Transcription termination factor Rho: a hub linking diverse physiological processes in bacteria

Aleksandra Grylak-Mielnicka, Vladimir Bidnenko, Jacek Bardowski and Elena Bidnenko

Abstract

Factor-dependent termination of transcription in bacteria relies on the activity of a specific RNA helicase, the termination factor Rho. Rho is nearly ubiquitous in bacteria, but the extent to which its physiological functions are conserved throughout the different phyla remains unknown. Most of our current knowledge concerning the mechanism of Rho’s activity and its physiological roles comes from the model micro-organism Escherichia coli, where Rho is essential and involved in the control of several important biological processes. However, the rather comprehensive knowledge about the general mechanisms of action and activities of Rho based on the E. coli paradigm cannot be directly extrapolated to other bacteria. Recent studies performed in different species favour the view that Rho-dependent termination plays a significant role even in bacteria where Rho is not essential. Here, we summarize the current state of the ever-increasing knowledge about the various aspects of the physiological functions of Rho, such as limitation of deleterious foreign DNA expression, control of gene expression, suppression of pervasive transcription, prevention of R-loops and maintenance of chromosome integrity, focusing on similarities and differences of the activities of Rho in various bacterial species.

Introduction

Transcription termination is a critical step in gene regulation in all living organisms. In bacteria, transcription termination is well known to be essential for the generation of different types of functional RNAs, the definition of the boundaries of the transcriptional units, the release of RNA polymerase (RNAP) and the regulation of gene expression via the mechanism known as transcription attenuation (Peters et al., 2011; Santangelo & Artsimovitch, 2011).

However, recent studies have revealed new roles of transcriptional termination, e.g. those linked to the maintenance of genome integrity or degradation of untranscribed mRNAs. Particular attention is now paid to transcription termination due to its crucial role in the control of pervasive transcription. This type of genome-wide transcription, not associated with annotated genome features such as protein-coding genes, is a universal phenomenon for all the three domains of life and viruses (Georg & Hess, 2011; Wade & Grainger, 2014). In eukaryotes, pervasive transcription arises mainly from bidirectional promoters that synthesize both mRNA and diverse non-coding RNAs, but this phenomenon is also controlled by selective transcriptional termination (Kapranov et al., 2007; Schulz et al., 2013). Recently, an essential role of transcription termination in the control of pervasive transcription was demonstrated for both Gram-positive and Gram-negative model micro-organisms Bacillus subtilis and Escherichia coli (Nicolas et al., 2012; Peters et al., 2012).

In bacteria, transcription termination is achieved by two mechanisms: factor-independent (intrinsic) and factor-dependent termination. Intrinsic termination is strongly associated with sequence-specific signals characterized by a GC-rich symmetrical element followed by a ‘T stretch’ sequence encoding a RNA terminator hairpin with a ‘U-tract’ essential for pausing and disruption of the transcription elongation complex (TEC) (Gusarov & Nuelder, 1999; Epshtein et al., 2007). Recognition of this structure by TEC with the consequent release of RNAP does not require any additional factors (reviewed by Peters et al., 2011; Santangelo & Artsimovitch, 2011). Intrinsic terminators and terminator-like sequences were identified in >2000 sequenced bacterial chromosomes (Mitra et al., 2011). However, numerous bacteria are devoid of canonical intrinsic terminators downstream from putative transcriptional units, suggesting the existence of other termination mechanisms (Washio et al., 1998; de Hoon et al., 2005; Mitra et al., 2009; Peters et al., 2011).

Several factors were reported to act on RNAP during the elongation stage, causing TEC dissociation and release of...
E. coli Rho is known to be involved in the control of a variety of important biological processes, including (i) enforcement of transcription–translation coupling and termination of transcription of untranslated mRNAs (well known also as a phenomenon of Rho-dependent transcriptional polarity) (reviewed by Ciampi, 2006; Peters et al., 2011); (ii) suppression of pervasive antisense transcription (Peters et al., 2012); (iii) assistance in preventing deleterious R-loops (Harinarayanan & Gowerthshanak, 2003; Leela et al., 2013) and maintenance of genome integrity by prevention of conflicts between transcription and replication machineries (Dutta et al., 2011; Washburn & Gottesman, 2011); (iv) silencing of horizontally transferred DNA (Carminade et al., 2008; Menouni et al., 2013); and (v) regulation of gene expression mediated by small regulatory RNAs (sRNA) and riboswitches (Bossi et al., 2012; Hollands et al., 2012; Proshkin et al., 2014). Thus, Rho-dependent transcription termination plays an important role in linking transcription to other vital cellular processes.

Several of these processes have been shown to rely on Rho cooperating with various endogenous factors. Amongst them, the most important role was attributed to the transcription elongation factors NusG and NusA, which bind to Rho and RNAP, and modulate Rho-dependent termination at certain terminators (Burns et al., 1998; Ciampi, 2006; Peters et al., 2012; Boudvillain et al., 2013). Both nusG and nusA are conserved, but not always essential in different bacteria. In E. coli, other putative protein partners of Rho were identified as components of the interaction network containing conserved and essential proteins (Butland et al., 2005).

At present, the extent to which molecular mechanisms and physiological functions of Rho are conserved throughout the bacterial phyla remains elusive despite the fact that phylogenetic analysis has shown that rho is nearly ubiquitous in bacteria (D’Heygere et al., 2013). The current state of knowledge concerning Rho from bacterial species other than E. coli is restricted to several representatives. In each case, only discrete aspects of Rho functionality were examined. However, although limited, these analyses indicate that other bacteria deviate from the E. coli paradigm as the activity and function of Rho may be influenced by species-specific features, such as its cellular abundance, conservation of Rho protein partners or the occurrence of sequence determinants associated with regions of Rho-dependent termination (de Hoon et al., 2005).

