Regulation of transcription initiation in mycobacteria

Priyanka Tare[1] and Valakunja Nagaraja[1,2,*]

[1]Department of Microbiology and Cell Biology, Indian Institute of Science, Bangalore 560 012, India
[2]Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560 064, India

The success of *Mycobacterium tuberculosis* as a deadly pathogen lies in its ability to survive under adverse conditions during pre- and post-infectious stages. The transcription process and the regulation of gene expression are central to the survival of the pathogen through the harsh conditions. Multiple sigma factors, transcription regulators, diverse two-component systems contribute in tailoring the events to meet the challenges faced by the pathogen. Although the machinery is conserved, many aspects of transcription and its regulation seem to be different in mycobacteria when compared to the other well-studied organisms. Here, we discuss salient aspects of transcription and its regulation in the context of distinct physiology of mycobacteria.

Keywords: Mycobacteria, pathogen, transcription regulators.

Introduction

FROM the time *Mycobacterium tuberculosis* was identified as the causative agent for tuberculosis, understanding the organism and controlling the devastating disease has been one of the major challenges. The organism continues to be the leading pathogen posing a serious threat, defying all measures to counter the infection. With millions getting afflicted each year, countering the organism and the treatment of the disease poses a great challenge. The pathogen continues to torment mankind, retaining the status of number one global killer. Its slow growth rate, formidable cell wall, latency and development of resistance to multiple drugs are only some of the hurdles presented by the organism, discussed and deliberated in scientific panels and other forums on basic biology as well as in clinical setting[1–6]. The exceptional success of the microbe lies in its inimitable lifestyle and the design of nearly perfect mechanism of pathogenesis. *M. tuberculosis* has evidently evolved elaborate strategies to overcome the adverse conditions faced during infection periods by efficiently modulating its regulatory network, gene expression, transport and signalling[5,6]. The nexus between *M. tuberculosis* and human immunodeficiency virus and the emergence of multiple drug-resistant strains has further aggravated the problem and demands massive counter measures to tackle it at various levels[7].

Although the study of the organism was a daunting challenge, efforts have been made over the last several decades to combat the disease and to understand the biology of the organism. These are summarized in the cited reviews and book chapters[1–4]. This early work was followed by several studies on the intermediary metabolism of the organism and molecular processes. Advent of newer approaches in the last few decades has led to major efforts to understand the organism better and also newer strategies to develop anti-tubercular molecules[8–11]. In this review, we summarize the more recent progress made to understand transcription in mycobacteria, one of the central processes in all the organisms with emphasis on the events during the initiation step.

RNA polymerase, the molecular machine of transcription

Transcription in bacteria is initiated after binding of the RNA polymerase (RNAP) to promoter DNA and melting of 12–14 base-pairs around the transcription start site to form a single-stranded 'transcription bubble' within a catalytically active RNAP–DNA open complex (RP_o). Initiation involves a series of sequential steps illustrated

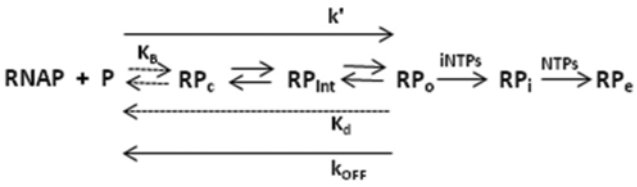

Figure 1. Schematic of transcription initiation pathway. When a promoter (P) is encountered by the RNA polymerase (RNAP) after the search, the enzyme forms a reversible complex with the DNA (closed complex, RP_c). A series of steps involving the separation of strands at –10 region results in a number of different intermediates collectively represented as RP_{int}. The conformational changes in DNA and RNAP eventually lead to the formation of open complex (RP_o). Binding of initiating ribonucleotides (iNTPs) to the complex leads to the formation of initiation complex (RP_i) having nascent RNA dinucleotide. The dinucleotide is extended into small RNA products (abortive transcription) and finally into the growing chain of productive RNA message by elongating RNAP (RP_e). Solid arrows, kinetic events; dotted arrows, equilibrium events; K_B, equilibrium binding constant; K_d, equilibrium dissociation constant; k′, association constant; k_{off}, dissociation constant.

*For correspondence. (e-mail: vraj@mcbl.iisc.ernet.in)

in Figure 1. All bacterial RNAPs have the subunit composition σ, β, β', 2α, ω and the mycobacterial enzyme is no exception[12,13]. Being a principle component of the essential process of transcription, one would expect the subunits of the core polymerase to be well conserved and that is indeed the case[14]. The major difference emerges in the sigma subunit composition as mycobacteria possess an unusually large number of sigma factors including the primary sigma factors, required for the transcription of a majority of the genes necessary for the housekeeping functions and alternative sigma factors, which direct the transcription initiation of genes involved in specialized functions[15–18]. Out of the 13 sigma factors in *M. tuberculosis*, SigA, B and F are representatives of the primary group, i.e. they belong to Sig 70 family[16], whereas others are grouped as alternative sigma factors. A number of them in the latter group have the characteristic features of extra-cytoplasmic factors (ECF)[15–18]. Table 1 summarizes the available information regarding the recognition sequence of each of the sigma factors, their function and response to different conditions including environmental and stress. Unlike in *Escherichia coli*, a majority of these sigma encoding genes seem to be expressed in the exponential phase although at varied levels[15,16,19,20]. For example, the qRT-PCR studies revealed that SigC transcripts were at the highest level followed by SigA, B, D, E mRNAs whereas SigF, H and M mRNAs were found to be present in lower levels[16]. Among these, the levels of SigA transcripts remained almost constant throughout the growth phases and in different stresses except during starvation and low aeration[16]. Alteration in the levels of the ECF sigma transcripts in response to different stress indicates their involvement in regulation of stress-specific genes (Table 1). A number of these sigma factor transcripts are translated during the exponential phase itself. Western blotting using antibodies against the SigA, B, E, F, L and H indicated their presence albeit at varied levels[13]. Although SigC mRNA was found to be abundant in exponential growth phase[16], no information regarding the level of the protein is available. However, the presence of SigA and B proteins in nearly equal quantity emphasizes their role as major factors for transcription from a large number of promoters in exponential phase[13]. The expression pattern of these sigma factors appears to be vastly different from any other bacteria, particularly *E. coli*[20]. For example, RpoS, the stationary phase sigma factor in *E. coli* is expressed in higher levels only during the starvation and stress conditions. SigB, which is considered to be a homologue of RpoS of *E. coli*, is expressed abundantly in exponential phase and the level goes up further at the later stages of growth or under stress conditions[16,20,21] (Table 1). Why SigC mRNA is present in high quantity in the exponential phase given that it belongs to the group of alternate sigmas, is not known. Notably, *sigC* mutation decreased the expression of a large number of genes suggesting its important role. The availability of the mRNA for immediate translation when needed could be a regulatory strategy as seen in the case of SigH transcripts in *E. coli*[22]. From the above discussion, it is evident that a number of diverse promoters are transcribed by their specific sigma factors even in exponential phase. Expression and utilization of alternative sigma factors during all phases of growth reflect the ability of the organism to adapt to various environmental and stress conditions. However, a number of these sigma factors are subjected to regulation by anti-sigma or even anti-anti sigma and other factors (Figure 2). Thus it is conceivable that being a clever pathogen, *M. tuberculosis* has tailored its transcription machinery to adjust to the changing environment rapidly during pre- and post-infection stages.

