Review

Correspondence

M. Sofia Ciampi

ciampi@biologia.uniba.it

Rho-dependent terminators and transcription termination

M. Sofia Ciampi

Dipartimento di Genetica e Microbiologia, Università di Bari, Via Amendola 165/A, 70126 Bari, Italy

Rho-dependent transcription terminators participate in sophisticated genetic regulatory mechanisms, in both bacteria and phages; they occur in regulatory regions preceding the coding sequences of genes and within coding sequences, as well as at the end of transcriptional units, to prevent readthrough transcription. Most Rho-dependent terminators have been found in enteric bacteria, but they also occur in Gram-positive bacteria and may be widespread among bacteria. Rho-dependent termination requires both *cis*-acting elements, on the mRNA, and *trans*-acting factors. The only *cis*-acting element common to Rho-dependent terminators is richness in rC residues. Additional sequence elements have been observed at different Rho termination sites. These 'auxiliary elements' may assist in the termination process; they differ among terminators, their occurrence possibly depending on the function and sequence context of the terminator. Specific nucleotides required for termination have also been identified at Rho sites. Rho is the main factor required for termination; it is a ring-shaped hexameric protein with ATPase and helicase activities. NusG, NusA and NusB are additional factors participating in the termination process. Rho-dependent termination occurs by binding of Rho to ribosome-free mRNA, C-rich sites being good candidates for binding. Rho's ATPase is activated by Rho-mRNA binding, and provides the energy for Rho translocation along the mRNA; translocation requires sliding of the message into the central hole of the hexamer. When a polymerase pause site is encountered, the actual termination occurs, and the transcript is released by Rho's helicase activity. Many aspects of this process are still being studied. The isolation of mutants suppressing termination, site-directed mutagenesis of *cis*-acting elements in Rho-dependent termination, and biochemistry, are and will be contributing to unravelling the still undefined aspects of the Rho termination machinery. Analysis of the more sophisticated regulatory mechanisms relying on Rho-dependent termination may be crucial in identifying new essential elements for termination.

Introduction

Bacteria and their phages use two main modes of terminating transcription: Rho-independent or 'intrinsic' termination, mainly requiring elements located on the mRNA, and Rho-dependent termination, relying on both mRNA elements and *trans*-acting factors (reviewed by Richardson & Greenblatt, 1996). Rho-dependent termination has been intensively studied, and it has previously been extensively reviewed (Richardson, 2002, 2003). About half of the transcription terminators identified in *Escherichia coli* are Rho-dependent. These terminators lie at the natural end of genes or within cistrons and in control regions preceding the coding sequences of genes, where they play an important role in the regulation of gene expression.

Rho is a homohexameric protein with RNA-dependent ATPase and helicase activities (reviewed by Platt & Richardson, 1992), which binds to the mRNA. Essential sites on the mRNA are the Rho-binding site, known as the

Rho utilization site (*rut*), and a distal region where the transcripts are terminated at multiple transcription stop points (*tsp*) and released, mediated by Rho interaction with RNA polymerase (reviewed by Richardson & Greenblatt, 1996). This process requires a ribosome-free transcript of at least 85–97 nt (Hart & Roberts, 1994). This transcript exhibits a high-C low-G content, which is consistent with the known requirements for rC residues and for an unstructured RNA by the Rho activities (reviewed by Richardson & Greenblatt, 1996; for a summary see Guérin *et al.*, 1998).

Rho loads onto the mRNA at a binding site of 70–80 nt (Zhu & von Hippel, 1998a), smaller than the ribosome-free transcript required for the termination process. The Rho-binding site seems to be the primary and essential determinant for Rho termination activity (Guérin *et al.*, 1998). Thus, the C richness of Rho site sequences could very well be the main requirement for Rho binding, consistent with the fact that this is the only feature common to Rho-dependent terminators.

In addition to C richness, other sequence elements, termed ‘auxiliary elements’ in this review, have been observed only at some Rho-dependent termination sites. For instance, *rutA* and *rutB* elements have been described in the λ *tR1* terminator (Chen & Richardson, 1987), and suggested to provide a strong attachment site for Rho (Graham & Richardson, 1998) on the mRNA. Similar *rut* elements have been found at the *trp t'* terminator of *E. coli* (Zalatan *et al.*, 1993). Some Rho-dependent terminator regions show sequences with homology to *boxA* (Friedman & Olson, 1983), which has an antitermination function in the bacteriophage λ (Das, 1992) and in other systems (Heinrich *et al.*, 1995; Vogel & Jensen, 1995); interestingly, *boxA*-like sequences in the *tna* operon of *E. coli* (Gong & Yanofsky, 2002) and in the *his* operon of *Salmonella typhimurium* (Ciampi *et al.*, 1989; Carlomagno & Nappo, 2001) seem to be required for Rho-dependent termination. These and other potential ‘auxiliary elements’ (listed in Table 1), which are not common features of Rho-dependent terminators, could be part of a pool of elements, different combinations of which may contribute to Rho loading efficiency and to the efficiency of termination. Which ‘auxiliary element(s)’ occur at a specific Rho site may depend on the ‘local’ context of the terminator, and this would explain why not all terminators exhibit the same elements, other than C richness.