In this review, we present general information about Rho and Rho-dependent termination in bacteria and summarize the ever-increasing knowledge about the various aspects of the physiological activity of Rho, such as limitation of deleterious foreign DNA expression, control of gene expression, suppression of pervasive transcription, prevention of R-loops and maintenance of chromosome integrity. Special attention is paid to the role of Rho in Bacillus subtilis where its inactivation has been shown to significantly modify the transcriptome (Nicolas et al., 2012). Based on updated data, we discuss the similarities and differences of Rho activity in different bacterial species.

Features and phylogeny of Rho

Transcription termination factor Rho was initially characterized in 1969 by J. W. Roberts as a factor that boosts the ‘accuracy’ of in vitro transcription on the bacteriophage λ DNA template by terminating RNAP at specific sites (Roberts, 1969). The progress in understanding the structure and mechanisms of action of Rho, based on E. coli Rho as an experimental model, provides a rather comprehensive view of Rho that is convenient to export to other bacterial species.

Rho is a homo-hexameric protein with ATP-dependent RNA helicase-translocase activity that causes the TEC to dissociate (reviewed by Richardson 2002; Peters et al., 2011; Boudvillain et al., 2013). Complex multistep binding of Rho to the nascent transcript involves different structural regions within the Rho hexamer known as primary and secondary binding sites (Richardson, 1982; Skordalakes & Berger, 2003; Skordalakes et al., 2005). The primary binding site is jointly formed by the N-terminal subdomains of Rho monomers. It is responsible for the initial, ATP-independent binding of Rho to rut sites – complex RNA sequences with a high cytidine/low guanosine content and relatively little secondary structure. Each monomeric subdomain comprises a characteristic OB-fold (oligonucleotide/oligosaccharide binding fold) able to bind two pyrimidine bases, preferentially cytosines (Bogden et al., 1999; reviewed by Ciampi, 2006; Peters et al., 2011). This explains Rho’s utilization of the rut sites for RNA binding. Upon RNA binding, the Rho hexamer adopts an asymmetrical ring conformation, with RNA enclosed within its central channel containing the secondary RNA-binding site (reviewed by Peters et al., 2011; Boudvillain et al., 2013). Contacts between RNA and the secondary site lead to activation of ATP binding by the C-proximal ATPase domains of Rho. ATP hydrolysis stimulates 5′→3′ translocation of Rho along the RNA and, finally, results in...
dissociation of TEC (reviewed by Richardson, 2002; Peters et al., 2011) (Fig. 1). Important details of the mechanism of Rho translocation and RNAP dissociation on Rho-dependent terminators are still debated and different models have been proposed (Park & Roberts, 2006; Epshtein et al., 2010; Koslover et al., 2012; Gocheva et al., 2015). Discussion of these models is beyond the scope of this review. As mentioned earlier, molecular mechanisms used by Rho were studied mainly for *E. coli* Rho protein, with several exceptions of *Micrococcus luteus* (Nowatzke & Richardson, 1996; Nowatzke et al., 1996, 1997a, b), *Rhodobacter sphaeroides* (Ingham, 1999) and *Mycobacterium tuberculosis* (Kalarickal et al., 2010; Mitra et al., 2014; D’Heygère et al., 2015) Rho proteins. Despite some controversial data on helicase activity of *Mycobacterium tuberculosis* Rho (Kalarickal et al., 2010; D’Heygère et al., 2015) and the existence of structural differences between Rho proteins (see below), the basic principles of the action of Rho are conserved across species (D’Heygère et al., 2015). Intimate knowledge of the structure of Rho was used for rational design of potential inhibitors of Rho activity in the Gram-negative coccobacillus *Brucella melitensis* – the infectious agent of brucellosis disease (Pradeepkiran et al., 2015).

The most complete phylogenetic analysis of Rho so far performed revealed that Rho is a well-conserved protein across different bacterial phyla, with the corresponding gene found in >90% of sequenced bacterial genomes (D’Heygère et al., 2013). Bacteria devoid of Rho, such as all *Cyanobacteria* and *Mollicutes*, but also some members of the *Clostridia*, *Bacilli* and *Negativicutes*, frequently contain small AT-rich genomes. Although it was proposed that Rho conservation is linked to some form of genome complexity, the evolutionary loss of *rho* by some bacteria defies explanation.

In this context it should be noted that the relative importance of Rho-dependent termination of transcription differs between bacterial species. Whilst most bacteria contain genes homologous to *rho*, the homologue is not necessarily an essential gene. Alongside *E. coli*, Rho is essential for the viability of *Salmonella enterica*, *Klebsiella aerogenes*, *Shigella flexneri* (Miloso et al., 1993), *R. sphaeroides* (Gomelsky & Kaplan, 1996), *Mycobacterium tuberculosis* (Sassetti et al., 2003; Griffin et al., 2011), *Bacteroides fragilis* (Veenanagouda et al., 2014), *Pseudomonas aeruginosa* (Morita et al., 2010) and *Micrococcus luteus* (Nowatzke et al., 1997a, b). However, it was found to be dispensable under conditions of growth in rich media in *Bacillus subtilis* (Quirk et al., 1993; Nicolas et al., 2012), *Streptomyces lividans* (Ingham et al., 1996) and *Staphylococcus aureus* (Washburn et al., 2001). Alternatively, a requirement for Rho activity can increase under specific conditions, e.g. in *Caulobacter crescentus*; otherwise dispensable Rho becomes essential for survival under oxidative stress (Italiani & Marques, 2005).