However, the expression of most of the sigma factors, if not all, during the exponential phase itself poses a problem for *in vitro* studies with the enzyme, especially in understanding the RNAP structure or specific gene transcription and analysis of the interaction of a particular promoter with the polymerase. Purified RNAP preparations from mycobacteria are a mixture of several holoenzymes although proportion of sigma subunits is highly varied[13]. SigA containing holo-enzyme constitutes only about 30% of the population hampering the quantitative *in vitro* transcription studies[13]. Although cloning, heterologous expression in *E. coli*, purification and reconstitution of the enzyme is achieved by different groups, these preparations had lower specific activity and were found to be unsuitable for quantitative *in vitro* studies[13]. To overcome this serious problem, we over-expressed SigA in *Mycobacterium smegmatis* and isolated holoenzyme with near stoichiometric amounts of the factor[13]. This was also facilitated by a *M. smegmatis* strain in which *rpoC* gene encoding β' subunit of the enzyme was chromosomally His tagged for the facile purification of the enzyme[23]. Such a single sigma subunit enriched holoenzyme has enabled promoter-specific transcription studies described in the next section.

Promoter architecture in mycobacteria and promoter–polymerase interactions

The existence of a large number of sigma factors for transcription initiation also points at the diversity of the promoters in the mycobacteria. In contrast to *E. coli* where, G or A are predominantly the transcription start site, mycobacterial RNAP prefers G as the initiating nucleotide for synthesis of RNA[24,25]. A matrix for –35 and –10 elements recognized by these factors has been generated based on the sequence analysis of the promoters functional under different conditions[25]. Based on these studies, it appears that the SigA recognized promoters have –10 element resembling that of *E. coli* Sig70 consensus but a less conserved –35 sequence[25] (Table 2). The consensus sequence for the promoters of most of the sigma factors in mycobacteria has been elucidated (Table 2). Based on the architecture and sigma factors involved in

Table 1. Consensus sequence of sigma factors, levels in different conditions and their roles

Sigma	Consensus –35---spacer---10	Response to stress	Functions	Ref.
SigA	TTGACW---N_{18}---TATAMT	Phagocytosis, low aeration, starvation ↑	House-keeping, essential for virulence	25, 83, 84
SigB	NGTGG---$N_{14–18}$---NNGNNG	Heat shock, SDS, oxidative stress ↑ ↓	Adaption to carbon starvation and general stress response	85
SigC	SSSAAT---$N_{16–20}$---CGTSSS	Stationary phase, heat shock, cold shock, oxidative, surface stress ↓	Virulence, maintenance of toxicity	86, 87, 88
SigD	GTAACGct---AT rich stretch	Starvation ↑	Survival during nutrient deprivation. Stringent response	89
SigE	gGGAACYa–$N_{15–16}$–cGTT	Heatshock, SDS exposure, vancomycin stress ↑	Growth in macrophages, response to heat shock, oxidative stress	19, 90, 91
SigF	GGWWT–$N_{16–17}$–GGGTAY	Anaerobic, antibiotic, metronidazole, nutrient deprivation ↑	Cell wall synthesis, growth during starvation	92–94
SigG	GCGNGT–$N_{15–18}$–CGANCA	Macrophage infection ↑	Growth in macrophages and SOS response	92
SigH	gGGAAYA–$N_{16–17}$–cGTT	Heat shock, diamide stress, macrophage infection ↑	Histopathology, host immune response, heat shock response	95, 96
SigJ	GTCACA–N_{16}–CGTCCT	Stationary phase, growth in macrophage ↑	Oxidative stress response	97, 98
SigK	CCATCC–N_{15}–CCGAAT	*NA	*NA	99
SigL	TGAACC–$N_{16–18}$–CGTCR	*NA	Oxidative and detergent stress	100
SigM	GGAAC–$N_{16–18}$–CGTCR	Stationary phase and heat shock ↑	Expression of four *esat-6* genes	101

*NA – Not available, ↑, upregulation, ↓, down regulation. R-A/G; W-A/T; M-A/C; S-C/G; Y-C/T.

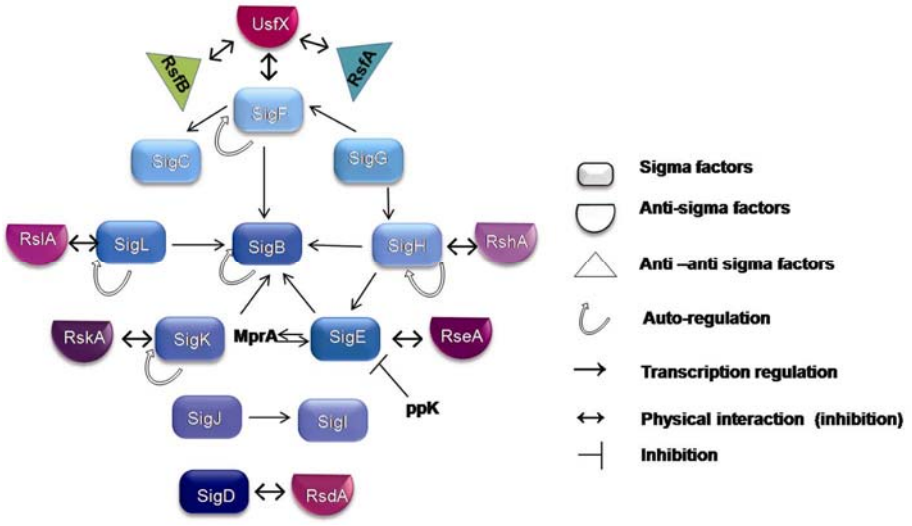

Figure 2. Sigma factors regulatory network in *M. tuberculosis*. The participating proteins are represented in different shapes and colors. For example, sigma factors are in different shades of blue. Out of thirteen sigma factors in the genome, information is not available for three. Double headed arrow indicates inhibition by physical sequestration; single headed arrow indicates regulation which could be either positive or negative. MprA is a response regulator of the two component system, MprAB; ppK, polyphosphate kinase.

recognition, the mycobacterial promoters have been categorized into four groups, viz. A, B, C and D. Those in group A have both, –35 and –10 elements similar to *E. coli* Sig70 consensus[25], whereas the promoters categorized in group B do not have a conserved –35 element[3]. Promoters in group A include principle promoter $P_{rrnPCL1}$

Table 2. Classification of mycobacterial promoters

Group	Organism	Consensus	Ref.
Group A	*M. tuberculosis*	$T_{62}T_{42}G_{76}A_{66}C_{71}A_{33}$---$T_{76}A_{81}K_{66}R_{66}A_{42}T_{81}$	25
	M. smegmatis	$T_{73}T_{58}G_{68}A_{26}C_{57}A_{36}$---$T_{94}A_{84}T_{63}A_{42}A_{42}T_{78}$	
Group B	*M. tuberculosis*	*NC---$T_{80}A_{80}G_{50}G_{50}C_{60}T_{99}$	3
	M. smegmatis	*NC---$T_{85}A_{85}NA_{57}C_{42}T_{71}$	
Group C	*M. tuberculosis*	*NC---NC	31, 32
	M. smegmatis		
Group D	*M. tuberculosis*	$T_{91}G_{64}C_{55}C_{55}G_{73}T_{27}$---$C_{82}R_{72}S_{90}C_{45}M_{63}S_{90}$	25
	M. smegmatis		

*Not conserved, R-A/G, W-A/T, M-A/C, S-C/G, Y-C/T.

and P_{rrnB} of *rrn* operons[26,27], initiator tRNA promoter, P_{metU} (ref. 28) and the principle promoter of *gyr* operon, P_{gyrB1} (ref. 25). Promoters transcribing *ideR*, *erp*, *purL* belong to group B promoters[3,29,30]. Group C is constituted by the promoters which do not resemble the *E. coli* consensus and control the transcription of regulated genes encoding HSP60 and GroEL1 (refs 3, 31, 32). Group D includes the promoters having unusual hexamer elements. One of them, termed SigGC promoter found upstream of *gyr* A was the first member to have these sequences[25]. The –35 and –10 elements of the promoter are highly GC rich and no cognate sigma factor has been identified so far. Subsequently, *par* genes of *Mycobacterium bovis* and *M. smegmatis* involved in partitioning of plasmids and chromosomes were shown to have these unusual GC rich promoter sequences[33].