Rho binding to the message activates Rho’s 5’ to 3’ ATP dependent RNA–DNA helicase (Delagoutte & von Hippel, 2003), allowing translocation of Rho along the message, toward polymerase pause sites. At these key sites interactions with RNA polymerase (Guarente, 1979) lead to termination and release of transcripts (Yager & von Hippel, 1987). Rho protein also interacts with other Rho-dependent termination factors, such as NusG (Sullivan & Gottesman, 1992; for a summary see Burns *et al.*, 1999), NusA (Schmidt & Chamberlin, 1984; for a summary see Ingham *et al.*, 1999) and NusB (Carlomagno & Nappo, 2001), which play an important role in the release of the terminated transcripts.

In this article I review the features of several Rho-dependent terminators, with the intent of drawing a picture of the various contexts in which terminators occur and of the different roles that they play, in relationship to the

mechanism of Rho-dependent termination and to the regulation of gene expression. I also summarize older and more recent data on the various elements involved in factor-dependent termination, emphasizing important insights that affected this field and remaining problems. Models for Rho-dependent termination are also discussed.

I start with a survey of Rho-dependent terminators: for each terminator the facts most relevant to the role of the terminator and to the mechanism of termination are presented; similarities and differences among the terminators are highlighted. The terminators are classified into two groups: constitutive and regulated terminators (see Table 2). The regulated terminators are divided into two subgroups: conventionally regulated Rho-dependent terminators and intragenic terminators, most of the latter being those responsible for transcriptional polarity effects. The main focus of the discussion will not be on the regulatory mechanisms, rather on the features of the Rho-dependent terminators themselves. I then review *cis*-acting sequence elements and *trans*-acting factors required for Rho-dependent termination. The ‘auxiliary elements’ identified at some terminators are presented and discussed for the individual terminators. Finally, ideas and facts leading to the current models for Rho-dependent termination are summarized.

Constitutive Rho-dependent terminators

Some Rho-dependent terminators occur at the end of transcriptional units, where transcription stops independently of any regulatory mechanism. As expected, these terminators are constitutive. Three such terminators are described in this section: the *trp* and *tyrT* operon terminators of *E. coli*, and the gene IV transcript terminator of the *E. coli* bacteriophage ϕ 1 (Table 2).

trp operon of *E. coli*. One of the most investigated Rho-dependent terminators is *trp t'*, located at the very end of the *trp* operon of *E. coli*, terminating the synthesis of the polycistronic *trpA–E* mRNA (Platt, 1981). *In vitro* studies of a DNA fragment carrying the terminator region revealed that the terminator is Rho-dependent and that

Table 1. ‘Auxiliary elements’ associated with Rho-dependent termination sites

Location of terminator (name)		Auxiliary elements	
<i>trp</i> operon of <i>E. coli</i> (<i>trp t'</i>)	<i>rutA</i> -like and <i>rutB</i> -like	R1 and R2	
<i>tyrT</i> operon of <i>E. coli</i>			CAAUCAA
<i>cro</i> gene of λ (<i>tR0</i>)			CAACAA
Colicin E1 gene of <i>E. coli</i>			CAAACAA
Rightward operon of λ (<i>tR1</i>)	<i>rutA</i> and <i>rutB</i>	<i>boxA</i> -like	CAAUCAA
<i>hisG</i> gene of <i>S. typhimurium</i>		<i>boxA</i> -like	
<i>tna</i> operon of <i>E. coli</i>		<i>boxA</i> -like	
<i>ilv</i> operon of <i>E. coli</i>			CAAUCGA