Comparative analysis of Rho homologues has demonstrated that the key residues involved in Rho’s oligomerization, RNA binding, ATP hydrolysis and RNA translocation are conserved through different species, consequently

![Fig. 1](http://mic.microbiologyresearch.org)

**Fig. 1.** Schematic representation of Rho-dependent termination (elements not to scale). During coupled transcription–translation of protein-coding genes, RNAP (in blue) is closely followed by ribosomes; the Rho utilization site (*rut*, in dark violet) is not accessible. In the absence of translating ribosomes, the open form of the Rho hexamer (in green) loads on nascent RNA (in brown) at the *rut* site through its primary binding site. Adaptation of the ring conformation and contacts between RNA and the Rho secondary binding site stimulate activation of ATP hydrolysis and 5′→3′ translocation of Rho along the RNA until Rho catches up with RNAP and dissociates the TEC.
suggested conserved mechanisms of action (D’Heygère et al., 2013). However, the RNA-binding domain of Rho shows more variability than its ATP-binding domain (Opperman & Richardson, 1994; Italiani & Marques, 2005). The most significant difference between Rho proteins corresponds to a large and variable insertion within the N-terminal domain found in ~30% of bacterial genomes. At present, the functional significance of these structural modifications of Rho is not well understood. Such an insertion was proposed to facilitate Rho–RNA binding in Actinobacteria (Nowatzke et al., 1997a, b; D’Heygère et al., 2013). A recent study in the Mycobacterium tuberculosis system has shown that deletion of the N-terminal insertion subdomain of Rho provokes RNA-binding defects and modifies ATPase activity (D’Heygère et al., 2015).

In the majority of bacteria, the size of Rho is ~420 aa (D’Heygère et al., 2013), but it varies amongst different species. For example, it is longer in Actinobacteria and Bacteroidetes due to extensions and/or insertions of the N-terminal domain. The longest Rho sequence was found in Thermoaerobacter marianensis (865 aa) — a marine extremophile belonging to the Firmicutes. The shortest Rho sequences, lacking some RNA-binding motifs, were detected in Colwellia psychrerythraea (314 aa) and Marinomonas sp. MWY1 (318 aa), both belonging to the Gammaproteobacteria. However, both genomes contain an additional full-sized rho ORF, carrying supplementary motifs (D’Heygère et al., 2013).

With regard to the involvement of Rho in the control of diverse processes in bacteria (see below), a better understanding of its evolutionary loss and/or conservation across bacterial phyla appears to be important.

**Regulation of gene expression**

Rho is known to enforce transcription–translation coupling by interrupting transcription of messages that are not translated (reviewed by Richardson, 1991). According to the E. coli model, under optimal translational conditions, the ribosome immediately following RNAP occludes the nascent RNA and consequently physically blocks the access of Rho to the rut sites which are presumed to occur frequently in mRNAs (Boudvillain et al., 2013). In the absence of translating ribosomes, Rho binds to the available rut sites of the nascent transcript and proceeds to terminate transcription. This underlies the well-known phenomenon of transcriptional polarity, when a nonsense mutation within a gene represses expression of the downstream genes in the same operon (Proshkin et al., 2010).

The above mechanism of Rho function was deduced for E. coli where Rho is an abundant protein, and the amount of Rho, in the form of a hexamer, corresponds to ~38–64% of the RNAP level (Li et al., 2014; Wang et al., 2015a). However, this model of Rho activity cannot be directly extrapolated to Gram-positive bacteria such as Bacillus subtilis and Staphylococcus aureus, given the low cellular abundance of Rho and low Rho/RNAP ratio in these bacteria (Ingham et al., 1999; Maass et al., 2011; Nicolas et al., 2012; Muntel et al., 2014). In Bacillus subtilis, the cellular level of Rho estimated using immunoblotting analysis does not exceed 80 hexamers per cell (Ingham et al., 1999). Similarly, the copy number of Rho determined by the absolute quantification of the Bacillus subtilis cytosolic proteome does not exceed 80 Rho hexamers per cell, which corresponds to ~0.8% of RNAP (Muntel et al., 2014). Measurement of Bacillus subtilis Rho cellular abundance using a GFP-Rho fusion estimated the Rho hexamers at ~5% of the level of RNAP (Nicolas et al., 2012). Thus, Bacillus subtilis Rho cannot be present at the majority of TECs, contrary to other general transcription factors such as NusG and NusA, whose cellular levels are near-equimolar to RNAP (Doherty et al., 2006; reviewed by Lewis et al., 2008).

Transcriptional polarity in Bacillus subtilis was described for the tryptophan biosynthesis trpEDCFBA operon (Babitzke & Gollnick, 2001; Yakhnin et al., 2001). Another example was reported for the rplJL operon encoding the ribosomal L10(L12)4 complex. This operon is regulated by an increased translation of the leader region that leads to reduced Rho-dependent termination and relief of transcriptional attenuation (Yakhnin et al., 2015). Despite the limited number of examples of Rho-dependent regulation, the potential role of Rho in transcription–translation coupling in Bacillus subtilis and other Gram-positive bacteria should not be underestimated. In Bacillus subtilis, Rho prevents synthesis of the untranslated antisense RNAs (asRNAs) initiated from 31 promoter sequences (Nicolas et al., 2012). This argues that despite its low availability, Rho acts to control transcription–translation coupling in Bacillus subtilis at the genome scale. One can suggest that in Bacillus subtilis some endogenous cellular factors may increase Rho affinity to individual untranslated mRNAs via protein-mediated recruitment mechanism.