Occurrence of a large number of sigma factors and diverse promoter architecture indicate a complex pattern of promoter–polymerase interaction. With the development of optimized RNAP preparation enriched with SigA, the *in vitro* analysis of specific promoter–polymerase interaction has become feasible and the kinetic analysis of transcription initiation has been carried out. The information from these studies has provided key insights into the individual steps of the process to unravel the regulation of the synthesis of stable RNA and enzymes having housekeeping functions.

In the total RNA pool of the cell, rRNA transcripts constitute the major fraction[34–36]. Rest of the bulk is contributed by the other stable RNA, tRNA and transcripts for house-keeping essential functions, which are also driven by efficient promoters[36]. A majority of bacteria possess multiple operons for rRNA with more than one promoter to transcribe each operon[34–38]. Some of the eubacteria have one or more *rrn* operons located near to *oriC*, an arrangement to take advantage of the gene-dosage effect during replication[39,40]. Fast growing bacteria such as *E. coli* and *Bacillus subtillis* have seven and ten rRNA operons respectively, driven by two promoters in all the operons[36–39]. However, all the members of the genus mycobacteria do not have the luxury of possessing multiple *rrn* operons[26,27,41,42]. The number of *rrn* operons varies from one or two depending on the growth rate of the members of the genus[27,41,42]. The slow growers have only one while the fast growers have two operons, to cope up with their growth characteristics[27]. Most notably, in many species, transcription is driven by several promoters ranging from two to four[26,27,41,42], thus having flexibility to generate more *rrn* transcripts when needed, depending upon the cellular requirements. As if to suit its very slow growth rate, *M. tuberculosis* has single operon, that too located about 1,500 kb away from the *oriC*[40]. However, two promoters P1 and PCL1 drive the transcription from the operon. In contrast, in *M. smegmatis*, a fast grower, the entire requirement of rRNA transcripts is met by two rRNA operons[27,41,42]. Of the two operons, *rrnBf* is driven by a single promoter P_{rrnB}, whereas the second operon *rrnAs* has three promoters P1, P2 and PCL1 (ref. 27). Besides the difference in the distribution, *rrn* promoters from the two organisms also show distinct properties[43,44]. For example, $P_{rrnPCL1}$, the principle promoter of *rrnA*s operon in both *M. smegmatis* and *M. tuberculosis* differs in the kinetics of transcription initiation and promoter activity between the two species[43,44]. The rate limiting step at *M. tuberculosis* $P_{rrnPCl1}$ was found to be the stability of the open complex, whereas the promoter clearance was the slowest step in *M. smegmatis*[43,44]. Both promoters respond to the small molecule regulators, initiating NTPs (iNTPs) and guanosine pentaphosphate (pppGpp), albeit to a different extent. iNTPs upregulate while pppGpp downregulates the transcription from these promoters[43,44]. The action of these regulators at rRNA promoters alters their strength in different growth phases. A GC rich sequence between –10 and transcription start site termed as discriminator is an important element in determining the response of the promoters to these small molecule regulators in *E. coli*[45]. Previous studies with the *rrn* promoters of mycobacteria reported a lack of discriminator sequences[39,46]. However, we have considered the existence of an alternate discriminator sequence in promoters of mycobacteria that are subjected to regulation by iNTPs and pppGpp. Indeed our ongoing analysis suggests the presence of different sequence determinants not only in rRNA operons, but

also in other promoters regulated by growth phase dependent control of gene expression (Tare and Nagaraja, MS in preparation). In addition to these unique sequence determinants, rRNA operon may also be differentially controlled by regulators that act in a manner unique to the organism. In *E. coli*, the *cis*-acting AT-rich UP elements and a nucleoid associated protein, Fis are involved in increasing the transcriptional efficiency of *rrn* promoters, adding another layer of regulation[47,48]. However, neither an AT rich UP element upstream of the operon nor a Fis homologue has been identified in mycobacteria[40,49]. Involvement of additional elements and regulatory proteins, if any, in rRNA regulation of mycobacteria is yet to be explored.

P_{metU} is one of the strong promoters in *M. tuberculosis* and *M. smegmatis*, encoding the crucial initiator tRNA responsible for translation initiation[28]. Unlike its *E. coli* counterpart, it is not arranged as an operon and is present in a single copy in both the species, representatives of fast and slow growing mycobacteria respectively[28]. Although these two organisms have different number of rRNA operons to suit their growth needs, presence of a single functional initiator tRNA gene in both of them is rather intriguing[26–28,41,42]. However, *M. smegmatis* P_{metU} is stronger than its counterpart in *M. tuberculosis*[28], which would allow synthesis of higher amounts of transcripts to cope up with the needs of faster dividing cells. Another surprising feature of the initiator tRNA promoters is their insensitivity towards iNTPs and pppGpp[44] (Tare *et al.*, unpublished data). The inhibition of the single promoter of an essential component during stationary phase could be detrimental for cell survival which may explain in part why these promoters are unresponsive to iNTP/pppGpp mediated regulation.

Gyrase promoters illustrate yet another example where the transcription machinery seems to be adapted differently to meet the needs posed by the distinct lifestyle of mycobacteria. Gyrase, a type II topoisomerase is a heterodimer of GyrA and GyrB, required to maintain the supercoiling homeostasis of the genomes of all eubacteria[50,51]. The transcriptional unit for the two genes – *gyrB* and *gyrA* in mycobacteria is arranged as an operon, an organization distinct from *E. coli* and many other organisms where the two genes are present at separate loci, far apart from each other[25]. In *M. smegmatis*, the dicistron is directed by a single promoter, whereas the *M. tuberculosis gyr* operon has multiple promoters[25,52]. The principle promoter P_{gyrB1} is accompanied by an overlapping and divergent promoter P_{gyrR}. The –35 and –10 elements of the principle promoter are similar to the SigA recognized sequences, but the promoter elements at P_{gyrR} do not resemble the consensus[25]. The reverse promoter which was shown to be 13 times weaker than the principle promoter may have an important function in the regulation of transcription at *gyr* operon in the conditions that alter the topology of the cell[25] (see later section). Besides the differences in the organization, the promoters of *M. tuberculosis* and *M. smegmatis* show distinct rate limiting steps during transcription initiation and regulation[43,44]. While the kinetics of open complex formation at P_{gyr} of *M. smegmatis* is slower[43], *M. tuberculosis* P_{gyrB1} is rate limited at promoter clearance step[44], similar to *E. coli gyr* promoters[53], highlighting subtle differences exhibited by promoters from closely related species transcribing the same gene. In this regard, a hitherto unknown mode of gyrase regulation during transcription initiation of the operon appears to contribute to the maintenance of distinct lifestyle of *M. tuberculosis*. The activity of the *M. tuberculosis* P_{gyrB1} is maximum in the exponential phase and decreases in the stationary phase rendering the promoter sensitive to the growth phase dependent regulation and respond to the small molecule effectors of the regulation, iNTPs and pppGpp[44]. This sensitivity to iNTPs and pppGpp seen with *M. tuberculosis* gyrase promoter is not observed for *gyr* promoters from *M. smegmatis*, *E. coli* or any other organism studied so far[43,44,53,54].