Table 2. Classification and location of Rho-dependent terminators

Location of terminators (name)	Organism	Position of terminators
Constitutive terminators		
<i>trp</i> operon (<i>trp t'</i>)	<i>E. coli</i>	End of operon
<i>tyrT</i> operon	<i>E. coli</i>	End of operon
Gene IV	<i>E. coli</i> phage ϕ 1	End of gene
Conventionally regulated terminators		
Rightward operon (λ tR1)	<i>E. coli</i> phage λ	Between <i>cro</i> and <i>cII</i> genes
<i>tna</i> operon	<i>E. coli</i>	Between leader region and structural genes
<i>rho</i> gene	<i>E. coli</i>	Leader and 5' end regions
<i>rho</i> gene	<i>B. subtilis</i>	Leader and 5' end regions
Intragenic regulated terminators		
<i>his</i> operon	<i>S. typhimurium</i>	Within <i>hisG</i> and <i>hisC</i> coding regions
<i>trp</i> operon	<i>E. coli</i>	Within <i>trpE</i> coding region
<i>trp</i> operon	<i>B. subtilis</i>	Within <i>trpE</i> coding region
<i>ilv</i> operon	<i>E. coli</i>	Within <i>ilvGM</i> region
<i>lac</i> operon	<i>E. coli</i>	Within <i>lacZ</i> coding region
<i>gal</i> operon	<i>E. coli</i>	Within <i>galE</i> coding region
<i>malK-lamB</i> operon	<i>E. coli</i>	Within <i>lamB</i> coding region
<i>nifLA</i> operon	<i>K. pneumoniae</i>	Within <i>nifA</i> coding region
Left operon (<i>tim</i>)	<i>E. coli</i> phage P4	Within <i>kil</i> coding region

the actual sites of termination are spread over a region of about 100 nt (Wu *et al.*, 1981).

Transcription of the DNA fragment from the *gal* promoter showed that the activity of the terminator is not promoter-specific, allowing deletion analysis of the first 105 bp sequence of the region and identification of a Rho 'loading site' and of a pair of elements within it (summarized by Zalatan *et al.*, 1993). These elements are examples of what I have defined above as 'auxiliary elements'; that is, they are not common to all terminators, but they are found in the specific context of one or a few terminators. Indeed, the pair of elements observed at the *trp t'* terminator are similar to the *rutA* and *rutB* sequences identified in the λ tR1 terminator (see later) and proposed to be required for the ATPase activity of Rho (Chen & Richardson, 1987). In *trp t'*, the *rutB*-like element seems to be more important than the *rutA*-like element, which appears to be non-functional unless the *rutB*-like sequence is present. Two similar subelements, R1 and R2, within the Rho-loading site, have also been proposed to be important for Rho-dependent termination at *trp t'* (Zalatan *et al.*, 1993).

Efficient loading of Rho at *trp t'* guarantees efficient and precise termination at the downstream termination zone; in fact, introduction of a secondary structure within the Rho loading site both reduces termination efficiency and shifts the initial termination site downstream (Zhu & von Hippel, 1998a). Zhu & von Hippel (1998a) also showed that approximately 97 nt of transcript are required by *trp t'* to cause termination. They also found that termination efficiency depends on the sequence context at the 3' end of the transcripts; the transcript terminus probably affects the

interactions of components of the transcription complex with the transcription bubble and thereby affects pausing. Indeed, the efficiency of RNA release at the individual termination positions is controlled by a kinetic competition between RNA polymerase elongating the transcripts and Rho terminating them (Zhu & von Hippel, 1998b). These authors found no requirement for specific sequences, like the previously identified *rut* sites, under the conditions used in their experiments, which is consistent with the view that Rho loading occurs at sites that are simply devoid of secondary structures and that contain a number of rC residues compatible with the activation of the RNA-dependent ATPase of Rho.

tyrT operon of *E. coli*. Transcription termination of the *tyrT* operon of *E. coli*, encoding tRNA^{Tyr}, occurs at a Rho-dependent terminator, both *in vivo* (Rossi *et al.*, 1981) and *in vitro* (Küpper *et al.*, 1978), and is completely dependent on Rho-factor. The site of termination is A·T rich and the adjacent sequences on both sides are G·C rich (Küpper *et al.*, 1978). An A·T rich region preceded by a small cluster of G·C pairs is common to some other Rho-dependent terminators, thought to cause a pause in transcription downstream. Small hairpin structures have been found at some termination sites, just before the end of the transcript, but they are absent at the tRNA^{Tyr} terminator (Küpper *et al.*, 1978). On the other hand, the potential 'auxiliary sequence' CAAUCAA, observed at some Rho sites (see later), is also found at the *tyrT* operon terminator (Küpper *et al.*, 1978). Madden & Landy (1989) showed that progressive removal of the region preceding the tRNA^{Tyr} terminator progressively reduces Rho-dependent

termination activity, consistent with the idea that upstream sequences mediate the efficiency of Rho-dependent termination.

