase, and IV) [9]. Minimal cells commonly present no topoisomerase III and a significative reduction of NAPs [10,11••]. With such a low number of DNA-binding proteins it was questionable whether small bacteria would preserve a chromosomal organization. A recent study in *Mycoplasma pneumoniae* found that small bacteria have enough components to maintain a defined chromosome structure and the presence of CIDs. In addition, this study provide the first evidence that genes inside CIDs tend to be co-regulated but the underlying mechanism to achieve this remains unknown. Interestingly, CIDs in *M. pneumoniae* are smaller (15–33 kb) but more frequent (44 CIDs) than *C. crescentus* and *B. subtilis* [11••]. Additionally, promoters are sensitive to local superhelical state as it regulates the distance between the elements participating in the promoter [12]; even in small-genome bacteria with reduced number of topoisomerases ([13], Yus et al., in preparation). Finally, ATP controls the ratio of ATP dependent/independent topoisomerases with direct effect on supercoiling and could imply a regulatory link between metabolism and genome topology and, consequently, expression [13].

## Genome organization in operons

Genome organization in operons constitutes a first level of gene regulation in prokaryotes. As transcription and translation occur simultaneously in bacteria, positional effects exist, and expression levels of the individual proteins in an operon are inversely proportional to the distance to the transcription initiation site of the operon [14]. This represents a level of regulation that is used not only in small but in all bacteria.

Traditionally, operons have been treated as static entities. However, recent research has shown that these structures are highly dynamic, being able to adapt in response to changing conditions, mainly thanks to termination, generating large transcripts or super-operons in some conditions, while producing short transcripts of sub-operons in others (Figure 1b) [15•]. In *M. pneumoniae*, this condition-dependent transcriptional read-through can explain a large part of how transcription is regulated [15•]. This mechanism has been shown to occur also in larger bacteria such as *E. coli* and *B. subtilis* [16].

## Bacterial promoters and transcription initiation

Promoter regions require certain features that make them recognizable by the RNA polymerase (RNAP) and the different TFs. Besides specific motifs binding sites for TFs, the most important sequence features are the boxes recognized by the RNAP complex and the different sigma factors. The housekeeping sigma factor binds two regions: the −10 box or Pribnow motif, and the −35 box. In genome-reduced bacteria, promoters have evolved towards the elimination of the −35 box, as this is non-existent or highly degenerated (Figure 1b) [8,17••,18]. In *Buchnera aphidicola*, an aphid symbiont with a minimal genome, regions similar to the −10 box of *E. coli* have been found, while a −35 motif has been only found upstream the rRNA genes [19]. In Gram-positive bacteria like *B. subtilis*, absence of a −35 element has been shown to be compensated if the Pribnow motif is preceded by a 'TG' dinucleotide (the so-called extended −10 box), but this short motif is present in only a handful of promoters in *Mycoplasma gallisepticum* [17••] and is not essential in determining promoters in *M. pneumoniae* [18]. This reduction in promoter complexity could be due to the scarcity of alternative sigma factors. This raises a question as to what makes promoters determine initiation of transcription and recognition by the RNAP complex. A recent study in *M. pneumoniae* points to the importance of the bases immediately surrounding the Pribnow motif, which tend to be A/T rich [20].

The structure of these regions is also important to trigger transcription. The double-stranded DNA should be less stable at the promoter region to unwind and accommodate the RNAP complex. Although the unwinding of the double helix is energetically favored at the promoters, the open complex formed between the promoter and the RNAP can be unstable. Unstable complexes require high concentrations of the initiating NTP (iNTP) to be stabilized so that RNA synthesis can be launched immediately. Otherwise, these complexes rapidly dissociate and transcription initiation is not produced. In contrast, very stable complexes require lower concentrations of the iNTP, as they will not easily dissociate [21]. Later, it was shown that the +2 nucleotide also modulates transcription initiation [22]. This mechanism establishes a link between cellular metabolism and transcriptional regulation and is not unique of genome reduced bacteria, but in the absence of major regulators this might be an elegant way to coordinate the expression of large groups of transcripts with identical +1 and +2 bases. An example of this nucleotide-based regulation includes the response to amino acid starvation (stringent response) in *B. subtilis*. In this scenario, concentration of ATP increases while GTP decreases as a consequence of the synthesis of (p)ppGpp (Figure 1b) [23]. Upregulated genes in this condition have adenosine in the +1 position, while downregulated promoters have guanosine. This effect could also be present and play a major role in the absence of many TFs in minimal bacteria as a regulatory mechanism dependent only in sequence composition.