Apart from the growth phase-dependent regulation, the synthesis of gyrase mRNA is subjected to another layer of regulation termed as relaxation stimulated transcription (RST). This regulatory strategy is an adaptation to allow the homeostatic maintenance of the levels of gyrase[53,54]. In this mechanism, the transcription from the gyrase promoters is stimulated in response to relaxed status of the genome[53]. Unlike growth phase dependent control which seems to be incorporated only to *M. tuberculosis gyr* promoters, RST seems to be occurring at *gyr* promoters of *E. coli*, *M. smegmatis* and *M. tuberculosis*, though the mechanism appears to be different. In *M. smegmatis* formation of Cruciform Hairpin Palindromic Structure (CHPS) with an 8-base pair stem and a 4-base loop in the 5′ translated region results in inhibition of transcription when the genome is negatively supercoiled. As the genome gets relaxed, the structure is resolved, relieving the inhibition. In addition, studies in the plasmid context have shown that DNA elements present at 600 bp downstream within the ORF are essential for the response of the promoter to RST in *M. smegmatis*[52]. The exact mechanism by which both, promoter distal and proximal elements play a role in RST is yet to be elucidated. In *M. tuberculosis*, RST operates at a slower rate and the elements shown to be crucial for the response in *M. smegmatis* are not involved[25,52]. Rather, the overlapping promoters appear to contribute to RST. The overlapping and divergently oriented promoter P_{gyrR} could potentially occlude the binding of RNAP at the principle promoter in *gyr* operon reducing the transcription. Earlier study with the promoter provided evidence that in spite of the absence of any ORF downstream to P_{gyrR}, the RNAP at divergent promoter directs the transcript synthesis. The negative supercoiling appears to increase the transcription from the weaker promoter with a concomitant decrease at the P_{gyrB1} (ref. 25).

Role of transcription factors

During transcription initiation and elongation, distinct sets of transcription factors associate with various subunits of RNAP to modulate the DNA binding and catalytic properties. Between the initiation and elongation phases of the cycle, many of the factors have to be exchanged to ensure the continuation and completion of the process. In most of the promoters, the control of transcription initiation is not confined to promoter–polymerase interactions, but rather subjected to the action of the regulators that act either positively or negatively to influence the process. From the genome analysis of *M. tuberculosis* and related bacteria, it is apparent that more than 100 proteins are involved in the regulation of transcription[40]. Although the number appears to be smaller compared to the factors characterized in *E. coli*, *B. subtillis* and other well-studied organisms, very little is known about their function and molecular mechanism of action. From the studies carried out by several groups, it is evident that a number of regulatory factors are being characterized. Some of these factors are found only in mycobacteria and related species (WhiB family, DevR, CarD), whereas others are conserved in other bacteria as well. In the following section, we deal with a few of these factors, whose role has been subjected to detailed investigation. Broadly, these factors could be fitted into two groups – those which bind RNAP and influence the process during initiation, elongation and termination and the second set that bind to DNA to exert their regulatory role either positively or negatively. We are confining the following brief discussion to a few of the regulators that bind specific sequences in the DNA to exert global gene expression control and another selected set of factors that bind to the polymerase to influence its function. The factors and conditions responsible for the induction or activation of the regulatory proteins are presented in Table 3.

Cyclic AMP receptor protein (CRP) is one of the most well studied global regulators of transcription[55,56]. The protein gets activated after binding of cAMP. CRP binds to 16 bp sequence and modulates the transcription of the target promoters in mycobateria[56]. The genes targeted by cAMP–CRP complex appear to be involved in the persistence and/or emergence from the dormant state in *M. tuberculosis* as opposed to the genes involved in carbon metabolism in *E. coli* and number of other bacteria[55,57,58]. For example, CRP regulates the transcription of *WhiB1* and *rpfA*, genes having important functions in the growth cycle of *M. tuberculosis*[57,59]. The transcript of *rpfA* encodes a protein that is involved in reviving dormant bacteria, while the WhiB1 protein has a role in relaying the nitric oxide signal – a stress generated during the growth of the pathogen in macrophages[60]. WhiB1 is encoded by a gene belonging to the *wbl* family having seven genes numbered 1 to 7 (ref. 59). WhiB1 is sensitive to NO and when bound by the small molecule represses the transcription from its own promoter[59]. Recently it has been shown to repress the transcription from GroEL2 promoter encoding the essential gene[61]. Based on the targets affected by WhiB1, it appears that the protein is essential during the survival in macrophage and probably in maintaining the dormant state of the bacteria. WhiB3 binds to the promoters of polyketide biosynthetic genes in redox-dependent manner and regulates synthesis of inflammatory polyketides[62]. Besides, the protein also interacts with SigA to affect the host survival[63]. Recent studies have shown the role of WhiB4 during oxidative stress and maintenance of the redox balance in the cells[64]. Although WhiB1, WhiB3 and WhiB4 are well studied, the roles of other WhiB proteins are still being elucidated[61–64]. All the seven proteins of the WhiB family are redox-sensitive owing to [Fe–S] clusters stably bound to them[65]. The gene products of the family play important role in different physiological aspects of *M. tuberculosis*. These transcription factors influence essential processes, viz. cell division, survival during nutrient deprivation, pathogenesis, drug resistance, sensing different stress and are unique to *Mycobacterium*, *Streptomyces* spp. and other related actinobacteria[65]. These roles of WhiB family of proteins shed light on their exclusive function in complex physiology of mycobacteria. The key requirement of iron during infectious stages of the organism is yet another aspect of the unique lifestyle of mycobacteria. Iron is used as the cofactor for enzymes participating in the redox reactions. In humans, iron is majorly found in haem bound form and therefore the availability of iron for the pathogen is limited[66]. To compensate for the limited availability of usable iron in human blood, the pathogen produces siderophores, the high affinity metal chelators[66]. To regulate the amount of iron, which in higher amounts can be detrimental to the cells, mycobacteria contain a number of iron-dependent regulators. IdeR is one of those metallo-regulatory proteins, which is a DNA-binding transcription factor[66]. The genes regulated by IdeR include *mbtA*, *mbtB*, *mbtI*, *rv3402c*, *bfd*, *bfrA*, *mmpL4* and *mmpS4* and the list is likely to be much larger[67]. The IdeR controlled genes encode proteins with diverse functions such as transport and lipid metabolism. However, the primary targets of IdeR seem to be the ones involved in iron metabolism[66]. The essentiality of IdeR and its role in iron metabolism and oxidative stress response indicates its role during the survival of *M. tuberculosis* in human system and modulating the transcription during the infection periods to aid in better survival of the organism.

Besides these regulators, *M. tuberculosis* is armed with 11 two-component systems – a possible adaptation for its survival in the complex environment thereby modulating the gene expression in response to the diverse cues. Of these, PhoP–PhoQ and DevR–DevS are the most studied systems due to their regulatory effect on a large number

Table 3. Environmental conditions and factors responsible for activation/induction of transcription factors

Regulators	Conditions and factors	Ref.
CRP	cAMP, macrophage infection	57
CarD	Oxidative stress, DNA damage, starvation, genotoxins, ciprofloxacin, H_2O_2	78
DevRS	NO, hypoxia, dormancy	102
IdeR	Excess iron	66
PhoP	Low Mg^{2+}, low pH, antimicrobial peptides	69
RbpA	Stationary phase, aerated starvation, *in vitro* hypoxia, mouse macrophages, heat shock, di-amide stress, dormant stages	82 103
WhiB1	Hypoxia, cAMP, mice infection	104
WhiB2*	cAMP	
WhiB3	Hypoxia, NO	
WhiB4	cAMP, mice infection	
WhiB5	cAMP	
WhiB6*	Hypoxia, NO, macrophage infection	
WhiB7	Hypoxia, NO, mice and macrophage infection	

*WhiB2 and WhiB6 have been shown to be down-regulated during hypoxia and macrophage infection respectively.

of genes[68–72]. *M. tuberculosis* PhoP–PhoQ is essential for virulence and mutation in PhoP, the histidine kinase sensor, leads to attenuation of the strain, possibly by preventing the secretion of enzymes important for virulence. Notably, among the sequence variations between *M. tuberculosis* H37Rv and H37Ra, a single base mutation in PhoP is also considered important for virulence. PhoP regulates about 114 genes and many of these are involved in lipid metabolism, secretion and several other functions[68,69]. Another well-studied two-component system, DevR–DevS, is essential for survival of *M. tuberculosis* in host macrophages especially during dormancy. The expression of DevS, the sensory kinase of the two component system is induced by NO and hypoxia. DevR – the response regulator, binds to DNA in cooperative manner after it gets phosphorylated to alter the expression of the target genes[70]. The sequence in DNA crucial for binding of the regulator protein is termed as Dev box and its presence is shown in the promoters of several genes and the list is expanding as more genes are predicted to be the targets of DevR[70,71]. More detailed information about the two-component systems and their crucial role has been described in several papers[70–72].