Gene IV of bacteriophage $\phi 1$. A peculiar type of terminator lies at the end of gene IV of the *E. coli* bacteriophage $\phi 1$. Readthrough transcription at this terminator is observed in nonsense suppressor or *rho* hosts (Moses & Model, 1984; La Farina *et al.*, 1990), indicating that translating ribosomes are responsible for suppression of termination, through occupation of a region of RNA which is not normally translated. The observations further indicate that Rho factor is required for the activity of the terminator. The terminator is peculiar in the sense that it seems to be Rho-dependent, according to the above observations, but not according to other features of the terminator site. Indeed, the region of RNA between the nonsense codon for translation termination of gene IV and the 3' end of the Rho-terminated transcripts is 63 nt long and not sufficient to guarantee Rho binding, according to the minimum size of ribosome-free transcript required for Rho termination; this suggests that the common concept of a Rho-binding site needs to be refined. An additional, unusual feature of this terminator is that the termination site occurs within a predicted, fairly stable stem-loop structure (La Farina *et al.*, 1990), in contrast with the generally accepted idea of the requirement of a mostly unstructured RNA by the Rho activities. Thus, the context in which Rho-dependent termination occurs at this site seems to be quite different from that of other characterized Rho sites. On the other hand, an analysis of the distal portion of the gene IV coding sequence revealed similarity with sequences observed in other Rho-terminated transcripts (La Farina *et al.*, 1990).

The 'auxiliary elements' or potential 'auxiliary elements' found at the constitutive Rho sites described above are not common to the terminators of this class (Table 1); thus, these elements don't seem to be specific to the role of constitutive terminators as natural terminators or to the location of these terminators at the end of transcriptional units. The next section describes Rho-dependent terminators whose activity is regulated.

Regulated Rho-dependent terminators

Most of the Rho-dependent terminators identified and studied are regulated by various and sophisticated mechanisms. Some show a conventional regulation, requiring *trans*-acting factors, and others are regulated by *cis*-acting elements, located within the coding regions, the latter being terminators responsible for natural polarity.

Conventionally regulated Rho-dependent terminators

Rightward operon of bacteriophage λ . Rho plays a central role in the life cycle of the *E. coli* bacteriophage λ , as a key function in the regulation of gene expression. The *tR1* terminator of λ is one of the best-studied Rho-dependent terminators. *tR1* lies within the major rightward operon

of λ , in the intercistronic region between the *cro* and *cII* genes, and it controls N-mediated expression of the downstream genes in the operon. Two regions were identified in λ *tR1*, *rut* and *tsp*. *rut*, in the mRNA, binds Rho. It was dissected into two elements, *rutA* and *rutB*, separated by the N-utilization element *boxB* (Chen & Richardson, 1987). The roles of these 'auxiliary elements' were tested *in vivo*, assaying the effects of deletions of *rut* sequences both on the ability of the mutant transcripts to bind Rho and to activate Rho ATPase and on the overall activity of the terminator (Graham & Richardson, 1998). These studies suggested that the *rut* site guarantees correct positioning of the RNA with respect to the Rho protein, for activation of ATPase, in addition to providing a strong attachment site for Rho. Removal of *rutA* has a more dramatic effect on the function of *tR1* than does removal of *rutB* (Graham & Richardson, 1998). Indeed, *rutA* RNA is more C-rich than *rutB* RNA, consistent with the proposed role of C residues in the interaction of Rho with the mRNA (McSwiggen *et al.*, 1988) and in the function of Rho-dependent terminators (Hart & Roberts, 1991; Zalatan & Platt, 1992). Deletion of the *boxB* sequence of the λ *cro* gene, which forms a small stem-loop structure in the RNA, has no effect on the affinity of the *cro* RNA for Rho, nor on the efficiency of termination at *tR1* (Graham & Richardson, 1998); this contrasts with the observation that small hairpin structures are generally found in transcripts to which Rho binds (Schneider *et al.*, 1993). Since no differences were observed between the findings *in vivo* and those in a purified *in vitro* system with Rho factor only (Graham & Richardson, 1998), possible interaction of other factors does not, apparently, interfere with Rho action in the *rut* region. The *in vitro* results seem to fully account for the *rut* functions *in vivo* (Graham & Richardson, 1998), and suggest that no other *cis*-acting sequences outside *rut* are necessary for *tR1* function.

The *rutA* element carries a sequence with only one mismatch with the consensus sequence of *boxA*, a site of interaction of the host-encoded NusB, NusG and S10 proteins, for antitermination in bacteriophage λ (Mogridge *et al.*, 1998). Similar 'auxiliary elements' have been identified in the *his* operon of *Salmonella* (Ciampi *et al.*, 1989; Carlomagno & Nappo, 2001) and in the *tna* operon of *E. coli* (Gong & Yanofsky, 2002) (Table 1), where they seem to be required for Rho-dependent termination.