New mechanisms of gene regulation which use Rho-dependent termination were revealed recently. In Salmonella enterica, translational inhibition leading to Rho-dependent transcription termination is due to the action of sRNA, as was shown by analysis of ChiX sRNA which negatively regulates the chiPQ operon involved in oligosaccharide uptake (Bossi et al., 2012). ChiX binds to the ribosome-binding site of the first gene of the operon, chiP. As a result, ribosome binding and translation of chiP are abolished, inducing premature Rho-dependent transcription termination within the early portion of the chiPQ operon. Consequently, the expression of the downstream chiQ gene is also downregulated (Bossi et al., 2012). In E. coli, the 109 nt long Spot 42 sRNA encoded by the spf gene is known to downregulate expression of galK, the third gene of the galactose operon galETKM, thereby inhibiting GalK production (Møller et al., 2002). Recent studies suggested that Spot 42 regulates galK expression at two levels: RNA degradation and enhancement of Rho-dependent transcription termination at the
galT–galK junction (Wang et al., 2015b). To explain how Rho assures termination at the end of the galT gene, Wang et al. (2015b) proposed that the binding site of Spot 42 on the galT–galK junction RNA overlaps with a putative rut site. Spot 42 binding enhances the disassembly of the ribosome at the stop codon of galT and exposes the rut site. Rho would finally catch up with the RNAS transcribing downstream DNA and terminate transcription (Wang et al., 2015b).

A new Rho-dependent mechanism was described for expression of Salmonella genes involved in Mg$^{2+}$ transport, suggesting that Rho links Mg$^{2+}$ uptake to translational signals (Kriner & Groisman, 2015). Expression of the Mg$^{2+}$ channel gene corA was shown to be regulated by a Rho-dependent terminator located within its 5′ leader region. Accessibility of the rut site depending on RNA conformation was shown to be modulated by translocation of corL, a short orf located within the corA gene (Kriner & Groisman, 2015).

Another example of gene regulation implicating Rho is the regulation of the E. coli pgaABCD operon by the CsrA protein (Figueroa-Bossi et al., 2014). The 5′ UTR of the pgaABCD operon was shown to contain a rut site which is sequestered by stable RNA secondary structure. Binding of CsrA to the RNA prevents formation of this secondary structure, thus making the rut site accessible for Rho binding, and consequently promotes Rho-dependent transcriptional attenuation (Figueroa-Bossi et al., 2014). The authors suggested that in terms of regulatory responses, transcription termination and anti-termination can be equated to repression and activation of transcription initiation (Figueroa-Bossi et al., 2014). This hypothesis is supported by the involvement of Rho in the widespread regulatory system using riboswitches.

In bacteria, riboswitches are RNA-based regulatory elements that control expression of biosynthetic and transport proteins as a result of binding to particular ligands (ions or metabolites) (reviewed by Mellin & Cossart, 2015). The significance of riboswitch function lies in the transduction of ligand binding into changes in expression of the downstream gene. Recently, it was shown that some riboswitches use Rho to attenuate transcription, thus linking Rho to the process of sensing and regulating gene expression in response to environmental cues (Proshkin et al., 2014).

Regulation of Rho-dependent termination by riboswitches was described for the Mg$^{2+}$-sensing mgtA riboswitch from Salmonella enterica (Hollands et al., 2012, 2014) for the flavin mononucleotide-sensing ribB riboswitch from E. coli (Hollands et al., 2012) and flavin mononucleotide riboswitch from the Gram-positive Corynebacterium glutamicum (Takemoto et al., 2015). In general, depending on the ligand concentration (e.g. Mg$^{2+}$ or flavin), the leader region of the riboswitch can exist in two alternative conformations. When a ligand is highly available, the riboswitch binds to it and exposes the rut site. Consequently, Rho interacts with the nascent RNA and induces transcription termination. Otherwise, at low ligand concentrations, the rut site is inaccessible and thus regular gene transcription can occur.

Furthermore, it was speculated that Rho-dependent transcription termination is a common, integral part of riboswitches that is actually underestimated. This point of view is based on the fact that a number of riboswitches found in different mRNA leader sequences in E. coli are deprived of obvious intrinsic, Rho-independent terminators (Proshkin et al., 2014).

It should be noted that for several proven cases of Rho-dependent gene regulation, the corresponding molecular mechanisms remain unknown currently. Among them are the Rho-dependent mechanisms of oxidative stress survival in Caulobacter crescentus and E. coli (Italiani & Marques, 2005; Kawamura et al., 2005) or repression of the osmotically regulated proU operon in E. coli and Salmonella enterica (Rajkumari & Gowrishankar, 2001). Rho-dependent regulation of the pyrimidine de novo biosynthesis pyr operon was demonstrated recently in Corynebacterium glutamicum. Rho inactivation in this bacterium leads to a two- to fourfold increase of mRNA levels of the pyrimidine biosynthesis genes (Tanaka et al., 2015). Interestingly, an opposite situation is observed with the pyr operon of Bacillus subtilis cells (Turner et al., 1994): transcriptome analysis of the Bacillus subtilis rho mutant revealed a net decrease of pyr operon transcription (Nicolas et al., 2012). This strongly suggests that in both Corynebacterium glutamicum and Bacillus subtilis, Rho regulates pyr operon expression, but by diverse unravelled regulatory mechanisms with opposite physiological effects.

Gram-positive bacteria also use a variety of regulatory mechanisms based on the association of proteins, sRNAs or metabolites with mRNA for control of gene expression (Mandal et al., 2003). In Bacillus subtilis, ~70 genes are known to be controlled by riboswitches (Mandal et al., 2003; Irnov et al., 2010). Recently, a wide variety of sRNAs was identified in Bacillus subtilis, Bacillus anthracis, Listeria monocytogenes, Staphylococcus aureus and Streptomyces coelicolor (Irnov et al., 2010 and references therein; Oliva et al., 2015). However, involvement of Rho-dependent termination in these regulatory mechanisms in Gram-positive bacteria remains to be demonstrated.