A number of factors bind to RNAP at different stages of transcription. Among them, Nus factors and Rho could be considered as general factors that assist the enzyme during elongation and or termination. Although some information is available on some of the Nus factors, the function of transcription terminator factor Rho in *M. tuberculosis* is yet to be elucidated. At least two kinds of factors ensure that transcripts formed are devoid of any errors. Precise transcription by RNAP requires efficient removal of occasionally incorporated non-cognate nucleotide residues. Incorporation of a wrong nucleotide results in backtracking of transcription elongation complex (TEC). Hydrolysis of the mis-incorporated nucleotide at the 3′ end of the RNA and restoring the transcription elongation on track is a must to avoid deleterious consequences. Gre factors, originally named as transcript cleavage factors, assist RNAP active centre in the removal of mis-incorporated nucleotide to reset the process[73]. These transcription factors bind RNAP, approach the active centre Mg^{2+} through secondary channel and fine tune the centre to accelerate the proof reading process[73]. *E. coli* genome encodes two Gre factors (Gre A and B), both dispensable for the survival of the organism[73,74]. In contrast, a single Gre factor is found in the genomes of various mycobacteria and it seems to be essential for survival. The transcription factor enhances the inefficient promoter clearance and rescues the RNAP from a stalled complex to resume the transcription elongation. Though the cleavage signature of the single Gre factor resembles that of GreA in *E. coli*, it did not interact with *E. coli* RNAP[75]. The C-terminal domain responsible for interacting with RNAP shows considerable variations, which may confer the species specificity to these factors. Another factor, Rv3788, which shows 21% similarity to *M. tuberculosis* Gre, seems to be yet another secondary channel-binding protein[75,76]. The homologues of this protein have been found only in slow-growing mycobacteria and absent in fast-growing members[76]. Similar to the Gre factors, Rv3788 has conserved residues at the tip of N-terminal coiled coil domain[75,76]. Despite having similar domainal organization like Gre factors, the protein inhibited the transcription rather than inducing nascent transcript cleavage. The inability of the Rv3788 to induce cleavage may be attributed to reduced length of the N-terminus. The modulator binds at the entry site of the secondary channel, possibly competing with the binding of NTPs resulting in the transcription inhibition[76]. The exact mechanism of action of the inhibitor and environmental cues to which the protein responds are being elucidated.

Another set of factors help in recruiting nucleotide excision repair machinery when the RNAP stumbles upon DNA damage in template DNA itself. Mfd, also termed as transcription repair coupling factor based on its function binds to stalled RNAP, displaces it from the site of damage to recruit the repair machinery. Our studies reveal that although *M. tuberculosis* Mfd is similar to *E. coli* factor in many respects, some of its unique features may have important role during transcription elongation on GC-rich templates of mycobacteria[77]. Unlike in *E. coli* and other well studied systems, mycobacterial Mfd is found to be hexameric both *in vitro* and in intracellular milieu, suggesting its additional role in mycobacteria[77].

CarD, a highly expressed protein in *M. tuberculosis* and *M. smegmatis* is a modulator of transcription at many promoters including *rrn* and ribosomal proteins. The expression of CarD is stimulated under the conditions of stress and starvation[78,79]. CarD is essential for viability of *M. tuberculosis* and for persistence during infection. The protein binds at β subunit of RNAP and functions in a way similar to DksA of *E. coli* by affecting the stability of RNAP at *rrn* promoters[78,79]. Depletion of the essential protein causes sensitivity to oxidative stress, starvation, DNA damage and accumulation of rRNA transcripts. The essentiality of the two secondary channel-binding proteins, Gre and CarD in *M. tuberculosis* hint at their moonlighting functions. It has been suggested that they may serve to resolve the conflict between advancing replisome and transcription assembly[78]. Absence of the protein would therefore adversely affect the replisomes causing DNA damage leading to the cell death. However, this idea remains to be tested and established.

A novel RNAP binding factor, RbpA has been identified in some species of mycobacteria and *Streptomyces* and it seems to be confined to Actinomycetes. Though it is dispensable for *Streptomyces*, its essentiality in *M. tuberculosis* has been demonstrated[80]. In studies with *S. coelicolor*, RbpA was shown to alter the sensitivity of RNAP to rifampicin at rRNA promoters[81]. In yet another analysis using *M. smegmatis* RNAP, it was shown to bind close to the antibiotic-binding site excluding the drug from its site[82]. However, with *M. tuberculosis* RNAP, RbpA does not prevent rifampicin binding and the interaction is mapped to a distant site. The factor binds to β subunit of RNAP at a new target for activation – Sandwitch Barrel Hybrid Mottif (SBHM), a location distinct and far from rifampicin-binding residues[80]. Binding of RbpA to RNAP increases the longevity of the enzyme at $P_{rrnPCL1}$. It has been suggested that the protein enhances the binding affinity of SigA to the core RNAP to increase the transcription efficiency at the *rrnA* promoter. The working hypothesis of the latter study is that the stimulation of housekeeping gene expression may be the primary criteria for rifampicin tolerance and for adaptation to various stress during the infection. Although the mechanism by which the protein alters the response of cells to rifampicin is not completely understood, the effect of RbpA on transcription machinery and its indispensible nature hints at its significant role.

Conclusions

The studies on the transcription and its regulation in mycobacteria have provided substantial new insights into the sequential events and the main components that participate in the process. To begin with, abundance of sigma factors indicates a larger diversity in promoter recognition. The over representation of ECF sigma factors suggests the ability of the organism to respond to a variety of environmental stimuli. Presence of a large number of holo-enzyme species (core + different sigma factors) renders the transcription analysis more challenging. Function of a large number of regulatory proteins that may act as global or gene-specific factors is yet to be elucidated. Understanding the process and the elucidation of the mechanism of transcriptional regulation are the key advances likely to occur in the next few years. These developments would in turn facilitate the identification or development of the new inhibitors of the essential process – an important step in future drug-discovery efforts.