The other region identified in λ *tR1*, *tsp*, carries three heterogeneous RNA stop points, which correspond to polymerase pause sites (Lau *et al.*, 1983; Richardson & Richardson, 1996). This region can be substituted with an alien sequence, carrying polymerase pause sites, without affecting termination at the transcription stop points (Richardson & Richardson, 1996), which shows independence of the *rut* and *tsp* action in the termination. Termination at one of the RNA stop points of λ *tR1* occurs at the sequence CAAUCAA preceded by a potential weak

stem-loop structure in the RNA transcript (Rosenberg *et al.*, 1978). Similar putative 'auxiliary elements' have also been found at the λ *tR0* Rho-dependent terminator (CAACAA) (Calva & Burgess, 1980), near Rho-dependent transcription stop points within the *ilv* operon (CAAUCGA) (Wek *et al.*, 1987) and at the Rho-dependent terminators following the tRNA^{Tyr} (CAAUCAA) (Küpper *et al.*, 1978) and the colicin E1 (CAAACAA) (Ebina & Nakazawa, 1983) genes (Table 1); other terminators do not have CAAUCAA-like sequences. Lau *et al.* (1984) tested the possibility that these features are sufficient to signal Rho-dependent termination at λ *tR1*, and found that they are insufficient, suggesting that there are additional sequence requirements for Rho action.

Specific residues required for Rho-RNA interaction at λ *tR1* have recently been identified by Graham (2004): single nucleotide substitutions at eight positions in a 25 nt region, previously shown to be critical for termination, reduce termination. These substitutions occur at the 5' end of *rutA* and introduce, in most cases, guanines, confirming Rho's idiosyncrasy for Gs; not surprisingly, many substitutions result in the loss of cytidine residues, confirming Rho's 'sympathy' for Cs.

The λ *tR1* terminator region shows several 'auxiliary elements' or potential 'auxiliary elements' (Table 1), as mentioned above. This suggests that the terminator may require the contribution of more elements for its activities than those needed at other terminator sites. These requirements may be dictated by the specific location of the terminator within the rightward operon of λ , in relation to its role of regulator of gene expression.

Some conventionally regulated Rho-dependent terminators occur in regulatory regions located at the beginning of transcriptional units. Such terminators have been found in the *tna* operon of *E. coli* and in the *rho* genes of *E. coli* and *Bacillus subtilis* (Table 2), and they are described below.

tna operon of *E. coli*. A Rho-dependent terminator, playing a key role in regulation of expression of the degradative tryptophanase (*tna*) operon of *E. coli* (Stewart *et al.*, 1986), lies between the sequences of the *tnaC* leader and of the first structural gene of the operon, *tnaA*. It prevents expression of the operon when no tryptophan is available and termination is inhibited when the inducer tryptophan is provided. Studies of *tna* constitutive mutants disclosed elements that participate in Rho-dependent termination, mainly a Rho-utilization site (Gollnick & Yanofsky, 1990) and a sequence with similarity to *boxA* (Konan & Yanofsky, 2000). These studies suggest that, in the absence of inducer, Rho binds to the leader portion of the *tna* transcript, moves 5' to 3' along the transcript until it reaches a paused polymerase, and then drives the elongation complex to terminate transcription. Operon induction requires the presence of a single, crucial, *trp* codon in the leader-peptide coding region and requires translation of *tnaC*. In the presence of tryptophan, the TnaC-peptidyl-tRNA remains bound to the translating ribosome, which

is not released at the *tnaC* stop codon, preventing Rho factor access to the *boxA* and *rut* sites, and consequently inhibiting termination (Gong & Yanofsky, 2003). This ribosome release inhibition appears to be the crucial event that regulates Rho-dependent termination in the *tna* operon leader region.

rho gene of *E. coli* and *B. subtilis*. The expression of the *E. coli rho* gene itself is regulated through Rho-dependent termination (Matsumoto *et al.*, 1986). In fact, Rho-dependent terminators have been identified in the mRNA leader and N-terminal regions of the *rho* gene, and they have been proposed to be responsible for reducing expression of *rho* in response to high levels of the coded protein.

Similarly, the *rho* gene of *B. subtilis* seems to use Rho-dependent terminators for autoregulation of the gene (Ingham *et al.*, 1999). The ability to regulate Rho levels may allow the cell to modulate gene expression by using Rho-dependent terminators strategically located between genes and their promoters. Interestingly, Rho-dependent termination also plays a role in *trp* operon expression of *B. subtilis* (see later), these two cases being, to the best of my knowledge, the only known examples of gene regulation by Rho-dependent termination in Gram-positive bacteria.