**Limitation of deleterious foreign DNA expression**

Among different physiological functions of Rho, silencing of horizontally transferred foreign DNA was revealed in E. coli within the past 10 years. Inactivation of Rho by the Rho-specific antibiotic bicyclomycin (BCM) in E. coli resulted in a global increase in the expression of prophages genes (Cardinale et al., 2008). Additionally, RNAP chromatin immunoprecipitation and microarray (‘ChIP-chip’) experiments in BCM-treated cells revealed a significant association of Rho-dependent terminators...
with foreign DNA, suggesting that horizontally transferred gene islands are 'hotspots' for Rho-dependent termination (Peters et al., 2009). Alternative in silico analysis confirmed this hypothesis and suggested that Rho can act as a part of the 'cellular immune mechanism' protecting against phage-related or xenogenic DNA not only in *E. coli*, but also in other proteobacteria species (Mitra & Nagaraja, 2012).

So far several hypotheses explaining this function of Rho have been proposed (reviewed by Peters et al., 2011). First, foreign (mostly phage-related) DNA may be rich in Rho-dependent terminators involved in the regulation of gene expression. For example, in *E. coli*, the Rho-specific terminator t$_{imm}$ blocks induction of toxin genes from the *rac* prophage (Cardinale et al., 2008). Rho has been also shown to control the lysogenic state of *E. coli* prophage KpiE1 by inhibiting the expression of the *torI* gene that mediates excisive recombination (Menouni et al., 2013). A second hypothesis implies that codon usage in foreign DNA could inhibit translation and thus expose rut sequences, which are Rho targets. It was also hypothesized that the insertion of foreign DNA into active transcriptional units alters the activity of natural terminators, disrupting translation–transcription coupling and consequently increasing recruitment of Rho (Peters et al., 2009). Finally, it was proposed that there could be progressive selection against hairpin-encoding sequences (like intrinsic terminators) to facilitate Rho action within horizontally acquired islands (Mitra & Nagaraja, 2012). Regardless of which hypothesis is accurate, Rho activity appears to be important for controlling expression of foreign DNA in *E. coli*.

An involvement of Rho in the control of horizontally transferred genes in other bacteria has not been addressed. It should be noted that *Rhodobacter sphaeroides* 2.4.1, *Bacteroides fragilis* 638R and *Mycobacterium tuberculosis* H37Rv strains in which *rho* was identified as an essential gene are derived from lysogenic strains (Gomelsky & Kaplan, 1996; Cole et al., 1998; Patrick et al., 2010; Griffin et al., 2011; Kontur et al., 2012). In contrast, a *Staphylococcus aureus* rho knockout mutant was obtained from prophage-free RN4220 strain (Washburn et al., 2001; Nair et al., 2011).

However, this activity of Rho appears not to be universal in bacteria as a lack of Rho does not stimulate expression of prophage-related genes in *Bacillus subtilis* (Nicolas et al., 2012). The largest asRNAs detected in *Bacillus subtilis* rho mutant strain were specific to ICE Bs1 element, SPβ prophage and SKIN element. These asRNAs could negatively interfere with expression of the corresponding genes, thus preventing prophage expression (see below). No asRNAs were detected for PBXS prophage, but transcription of the operon encoding putative phage structural proteins, terminase subunits and lysis functions was strongly down-regulated (Nicolas et al., 2012). These observations suggest that Rho could take part in regulation of foreign DNA expression in *Bacillus subtilis*, but differently than in *E. coli*. Expression of toxin genes *txpA* and *yonT* from SKIN and SPβ prophages was not significantly modified in the absence of Rho.

Importantly, although *E. coli* strains cured from prophages are relatively resistant to BCM, rho remains an essential gene in these strains (Cardinale et al., 2008; Tran et al., 2011; Washburn & Gottesman, 2011). This pinpoints other activities of Rho which are vital for viability.

### Suppression of pervasive transcription

Genome-wide overlapping transcription has been described for different bacterial transcriptomes (Dornenburg et al., 2010; Sharma et al., 2010; Georg & Hess, 2011; Nicolas et al., 2012; Peters et al., 2012; Voigt et al., 2014). This widespread phenomenon was designated 'pervasive' transcription, where non-canonical transcription is not delimited by any defined ends, which means that it can occur at nearly any place in the genome. Resulting transcripts are usually non-coding RNA, not demarcated by gene boundaries, and are frequently antisense (Wade & Grainger, 2014). Although the number of asRNAs with identified function is limited (reviewed by Thomason & Storz, 2010; Georg & Hess, 2011; Schultze et al., 2014), it is generally assumed that asRNAs could play a role in the regulation of gene expression via a variety of mechanisms, such as transcriptional interference, transcription attenuation, and modulation of degradation by nucleases and of ribosome binding (Thomason & Storz, 2010). Recently, a novel role of pervasive transcription in the surveillance of genome damage and efficient nucleotide excision repair was proposed (Kamarthapu & Nudler, 2015). However, it has also been suggested that pervasive transcription may have no functional role and be a form of transcriptional noise (Peters et al., 2012; Raghavan et al., 2012).

In any case, a high level of pervasive transcription could have deleterious effects, interfering with sense transcription, chromosome replication and genome stability (see below); it could also compromise cellular energy levels. This implies the existence of molecular mechanisms to control such type of transcription.

The role of Rho in the suppression of pervasive, primary antisense, transcription was demonstrated for both Gram-negative and Gram-positive model micro-organisms *E. coli* and *Bacillus subtilis* (Peters et al., 2009, 2012; Nicolas et al., 2012). In *E. coli*, most antisense transcription suppressed by Rho has been shown to arise either from a large uncharacterized set of antisense promoters within genes or from continuation of sense transcription past the ends of genes (read-through) into divergently oriented downstream genes. The mean size of these asRNAs is ~700 nt (Peters et al., 2012). Peters et al. (2012) observed that an increase in antisense transcription caused by sub-lethal inhibition of Rho did inhibit sense transcription, which is consistent with the idea that most antisense transcription is transcriptional noise. ChIP-chip experiments also showed that, for preventing antisense transcription,
Rho-specific termination is more important than intrinsic terminators as coding requirements for factor-independent terminators are not always consistent with a protein-coding gene on the opposite strand (Peters et al., 2009).