1. Ramakrishnan, T., Murthy, P. S. and Gopinathan, K. P., Intermediary metabolism of mycobacteria. *Microbiol. Mol. Biol. Rev.*, 1972, **36**, 65–108.
2. Barksdale, L. and Kim, K., Mycobacterium. *Bacteriol. Rev.*, 1977, **41**, 217–372.
3. Gomez, M. and Smith, S., Determinants of mycobacterial gene expression. In *Molecular Genetics of Mycobacteria* (eds Hatfull, G. F. and Jacobs Jr, W. R.), ASM Press, Washington, DC, 2000.
4. Glickman, M. S. and Jacobs Jr, W. R., Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. *Cell*, 2001, **104**, 477–485.
5. Shinnick, T. M., King, C. H. and Quinn, F. D., Molecular biology, virulence, and pathogenicity of mycobacteria. *Am. J. Med. Sci.*, 1995, **309**, 92–98.
6. Smith, I., *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.*, 2003, **16**, 463–496.
7. Daley, C. L., Small, P. M., Schecter, G. F., Schoolnik, G. K., McAdam, R. A., Jacobs Jr, W. R. and Hopewell, P. C., An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *N. Engl. J. Med.*, 1992, **326**, 231–235.
8. Clark-Curtiss, J. E. and Haydel, S. E., Molecular genetics of *Mycobacterium tuberculosis* pathogenesis. *Annu. Rev. Microbiol.*, 2003, **57**, 517–549.
9. Zhang, Y., Yew, W. W. and Barer, M. R., Targeting persisters for tuberculosis control. *Antimicrob. Agents Chemother.*, 2012, **56**, 2223–2230.
10. Barry, C. E. *et al.*, The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.*, 2009, **7**, 845–855.
11. Connolly, L. E., Edelstein, P. H. and Ramakrishnan, L., Why is long-term therapy required to cure tuberculosis? *PLoS Med.*, 2007, **4**, e120.
12. Borukhov, S. and Nudler, E., RNA polymerase: the vehicle of transcription. *Trends Microbiol.*, 2008, **16**, 126–134.

13. China, A. and Nagaraja, V., Purification of RNA polymerase from mycobacteria for optimized promoter-polymerase interactions. *Protein Exp. Purif.*, 2010, **69**, 235–242.
14. Finn, R. D., Orlova, E. V., Gowen, B., Buck, M. and van Heel, M., *Escherichia coli* RNA polymerase core and holoenzyme structures. *EMBO J.*, 2000, **19**, 6833–6844.
15. Gomez, J. E., Chen, J. M. and Bishai, W. R., Sigma factors of *Mycobacterium tuberculosis*. *Tuber Lung Dis.*, 1997, **78**, 175–183.
16. Manganelli, R., Dubnau, E., Tyagi, S., Kramer, F. R. and Smith, I., Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*. *Mol. Microbiol.*, 1999, **31**, 715–724.
17. Helmann, J. D., The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.*, 2002, **46**, 47–110.
18. Borukhov, S. and Severinov, K., Role of the RNA polymerase sigma subunit in transcription initiation. *Res. Microbiol.*, 2002, **153**, 557–562.
19. Manganelli, R., Voskuil, M. I., Schoolnik, G. K. and Smith, I., The *Mycobacterium tuberculosis* ECF sigma factor sigma E: role in global gene expression and survival in macrophages. *Mol. Microbiol.*, 2001, **41**, 423–437.
20. Osterberg, S., del Peso-Santos, T. and Shingler, V., Regulation of alternative sigma factor use. *Annu. Rev. Microbiol.*, 2011, **65**, 37–55.
21. Mulder, N. J., Powles, R. E., Zappe, H. and Steyn, L. M., The *Mycobacterium tuberculosis* mysB gene product is a functional equivalent of the *Escherichia coli* sigma factor, KatF. *Gene*, 1999, **240**, 361–370.
22. Morita, M. T., Tanaka, Y., Kodama, T. S., Kyogoku, Y., Yanagi, H. and Yura, T., Translational induction of heat shock transcription factor sigma 32: evidence for a built-in RNA thermosensor. *Genes Dev.*, 1999, **13**, 655–665.
23. Mukherjee, R. and Chatterji, D., Stationary phase induced alterations in mycobacterial RNA polymerase assembly: a cue to its phenotypic resistance towards rifampicin. *Biochem. Biophys. Res. Commun.*, 2008, **369**, 899–904.
24. Hawley, D. K. and McClure, W. R., Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucl. Acids Res.*, 1983, **11**, 2237–2255.
25. Unniraman, S., Chatterji, M. and Nagaraja, V., DNA gyrase genes in *Mycobacterium tuberculosis*: a single operon driven by multiple promoters. *J. Bacteriol.*, 2002, **184**, 5449–5456.
26. Gonzalez-y-Merchand, J. A., Garcia, M. J., Gonzalez-Rico, S., Colston, M. J. and Cox, R. A., Strategies used by pathogenic and nonpathogenic mycobacteria to synthesize rRNA. *J. Bacteriol.*, 1997, **179**, 6949–6958.
27. Verma, A., Sampla, A. K. and Tyagi, J. S., *Mycobacterium tuberculosis rrn* promoters: differential usage and growth rate-dependent control. *J. Bacteriol.*, 1999, **181**, 4326–4333.
28. Dastur, A., Kumar, P., Ramesh, S., Vasanthakrishna, M. and Varshney, U., Analysis of the initiator tRNA genes from a slow- and a fast-growing Mycobacterium. *Arch. Microbiol.*, 2002, **178**, 288–296.
29. Jackson, M., Berthet, F. X., Otal, I., Rauzier, J., Martin, C., Gicquel, B. and Guilhot, C., The *Mycobacterium tuberculosis* purine biosynthetic pathway: isolation and characterization of the *purC* and *purL* genes. *Microbiology*, 1996, **142**, 2439–2447.
30. Berthet, F. X., Rauzier, J., Lim, E. M., Philipp, W., Gicquel, B. and Portnoi, D., Characterization of the *Mycobacterium tuberculosis erp* gene encoding a potential cell surface protein with repetitive structures. *Microbiology*, 1995, **141**, 2123–2130.
31. Kong, T. H., Coates, A. R., Butcher, P. D., Hickman, C. J. and Shinnick, T. M., *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 2608–2612.
32. Stover, C. K. *et al.*, New use of BCG for recombinant vaccines. *Nature*, 1991, **351**, 456–460.
33. Casart, Y., Gamero, E., Rivera-Gutierrez, S. and Gonzalez, Y. M. J. A. and Salazar, L., *par* genes in *Mycobacterium bovis* and *Mycobacterium smegmatis* are arranged in an operon transcribed from 'SigGC' promoters. *BMC Microbiol.*, 2008, **8**, 51.
34. Paul, B. J., Barker, M. M., Ross, W., Schneider, D. A., Webb, C., Foster, J. W. and Gourse, R. L., DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, 2004, **118**, 311–322.
35. Klumpp, S. and Hwa, T., Traffic patrol in the transcription of ribosomal RNA. *RNA Biol.*, 2009, **6**, 392–394.
36. Dennis, P. P., Ehrenberg, M., Fange, D. and Bremer, H., Varying rate of RNA chain elongation during *rrn* transcription in *Escherichia coli*. *J. Bacteriol.*, 2009, **191**, 3740–3746.
37. Krasny, L. and Gourse, R. L., An alternative strategy for bacterial ribosome synthesis: *Bacillus subtilis* rRNA transcription regulation. *EMBO J.*, 2004, **23**, 4473–4483.
38. Perron, K., Comte, R. and van Delden, C., DksA represses ribosomal gene transcription in *Pseudomonas aeruginosa* by interacting with RNA polymerase on ribosomal promoters. *Mol. Microbiol.*, 2005, **56**, 1087–1102.
39. Condon, C., Squires, C. and Squires, C. L., Control of rRNA transcription in *Escherichia coli*. *Microbiol. Rev.*, 1995, **59**, 623–645.
40. Cole, S. T., Comparative and functional genomics of the *Mycobacterium tuberculosis* complex. *Microbiology*, 2002, **148**, 2919–2928.
41. Gonzalez-y-Merchand, J. A., Colston, M. J. and Cox, R. A., The rRNA operons of *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*: comparison of promoter elements and of neighbouring upstream genes. *Microbiology*, 1996, **142**, 667–674.
42. Gonzalez-y-Merchand, J. A., Colston, M. J. and Cox, R. A., Roles of multiple promoters in transcription of ribosomal DNA: effects of growth conditions on precursor rRNA synthesis in mycobacteria. *J. Bacteriol.*, 1998, **180**, 5756–5761.
43. China, A., Tare, P. and Nagaraja, V., Comparison of promoter-specific events during transcription initiation in mycobacteria. *Microbiology*, 2010, **156**, 1942–1952.
44. Tare, P., China, A. and Nagaraja, V., Distinct and contrasting transcription initiation patterns at *Mycobacterium tuberculosis* promoters. *PLoS One*, 2012, **7**, e43900.
45. Travers, A. and Muskhelishvili, G., DNA supercoiling – a global transcriptional regulator for enterobacterial growth? *Nat. Rev. Microbiol.*, 2005, **3**, 157–169.
46. Josaitis, C. A., Gaal, T. and Gourse, R. L., Stringent control and growth-rate-dependent control have nonidentical promoter sequence requirements. *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 1117–1121.
47. Ross, W. *et al.*, A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, 1993, **262**, 1407–1413.
48. Ross, W., Thompson, J. F., Newlands, J. T. and Gourse, R. L., *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. *EMBO J.*, 1990, **9**, 3733–3742.
49. Arnvig, K. B., Gopal, B., Papavinasasundaram, K. G., Cox, R. A. and Colston, M. J., The mechanism of upstream activation in the *rrnB* operon of *Mycobacterium smegmatis* is different from the *Escherichia coli* paradigm. *Microbiology*, 2005, **151**, 467–473.
50. Gellert, M., O'Dea, M. H., Itoh, T. and Tomizawa, J., Novobiocin and coumermycin inhibit DNA supercoiling catalyzed by DNA gyrase. *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 4474–4478.
51. Gellert, M., Mizuuchi, K., O'Dea, M. H. and Nash, H. A., DNA gyrase: an enzyme that introduces superhelical turns into DNA. *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 3872–3876.
52. Unniraman, S. and Nagaraja, V., Regulation of DNA gyrase operon in *Mycobacterium smegmatis*: a distinct mechanism of relaxation stimulated transcription. *Genes Cells*, 1999, **4**, 697–706.