Intragenic regulated Rho-dependent terminators

It has been known for a long time that Rho-dependent terminators are responsible for premature blocks of transcription within operons, in response to premature translation termination or inefficient translation initiation caused by stress conditions. Such a response is the result of the fact that transcription and translation are coupled. These intragenic terminators are cryptic; in fact, they don't normally function, but are unmasked only by pathological conditions interfering with proper translation of polycistronic mRNAs and uncoupling transcription from translation as a consequence. Rho-dependent termination within operons is responsible for transcriptional polarity effects on expression of the genes downstream from the transcriptional block; such a mechanism may have evolved to conserve energy, by preventing the synthesis of unused transcripts (reviewed by Richardson, 1991). Rho-dependent terminators causing polarity effects have been identified and studied in a number of systems (Table 2). Several of these terminators are presented below.

his operon of *S. typhimurium*. In the first gene of the histidine (*his*) operon of *S. typhimurium*, *hisG*, a Rho-dependent terminator, is responsible for the polar effects caused by point mutations and by insertions of transposable elements Tn10 and Tn5 (Ciampi *et al.*, 1982; Ciampi & Roth, 1988; Wang & Roth, 1988). This terminator was identified by the isolation and characterization of mutants suppressing the strongly polar *hisG* frameshift mutation *hisG2148* (Ciampi *et al.*, 1982; Ciampi & Roth, 1988).

In vivo analysis of the transcripts produced by the polar mutant unveiled the existence of five major transcription end points in *hisG*, all affected by Rho to a different extent (Ciampi *et al.*, 1989). Three of these sites fall immediately downstream of sequences with the potential for forming secondary structures (Ciampi *et al.*, 1989) and one of them, the first in the gene and the most affected in a *rho* background, occurs within a sequence (Ciampi *et al.*, 1989) resembling the *boxA* consensus element. This *Salmonella his* sequence, cloned on a plasmid, was shown to enhance termination by interacting with the NusB factor of *E. coli* cells (Carlomagno & Nappo, 2001). *Salmonella hisG* sequences required for premature, Rho-dependent termination of transcription (Ciampi & Roth, 1988; Ciampi *et al.*, 1989) were also cloned on a plasmid and used to show that NusA modulates this intragenic termination, in *nusA E. coli* hosts (Carlomagno & Nappo, 2003). Three of the termination sites in *hisG* are also weakly active in the normally translating wild-type strain, similarly to what was observed in the *Salmonella hisC* cistron (Alifano *et al.*, 1988) and in the *E. coli lacZ* gene (Stanssens *et al.*, 1986). These sites may coincide with polymerase pause sites, as described in other systems (Yager & von Hippel, 1987).

The histidine operon of *S. typhimurium* proved particularly useful for the isolation and study of polarity suppressor mutants. Several mutations affecting the *hisG* terminator have been isolated, all of which are deletions, suggesting that no sequence-specific element is required for Rho-dependent termination. Indeed, four point mutations scattered along a stretch of 18 bp removed by the smallest of the deletions do not damage the terminator (Ciampi & Roth, 1988). Two of the deletions do not remove any of the five transcription end-points within *hisG* (G. Lisanti & M. S. Ciampi, unpublished results) and they may identify a binding site for Rho factor (Ciampi & Roth, 1988). Other deletions show that progressive removal of the sequence including the transcription stop points has an increasing suppressing effect on polarity, leading to progressive increase of downstream gene expression (G. Lisanti & M. S. Ciampi, unpublished results). Mutations suppressing polarity by reinitiation of transcription (Bianco *et al.*, 1998) or of protein synthesis (M. R. Pasca, A. La Teana & M. S. Ciampi, unpublished results) within *hisG* were also isolated, consistent with transcription terminating in *hisG* in polar mutants and with the known interference of translation on Rho-dependent termination.

A cryptic Rho-dependent terminator is also responsible for premature termination of transcription in the *hisC* cistron of *S. typhimurium* (Alifano *et al.*, 1988), where polarity suppressor mutations affecting initiation of translation or transcription have been isolated and studied (M. R. Pasca & M. S. Ciampi, unpublished results). Several 3' transcription end points were mapped within the *hisD* gene, immediately following *hisG* in the operon, and polar *hisA* mutants also produce shorter mRNAs (Alifano *et al.*, 1991), but possible action of Rho at these sites remains to be defined. Polar

effects are also caused by mutations in the three remaining internal genes of the *his* operon, *hisB*, *hisH* and *hisF* (Ciampi *et al.*, 1982); hence, cryptic Rho sites may be embedded in the coding sequences of these cistrons as well. *In vivo* studies with polarity suppressor mutants, including *rho* mutants, and *in vitro* transcription experiments with purified Rho on the *hisD*, *hisB*, *hisH*, *hisA* and *hisF* cistrons, would provide a complete picture of the distribution and features of Rho-dependent termination sites in the *his* operon. The scenery of the *his* operon could be common to most, if not all, bacterial operons; indeed, most operons may have evolved such a mechanism to cope with translation pathology.

trp operon of *E. coli* and *B. subtilis*. The first gene of the *trp* operon of *E. coli*, *trpE*, also shows intragenic Rho-dependent termination, which is the cause of polarity effects (Korn & Yanofsky, 1976). Investigation of *trpE* revealed several termination sites and some pause sites within the coding sequence of the gene; however, the study of Rho-dependent termination in the *trp* operon did not focus on the features of the terminator, but rather on the ability of the polynucleotide-binding sites of the Rho protein to interact with the mRNA. This was investigated in *rho* mutants showing increased polarity in the *trp* and *lac* operons, by Tsurushita *et al.* (1989), who found that the primary Rho-binding site directs the efficiency of termination, while the secondary site selects the termination points.