The major role of Rho in the suppression of pervasive transcription was also demonstrated in *Bacillus subtilis* (Nicolas et al., 2012). The study, aimed at examination of *Bacillus subtilis* WT and *rho* mutant transcriptomes, revealed that 13 % of the protein-coding genes of the WT strain were targeted by asRNAs. A majority of antisense transcripts arise from incomplete termination of transcription. Subsequently, transcriptome analysis of the *rho*-null strain revealed that in the absence of Rho, antisense transcription in *Bacillus subtilis* is largely increased. In total, Rho was shown to prevent antisense transcription in > 93 chromosomal regions, comprising 367 genes. In 62 regions, asRNAs corresponded to extended mRNA up to 16 000 nt long and were associated with Rho-specific or partially efficient intrinsic terminators at the 3' ends of transcriptional units. The rest of the Rho-controlled asRNAs were associated with a large increase in the activity of the promoters and could attain 12 000 nt with a mean size of ~ 5400 nt (Nicolas et al., 2012). This last observation led to the suggestion that *Bacillus subtilis* Rho might act very shortly after initiation and promote premature termination of transcription initiated at spurious promoter-like sequences across the genome due to the lack of coupling with translation. Interestingly, in some chromosomal regions of the *rho* mutant an increase of asRNA correlated with a decrease of sense RNA transcription (Nicolas et al., 2012; our unpublished results). This negative correlation between sense and antisense transcription suggests that, unlike *E. coli*, Rho-controlled asRNAs may influence gene expression in *Bacillus subtilis* cells. The biological significance of this observation needs further investigations.

Pervasive transcription can be also suppressed by several other mechanisms during transcription initiation and elongation, and by RNA degradation involving the activity of different bacterial proteins (reviewed by Wade & Grainger, 2014).

In *E. coli*, both the histone-like nucleoid-structuring protein H-NS and transcription factor NusG were shown to contribute to Rho-mediated suppression of antisense transcription (Saxena & Gowrishankar, 2011; Peters et al., 2012). H-NS binds DNA at high-affinity sites, and forms nucleoprotein filaments that spread on AT-rich DNA and bridge distant DNA sites (reviewed by Seshasayee, 2014; Landick et al., 2015). H-NS is known to play an important role in the silencing of horizontally acquired genes and in the suppression of non-coding transcription by inhibition of both transcription initiation and elongation (Saxena & Gowrishankar, 2011; Peters et al., 2012; Singh et al., 2014). H-NS filaments cause RNAP trapping at the promoter region by binding to AT-rich DNA, which is abundant in spurious promoters (Singh et al., 2014) and characteristic for many horizontally acquired genes in *E. coli* (Chandraprakash & Seshasayee, 2014). ChIP-chip analysis demonstrated a strong association between sites of H-NS filament formation and Rho-dependent termination (Peters et al., 2012). It was also shown that bridged H-NS filaments directly inhibit elongating RNAP and promote Rho-dependent termination by expanding the kinetic window for Rho action (Kotlajich et al., 2015).

Premature termination governed by Rho can be increased through the binding of NusG. NusG is a universally conserved transcription factor in prokaryotes. In *E. coli*, it was shown to be indispensable for termination of transcription governed by Rho (reviewed by Tomar & Artsimovitch, 2013). *E. coli* NusG physically couples transcription and translation as it binds to RNAP, and concurrently, either Rho or NusE—a ribosomal S10 protein (Burmann et al., 2010). Rho–NusG interaction was proven to increase efficiency of termination at weak *rut* sites, characterized by lower C/G ratio sequence (Peters et al., 2012; Shashni et al., 2014). It was also shown that NusG influences termination efficiency at ~ 20 % of antisense and sense factor-dependent terminators. However, some NusG homologous proteins regulate transcription processivity differently. In *Mycobacterium* species, NusG does not act as a transcription elongation factor and is unable to bind Rho, although it weakly stimulates intrinsic termination (Czyz et al., 2014; Kalyani et al., 2015). Similarly, *Bacillus subtilis* and *Thermus thermophilus* NusG proteins were shown to stimulate RNAP pausing rather than facilitate transcription elongation (Yakhnin et al., 2008; Sevostyanova & Artsimovitch, 2010).

Moreover, H-NS is absent from the *Bacillus subtilis* genome, while NusG is dispensable in this bacteria (Ingham et al., 1999). Therefore, the above mechanisms assisting Rho activity in *E. coli* are not conserved in other bacterial systems.

Finally, pervasive antisense transcription can be controlled at the level of RNA degradation (Lasal et al., 2011; Laalami et al., 2014). It is possible that, under certain circumstances, Rho can participate in this process, although the experimental data supporting this hypothesis are limited at present. For example, in *E. coli*, Rho was identified as a component of the RNase E-based ribonucleoprotein complex purified under specific oxygen-dependent conditions (Tuckerman et al., 2011). Cells expressing a mutant form of Rho showed a decrease in the half-life of bulk mRNA which was attributed to the altered RNA-binding activity of the mutated Rho protein (Sozhannan & Stitt, 1997). Deletion of *pccB*, encoding poly(A) polymerase I which polyadenylates and consequently destabilizes RNAs, or *rrpH*, encoding a pyrophosphohydrolase which triggers 5'-end-dependent mRNA degradation, renders *E. coli* cell more sensitive to Rho inactivation (Tran et al., 2011). However, Tran et al. (2011) do not exclude that these effects are indirect and can be due to activation of the alternative pathways for RNA degradation.