53. Menzel, R. and Gellert, M., Modulation of transcription by DNA supercoiling: a deletion analysis of the *Escherichia coli gyrA* and *gyrB* promoters. *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 4185–4189.
54. Menzel, R. and Gellert, M., Regulation of the genes for *E. coli* DNA gyrase: homeostatic control of DNA supercoiling. *Cell*, 1983, **34**, 105–113.
55. Shimada, T., Fujita, N., Yamamoto, K. and Ishihama, A., Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One*, 2011, **6**, e20081.
56. Nanchen, A., Schicker, A., Revelles, O. and Sauer, U., Cyclic AMP-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*. *J. Bacteriol.*, 2008, **190**, 2323–2330.
57. Stapleton, M., Haq, I., Hunt, D. M., Arnvig, K. B., Artymiuk, P. J., Buxton, R. S. and Green, J., *Mycobacterium tuberculosis* cAMP receptor protein (Rv3676) differs from the *Escherichia coli* paradigm in its cAMP binding and DNA binding properties and transcription activation properties. *J. Biol. Chem.*, 2010, **285**, 7016–7027.
58. Bai, G., McCue, L. A. and McDonough, K. A., Characterization of *Mycobacterium tuberculosis* Rv3676 (CRPMt), a cyclic AMP receptor protein-like DNA binding protein. *J. Bacteriol.*, 2005, **187**, 7795–7804.
59. Agarwal, N., Raghunand, T. R. and Bishai, W. R., Regulation of the expression of whiB1 in *Mycobacterium tuberculosis*: role of cAMP receptor protein. *Microbiology*, 2006, **152**, 2749–2756.
60. Rickman, L. *et al.*, A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. *Mol. Microbiol.*, 2005, **56**, 1274–1286.
61. Stapleton, M. R., Smith, L. J., Hunt, D. M., Buxton, R. S. and Green, J., *Mycobacterium tuberculosis* WhiB1 represses transcription of the essential chaperonin GroEL2. *Tuberculosis*, 2012, **92**, 328–332.
62. Singh, A., Crossman, D. K., Mai, D., Guidry, L., Voskuil, M. I., Renfrow, M. B. and Steyn, A. J. C., *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog.*, 2009, **5**, e1000545.
63. Steyn, A. J., Collins, D. M., Hondalus, M. K., Jacobs Jr, W. R., Kawakami, R. P. and Bloom, B. R., *Mycobacterium tuberculosis* WhiB3 interacts with RpoV to affect host survival but is dispensable for *in vivo* growth. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 3147–3152.
64. Chawla, M. *et al.*, *Mycobacterium tuberculosis* WhiB4 regulates oxidative stress response to modulate survival and dissemination *in vivo*. *Mol. Microbiol.*, 2012, **85**, 1148–1165.
65. Alam, M. S., Garg, S. K. and Agrawal, P., Studies on structural and functional divergence among seven WhiB proteins of *Mycobacterium tuberculosis* H37Rv. *FEBS J.*, 2009, **276**, 76–93.
66. Rodriguez, G. M., Voskuil, M. I., Gold, B., Schoolnik, G. K. and Smith, I., ideR, An essential gene in *Mycobacterium tuberculosis*: role of IdeR in iron-dependent gene expression, iron metabolism and oxidative stress response. *Infect. Immun.*, 2002, **70**, 3371–3381.
67. Gold, B., Rodriguez, G. M., Marras, S. A., Pentecost, M. and Smith, I., The *Mycobacterium tuberculosis* IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages. *Mol. Microbiol.*, 2002, **42**, 851–865.
68. Goyal, R., Das, A. K., Singh, R., Singh, P. K., Korpole, S. and Sarkar, D., Phosphorylation of PhoP protein plays direct regulatory role in lipid biosynthesis of *Mycobacterium tuberculosis*. *J. Biol. Chem.*, 2011, **286**, 45197–45208.
69. Ryndak, M., Wang, S. and Smith, I., PhoP, a key player in *Mycobacterium tuberculosis* virulence. *Trends Microbiol.*, 2012, **16**, 528–534.
70. Chauhan, S. and Tyagi, J. S., Interaction of DevR with multiple binding sites synergistically activates divergent transcription of narK2-Rv1738 genes in *Mycobacterium tuberculosis*. *J. Bacteriol.*, 2008, **190**, 5394–5403.
71. Chauhan, S., Sharma, D., Singh, A., Surolia, A. and Tyagi, J. S., Comprehensive insights into *Mycobacterium tuberculosis* DevR (DosR) regulon activation switch. *Nucleic Acids Res.*, 2012, **39**, 7400–7414.
72. Leistikow, R. L., Morton, R. A., Bartek, I. L., Frimpong, I., Wagner, K. and Voskuil, M. I., The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J. Bacteriol.*, 2010, **192**, 1662–1670.
73. Borukhov, S., Lee, J. and Laptenko, O., Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.*, 2005, **55**, 1315–1324.
74. Orlova, M., Newlands, J., Das, A., Goldfarb, A. and Borukhov, S., Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 4596–4600.
75. China, A., Mishra, S. and Nagaraja, V., A transcript cleavage factor of *Mycobacterium tuberculosis* important for its survival. *PLoS One*, 2010, **6**, e21941.
76. China, A., Mishra, S., Tare, P. and Nagaraja, V., Inhibition of *Mycobacterium tuberculosis* RNA polymerase by binding of a Gre factor homolog to the secondary channel. *J. Bacteriol.*, 2012, **194**, 1009–1017.
77. Prabha, S., Rao, D. N. and Nagaraja, V., Distinct properties of hexameric but functionally conserved *Mycobacterium tuberculosis* transcription-repair coupling factor. *PLoS One*, 2011, **6**, e19131.
78. Stallings, C. L., Stephanou, N. C., Chu, L., Hochschild, A., Nickels, B. E. and Glickman, M. S., CarD is an essential regulator of rRNA transcription required for *Mycobacterium tuberculosis* persistence. *Cell*, 2009, **138**, 146–159.
79. Weiss, L. A., Harrison, P. G., Nickels, B. E., Glickman, M. S., Campbell, E. A., Darst, S. A. and Stallings, C. L., Interaction of CarD with RNA polymerase mediates *Mycobacterium tuberculosis* viability, rifampin resistance, and pathogenesis. *J. Bacteriol.*, 2012, **194**, 5621–5631.
80. Hu, Y., Morichaud, Z., Chen, S., Leonetti, J. P. and Brodolin, K., *Mycobacterium tuberculosis* RbpA protein is a new type of transcriptional activator that stabilizes the sigma A-containing RNA polymerase holoenzyme. *Nucl. Acids Res.*, 2012, **40**, 6547–6557.
81. Newell, K. V., Thomas, D. P., Brekasis, D. and Paget, M. S., The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on *Streptomyces coelicolor*. *Mol. Microbiol.*, 2006, **60**, 687–696.
82. Dey, A., Verma, A. K. and Chatterji, D., Role of an RNA polymerase interacting protein, MsRbpA, from *Mycobacterium smegmatis* in phenotypic tolerance to rifampicin. *Microbiology*, 2010, **156**, 873–883.
83. Collins, D. M. *et al.*, Generation of attenuated *Mycobacterium bovis* strains by signature-tagged mutagenesis for discovery of novel vaccine candidates. *Infect. Immun.*, 2005, **73**, 2379–2386.
84. Rodrigue, S., Provvedi, R., Jacques, P. E., Gaudreau, L. and Manganelli, R., The sigma factors of *Mycobacterium tuberculosis*. *FEMS Microbiol. Rev.*, 2006, **30**, 926–941.
85. Lee, J. H., Karakousis, P. C. and Bishai, W. R., Roles of SigB and SigF in the *Mycobacterium tuberculosis* sigma factor network. *J. Bacteriol.*, 2008, **190**, 699–707.
86. Thakur, K. G. and Gopal, B., Crystallization and preliminary X-ray diffraction studies of two domains of a bilobed extracytoplasmic function sigma factor SigC from *Mycobacterium*