Rho-dependent termination plays a sophisticated role in the expression of the *trpE-A* operon of *B. subtilis*, the mechanism involving transcription attenuation and translational control. Translational control relies on the formation of an RNA hairpin that sequesters the *trpE* Shine-Dalgarno (SD) sequence, blocking translation. Formation of the SD-blocking hairpin causes transcriptional polarity, by allowing Rho's access to the nascent transcript, downstream from this RNA structure (Yakhnin *et al.*, 2001).

ilv operon of *E. coli*. Rho-dependent termination is responsible for transcriptional polarity in the *ilv* operon of *E. coli* K-12. Three Rho-dependent termination sites were identified and partially characterized by *in vitro* transcription experiments with a restriction fragment containing the promoter-distal portion of *ilvG* and the proximal portion of *ilvM*, fused to the *tac* promoter (Wek *et al.*, 1987). Several sequences with no dyad symmetry, required for Rho interaction with the mRNA, were identified: two of these sequences fall immediately upstream of two of the termination sites (Wek *et al.*, 1987). Regions with regularly spaced cytidine residues, observed in other Rho-dependent terminators (Platt, 1986), lie upstream of the transcription stop points in the *ilv'GM'* region, and the sequence CAAUCAA, found adjacent to some Rho-dependent termination sites (Lau *et al.*, 1984) (Table 1), is also present, with one mismatch (CAAUCGA), near the three transcription stop points in the *ilv'GM'* region (Wek *et al.*, 1987); however, a possible role for these potential

'auxiliary elements' in Rho-dependent termination remains to be established.

lac and *gal* operons of *E. coli*. Inefficient translation initiation causes premature transcription termination in the first gene of the *lac* operon of *E. coli*, *lacZ* (Stanssens *et al.*, 1986). Four cryptic Rho-dependent termination sites, which are turned on by upstream blocks of protein synthesis, as well as by amino acid starvation, were identified in *lacZ*, by *in vitro* and *in vivo* studies; they correspond to polymerase pause sites (Ruteshouser & Richardson, 1989). A detailed analysis of the sequence showed a complex interplay of polymerase-pause and terminator elements, widely discussed by these authors. Two of these terminators in the early region of *lacZ* (*tiZ1* and *tiZ2*) proved particularly useful in testing the requirement for the NusG and NusA factors in Rho-dependent termination: NusG has a stimulating effect on Rho-dependent termination in *lacZ*, while NusA inhibits termination (Burns *et al.*, 1998) (see section on *trans*-acting factors).

Cryptic Rho-dependent terminators responsible for polarity effects also lie in the *gal* operon of *E. coli* (De Crombrughe *et al.*, 1973). The termination sites have not been identified and characterized at the molecular level, but this system has mostly been used to study the requirement for the *E. coli* NusG protein in Rho-dependent termination. Rho-induced polarity in the *gal* operon is suppressed in NusG-deficient cells and Rho-dependent termination in *galE*, the first gene of the operon, is entirely prevented upon NusG depletion (Sullivan & Gottesman, 1992) (see also section on *trans*-acting factors).

malK-lamB operon of *E. coli*. The presence of a Rho-dependent terminator has been inferred within the *lamB* gene of the *malK-lamB* operon of *E. coli* K-12 (Colonna & Hofnung, 1981). Expression of *lamB*, the structural gene for the λ receptor (Thirion & Hofnung, 1972), results in phage sensitivity. Expression of the operon requires an activator, the *malT* gene product. A *malT* mutant is resistant to λ , since *lamB* is not expressed; however, in the absence of the *malT* gene product, expression of *lamB* and sensitivity to λ infection are restored in a *rho* mutant, due to the activity of a cryptic promoter, located at the distal end of *malK* and unmasked in the mutant *rho* (Colonna & Hofnung, 1981). Transcription originating at this promoter is terminated by Rho at the inferred terminator within *lamB*, in a *malT rho*⁺ strain, and this prevents expression of *lamB*. The terminator's structure and mechanism of function remain to be explored.