The most direct evidence for the involvement of Rho in RNA degradation comes from *Rhodobacter capsulatus*, where Rho was found to be a major component of the
RNase E-based RNA degradosome (Jäger et al., 2001). The level of Rho considerably increased under anaerobic growth, suggesting a role of Rho in the regulatory response to changing environment (Jäger et al., 2004). At present, no systematic analysis, either by classical genetic approaches or by transcriptomic studies, has been performed to verify and understand the intriguing link that appears to exist between Rho-dependent transcription termination and mRNA degradation.

Taken together, these data provide strong evidence that Rho is a major factor responsible for suppression of pervasive antisense transcription in bacteria. Future detailed analysis of Rho is certainly needed to understand the still elusive role of pervasive transcription in prokaryotes.

**Rho as a factor of genome stability**

Maintenance of genome stability is a recently discovered biological function of *E. coli* Rho. This activity of this Rho is tightly linked to the control of RNAP backtracking (i.e. spontaneous reversed translocation of RNAP on the template during transcription elongation) and prevention of R-loop formation. R-loops are the three-stranded RNA/DNA hybrids in which RNA is base-paired with its template DNA, leading to extrusion of the non-template single strand from the DNA duplex (Fig. 2). Although the mechanisms of R-loop formation are still debated, it is well documented that formation of RNA/DNA hybrids is a co-transcriptional process favoured by negative supercoiling of DNA, its high G + C content, the absence of RNA secondary structures, and the uncoupling of transcription and translation (reviewed by Drolet et al., 2003; Gowrishankar & Harinarayanan, 2004; Li & Manley, 2006; Dutta et al., 2011; Gowrishankar et al., 2013). Formation of R-loops is also stimulated by RNAP backtracking (Nudler et al., 1997; reviewed by Nudler, 2012). Normally, RNAP backtracking is minimized by translating ribosomes which closely follow the elongation complex and push it forward (Proshkin et al., 2010 and references therein). By premature termination of the untranslated transcripts, Rho plays an important role in the control of spontaneous RNAP backtracking and consequently limits R-loop formation (Gowrishankar & Harinarayanan, 2004).

R-loops have proved to be harmful for genome stability in all organisms, as they can provoke hyper-recombination, mutagenesis and formation of chromosomal double-

---

**Fig. 2.** Schematic representation of R-loop formation. Factors that favour R-loop formation (DNA high negative supercoiling, high DNA G+C content, uncoupled transcription–translation, RNAP backtracking and transcription–replication conflicts) are listed on the left. Anti-backtracking factors preventing R-loop formation (translating ribosomes, Rho, Mfd, GreA) and factors negatively affecting R-loops (RNase HI, helicase RecG active on DNA/DNA and RNA/DNA hybrids) are listed on the right.
strand breaks (DSBs), which are lethal to cells if not repaired (reviewed by Li & Manley, 2006; Aguilera & Garcia-Muse, 2012; Nudler, 2012; Gowrishankar et al., 2013; Wimberly et al., 2013). One of the main sources of DSBs in all genera is the recurrent collisions between the replication and transcription machineries. In bacteria, these collisions are frequent due to the absence of temporal separation of transcription and replication, and to a higher rate of replication fork progression, compared with the elongating RNAP (Washburn & Gottesman, 2011 and references therein). Replication–transcription collisions lead to the replication forks stalling and collapsing, and provoke the formation of DSBs by different mechanisms which, on their own, depend on the nature of the arrest (Washburn & Gottesman, 2011; De Septenville et al., 2012; Dutta et al., 2011).

It has been shown that detrimental effects of replication–transcription collisions can be suppressed by overproduction of RNase HI, an enzyme which specifically hydrolyses the RNA moiety of RNA/DNA hybrids (Boubakri et al., 2010; Dutta et al., 2011; Gan et al., 2011). These observations clearly indicate that R-loops contribute to DSB formation during replication–transcription collisions.

There is also a body of evidence implicating Rho in the avoidance of such deleterious events in E. coli. (i) Viability of BCM-treated cells depends on elongation factor GreA (Dutta et al., 2011) and DNA translocase Mfd (Washburn & Gottesman, 2011), both acting as anti-backtracking factors for RNAP in addition to Rho (reviewed by Borukhov et al., 2005). Cells become less sensitive to BCM (i.e. less dependent on Rho) by introduction of the rpoB35 mutation which destabilizes TECs and makes them less prone to backtracking (Trautinger & Lloyd, 2002; Washburn & Gottesman, 2011). (ii) E. coli rho and nusG missense mutants with reduced termination activity show an increased requirement for the anti-R-loop functions of RNase HI and RecG – an enzyme which unwinds the RNA moiety of the RNA/DNA hybrid (Harinarayan & Gowrishankar, 2003). Consistently, rho missense and recG knockout mutations are synthetically lethal (Harinarayan & Gowrishankar, 2003). Most strikingly, the lethal phenotype of E. coli cells deleted for rho or nusG genes can be suppressed by the phage T4-borne helicaseUvsW acting on the RNA/DNA hybrids (Leela et al., 2013). Leela et al. (2013) also provide evidence for an increased incidence of the R-loops in the chromosome of the nusG mutant defective for Rho-dependent termination. (iii) Inactivation of functions needed for the restoration of the collapsed replication forks and replication restart renders E. coli cells more sensitive to BCM (Washburn & Gottesman, 2011). This lethality is linked to the considerable increase of chromosomal DSBs upon BCM treatment (Washburn & Gottesman, 2011). Finally, (iv) DSB formation induced by BCM can be suppressed by overexpression of RNase HI or by addition of the replication inhibitor hydroxyurea, which indicates that the DSBs originate from collisions between transcription and replication machineries (Dutta et al., 2011; Washburn & Gottesman, 2011).