tuberculosis. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 2005, **61**, 779–781.
87. Sun, R., Converse, P. J., Ko, C., Tyagi, S., Morrison, N. E. and Bishai, W. R., *Mycobacterium tuberculosis* ECF sigma factor sigC is required for lethality in mice and for the conditional expression of a defined gene set. *Mol. Microbiol.*, 2004, **52**, 25–38.
88. Chang, A., Smollett, K. L., Gopaul, K. K., Chan, B. H. and Davis, E. O., *Mycobacterium tuberculosis* H37Rv sigC is expressed from two promoters but is not auto-regulatory. *Tuberculosis*, 2012, **92**, 48–55.
89. Calamita, H., Ko, C., Tyagi, S., Yoshimatsu, T., Morrison, N. E. and Bishai, W. R., The *Mycobacterium tuberculosis* SigD sigma factor controls the expression of ribosome-associated gene products in stationary phase and is required for full virulence. *Cell Microbiol.*, 2005, **7**, 233–244.
90. Song, T., Song, S. E., Raman, S., Anaya, M. and Husson, R. N., Critical role of a single position in the –35 element for promoter recognition by *Mycobacterium tuberculosis* SigE and SigH. *J. Bacteriol.*, 2005, **190**, 2227–2230.
91. Fontan, P. A. *et al.*, *Mycobacterium tuberculosis* sigma factor E regulon modulates the host inflammatory response. *J. Infect. Dis.*, 2008, **198**, 877–885.
92. Lee, J. H., Geiman, D. E. and Bishai, W. R., Role of stress response sigma factor SigG in *Mycobacterium tuberculosis*. *J. Bacteriol.*, 2008, **190**, 1128–1133.
93. Rodrigue, S., Brodeur, J., Jacques, P. E., Gervais, A. L., Brzezinski, R. and Gaudreau, L., Identification of mycobacterial sigma factor binding sites by chromatin immunoprecipitation assays. *J. Bacteriol.*, 2007, **189**, 1505–1513.
94. Williams, E. P., Lee, J. H., Bishai, W. R., Colantuoni, C. and Karakousis, P. C., *Mycobacterium tuberculosis* SigF regulates genes encoding cell wall-associated proteins and directly regulates the transcriptional regulatory gene phoY1. *J. Bacteriol.*, 2007, **189**, 4234–4242.
95. Raman, S., Song, T., Puyang, X., Bardarov, S., Jacobs Jr, W. R. and Husson, R. N., The alternative sigma factor SigH regulates major components of oxidative and heat stress responses in *Mycobacterium tuberculosis*. *J. Bacteriol.*, 2001, **183**, 6119–6125.
96. Dutta, N. K. *et al.*, The stress-response factor SigH modulates the interaction between *Mycobacterium tuberculosis* and host phagocytes. *PLoS One*, 2012, **7**, e28958.
97. Hu, Y. and Coates, A. R., Increased levels of sigJ mRNA in late stationary phase cultures of *Mycobacterium tuberculosis* detected by DNA array hybridisation. *FEMS Microbiol. Lett.*, 2001, **202**, 59–65.
98. Hu, Y., Kendall, S., Stoker, N. G. and Coates, A. R., The *Mycobacterium tuberculosis* sigJ gene controls sensitivity of the bacterium to hydrogen peroxide. *FEMS Microbiol. Lett.*, 2004, **237**, 415–423.
99. Veyrier, F., Said-Salim, B. and Behr, M. A., Evolution of the mycobacterial SigK regulon. *J. Bacteriol.*, 2008, **190**, 1891–1899.
100. Hahn, M. Y., Raman, S., Anaya, M. and Husson, R. N., The *Mycobacterium tuberculosis* extracytoplasmic-function sigma factor SigL regulates polyketide synthases and secreted or membrane proteins and is required for virulence. *J. Bacteriol.*, 2005, **187**, 7062–7071.
101. Agarwal, N., Woolwine, S. C., Tyagi, S. and Bishai, W. R., Characterization of the sigma factor SigM by assessment of virulence and identification of SigM-dependent genes. *Infect Immun.*, 2007, **75**, 452–461.
102. Saini, D. K., Malhotra, V., Dey, D., Pant, N., Das, T. K. and Tyagi, J. S., DevR-DevS is a bonafide two component system of *Mycobacterium tuberculosis* that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. *Microbiology*, 2004, **150**, 865–875.
103. Murphy, D. J. and Brown, J. R., Identification of gene targets against dormant phase *Mycobacterium tuberculosis* infections. *BMC Infect. Dis.*, 2007, **7**, 84–89.
104. Larsson, C., Luna, B., Ammerman, N. C., Maiga, M., Agarwal, N. and Bishai, W. R., Gene expression of *Mycobacterium tuberculosis* putative transcription transcription factors whiB1-7 in redox environments. *PloS One*, 2012, **7**, e37516.

ACKNOWLEDGEMENTS. We thank the members of the laboratory for their contribution to the work and discussions. The work is supported by the Centre of Excellence in Tuberculosis Research grant from Department of Biotechnology, New Delhi. V.N. is a J.C. Bose Fellow of DST, New Delhi.