Taken together, these data prove that, by limiting RNAP backtracking and R-loop formation, Rho acts to avoid replication–transcription collisions, and consequently to diminish replication stress and DNA damage by breakage. Removal of the backtracked RNAP in front of replication forks and/or prevention of excessive genome-wide formation of R-loops are considered as the reasons of Rho essentiality in E. coli cells (Washburn & Gottesman, 2011; Leela et al., 2013; Gowrishankar et al., 2013). In the absence of Rho, cells fail to withstand massive DNA damage.

The extent to which Rho is implicated in the maintenance of genome stability in bacteria where it is non-essential has not yet been addressed experimentally. In the case of Bacillus subtilis, it seems plausible that the pervasive transcripts accumulated in the rho mutant, most of which are kilobases in size, could engage in the formation of R-loops. It is also remarkable that 28 out of the 31 promoters activated in the absence of Rho are oriented oppositely to chromosome replication (head-on orientation) (Nicolas et al., 2012). However, the Bacillus subtilis rho mutant grows normally in rich medium (Quirk et al., 1993; Nicolas et al., 2012), which is in sharp contrast to the low viability of cells experiencing head-on replication–transcription collisions in either the E. coli (Boubakri et al., 2010) or Bacillus subtilis (Srivatsan et al., 2010) model systems. The mechanisms underlying the robustness of Bacillus subtilis rho mutant cells are unknown and need to be established. It was proposed that RNase HI and RecG enzymes responsible for R-loop removal might be more active in bacteria where Rho is non-essential compared with E. coli (Gowrishankar et al., 2013). This interesting hypothesis awaits experimental validation, alongside the analysis of other functions potentially able to prevent and to repair deleterious consequences of the loss of Rho-dependent transcription termination in these bacteria.

Intrinsic inhibitors of Rho

The importance of Rho-dependent transcription termination for the fine control of different cellular processes is further supported by the occurrence of intrinsic negative regulators of Rho. Initially, inhibitors of Rho activity were discovered in bacteriophages, which develop several original strategies to suppress Rho-dependent terminators present in their genomes. One of the best-studied examples is N protein encoded by lambdoid phages. N protein, assisted by bacterial elongation factors, modifies RNAP and suppresses intrinsic and Rho-dependent terminators present in phage DNA, thereby assuring transcription of the middle and late phage genes (Mason et al., 1992). N protein overcomes Rho action in multiple ways: it forms an inactive complex with Rho–NusA, prevents Rho–RNAP interaction, removes NusA from Rho-dependent termination pathway and perturbs the Rho–NusG interaction (Muteeb et al., 2012). E. coli
cells encode two more Rho-specific inhibitor proteins, Psu and YaeO. The polarity suppression protein Psu is encoded by the E. coli defective prophage P4. Psu interacts with Rho specifically, thus affecting ATP binding and RNA-dependent ATP hydrolysis which may reduce Rho translocation along the RNA and thereby the termination efficiency (Pani et al., 2006). YaeO protein binds to the Rho hexamer in a 1:1 monomer/monomer ratio in the vicinity to the primary binding site and inhibits the early stages of Rho binding to RNA (Gutiérrez et al., 2007).

YaeO protein exhibits some topological similarities with the pleiotropic regulator of gene expression, RNA-binding protein Hfq (Gutiérrez et al., 2007). Recently, it was shown that E. coli Hfq also plays a specific role in Rho-dependent transcription regulation by direct association with Rho and trapping the Rho–RNA complex into an inactive configuration (Rabhi et al., 2011). Rabhi et al. (2011) suggested that functional Rho–Hfq interactions are frequent in E. coli, although the specifically targeted transcription units remain currently unknown. Interestingly, despite high conservation of the hfq gene in a wide range of bacterial genomes, it does not play an important role in regulation in Gram-positive bacteria. For example, in Bacillus subtilis, the absence of Hfq had no global effects on the transcriptome (Hämmerle et al., 2014; Rochat et al., 2015). As mentioned earlier, Hfq is a highly conserved protein, and the yaeO and psu homologous genes are present in some genomes of E. coli and related enterobacteria and several prophages. However, one cannot rule out the possibility that other proteins with Rho-specific inhibition activity exist in other bacterial species.

**Conclusions**

The importance of Rho-dependent transcription termination in bacteria is now commonly recognized. During the last two decades considerable progress has been made in our understanding of the structure and the molecular mechanism of Rho action, thus providing a solid basis for the study of its physiological roles in bacterial cell (Fig. 3). However, whilst the molecular mechanism of Rho activity, based on the E. coli Rho model, seems to be mostly conserved, some Rho features may vary considerably amongst different species. Recent data also suggest that apart from the universal functions, e.g. in the control of transcription–translation coupling and pervasive transcription, Rho might manage other functions more or less characteristic for different bacteria. Additional.

---

**Fig. 3.** Involvement of Rho-dependent termination in different inter-related cellular processes. TU, Transcriptional unit.
questions must be addressed by future experimental studies. How does Rho assure transcription termination of untranslated mRNAs in bacteria where it is present in low amounts? What is the extent of Rho employment in the control of horizontally transferred genes across bacterial phyla? What are the compensatory functions and back-ups for Rho activity in the control of genome stability in species where Rho is dispensable? Does Rho-controlled pervasive transcription have a regulatory role for genes expression? This list of the questions is certainly not exhaustive. Elucidation of the species-specific activities of Rho, and its structural and functional interactions with other proteins, promises to be gratifying for fundamental and applied research, especially relating to the discovery of new antimicrobial agents.

Acknowledgements

A. G.-M. was supported by a co-tutelle PhD scholarship funded by the French Government and Program ‘Partenariats Hubert Curien Polonium 2015’ funded by the French Ministry of Foreign Affairs and International Development. We thank Dr M. Mangan for careful reading of the manuscript. We are grateful to Professor Jodi Lindsay and two unknown referees for providing helpful suggestions which improved this manuscript.

References