nifLA operon of *Klebsiella pneumoniae*. Intragenic Rho-dependent termination has been studied in the *nifLA* operon of *K. pneumoniae*, encoding the sensor-activator pair involved in the regulation of other *nif* genes. In this operon, altered translation of *nifL* activates Rho-dependent termination not in that gene, but in the downstream *nifA* (Govantes & Santero, 1996), which has been explained by

the fact that *L* and *A* are translationally coupled (Govantes *et al.*, 1998). These studies provide new insight into the putative role of regions of biased C and G composition in Rho-dependent termination. An analysis of transcripts terminated within the *his* operon, as well as in the *lacZ* cistron and at the Rho-dependent *trp t'* and λ *tR1* terminators, led Alifano *et al.* (1991) to hypothesize that a region of high cytosine-over-guanosine content in the mRNA is an essential feature of Rho-dependent terminators. Such C>G regions were also found at the *nifLA* operon's terminators; however, other C>G regions in the operon do not show termination activity (Govantes & Santero, 1996). Hence, a direct correlation between the occurrence of C>G regions and Rho-dependent termination could not be established by these authors, who suggest that the distribution of cytosines, yielding specific C/G profiles, may only be important in diagnosing the possible presence of a terminator.

Rho-dependent termination within coding sequences is not in all cases triggered by pathological conditions interfering with proper mRNA translation; in fact, intragenic termination can also be part of a physiological regulation of gene expression. One such case has been reported in the left operon of the *E. coli* bacteriophage P4.

Left operon of P4. A Rho-dependent terminator is involved in the establishment of P4 lysogeny, through a positive feedback mechanism on termination. The terminator, named *tim*, occurs within the coding sequence of the first translated gene, *kil*, of the left operon for replication of P4 and it seems to participate, with two upstream Rho-independent terminators, to attenuate expression of the left operon functions, early after infection (Briani *et al.*, 2000). Translation of *kil* inhibits termination at *tim* (Forti *et al.*, 1999), which may contribute to the production of a leader transcript, which is matured into a small, untranslated RNA, CI, a phage factor required for efficient termination. Once produced, CI acts as an anti-sense RNA, on the RNA leader of the left operon, where RNA-RNA interactions are essential for stringent termination at the downstream *tim* (Briani *et al.*, 2000), which in turn prevents expression of the replication genes and guarantees immunity against superinfecting phages.

Summary of Rho-dependent terminators

Some of the regulated Rho-dependent terminators described above lie in regulatory regions separating genes from their promoters, which may provide the cell with a convenient mechanism for the control of gene expression. Most regulated Rho-dependent terminators occur within coding sequences, contributing to preventing wasteful expression of genes, but this location is also used to direct gene expression. Rho-dependent termination can also serve to end transcriptional units; however, this role seems to be minor among the cases identified so far.

Looking back over all Rho-dependent termination sites, their features don't seem to depend on the common location or role of the terminators. 'Auxiliary elements' and potential 'auxiliary elements', observed at some of the Rho sites, do not appear to be specific to a specific class of terminators (Tables 1 and 2). These 'auxiliary elements' could be part of a pool of elements, one of which or combinations of which may contribute to Rho loading efficiency and to the overall efficiency of termination. The requirement for and effect on termination could be dictated by the 'local' context of the individual terminators, and this would explain why not all terminators exhibit the same 'auxiliary elements'. How the individual sequence elements contribute to the structure of Rho-dependent terminators and to the termination mechanism is not fully understood yet. Analysis of the more sophisticated mechanisms relying on Rho-dependent termination may be crucial in identifying new essential elements of the Rho-dependent termination machinery. The isolation of mutants relieving transcriptional polarity, site-directed mutagenesis of the various sequence elements, and biochemistry, will contribute to uncovering the secrets of Rho-dependent termination.

Cis-acting features of Rho-dependent terminators

Various *cis*-acting elements have been observed at different Rho-dependent termination sites and they are discussed, for the individual Rho sites, in the section on Rho-dependent terminators. However, the only feature common to Rho-dependent terminators seems to be richness of rC residues in the mRNA. In fact, the strongest activators of NTP hydrolysis by Rho are poly(rC) and rC-rich oligonucleotides. The requirements for rC residues by the Rho activity have been extensively studied (reviewed by Richardson & Greenblatt, 1996; summarized by Guérin *et al.*, 1998): site-directed mutagenesis of the *rut* site of the *trp* *t'* terminator did not identify a number or distribution of rC residues that could be directly correlated with the efficiency of termination (Zalatan & Platt, 1992). Recently, however, Graham (2004) showed that single nucleotide substitutions in the *rutA* region of λ *tR1*, which reduce termination, affect C residues and most of them introduce Gs. The C/G ratio of Rho-dependent terminators has been proposed to be an essential feature for termination (Alifano *et al.*, 1991), but it may only be indicative of the presence of a terminator (Govantes & Santero, 1996). Short rUrC oligomers stimulate Rho's ATPase more than do rC homo-oligomers (Wang & von Hippel, 1993); besides, a 30 nt run of pure rUrC, introduced between a promoter and a non-terminator site, is sufficient to cause transcription termination (Guérin *et al.*, 1998). This termination is dependent on the Salmonella Rho factor (identified by Housley & Whitfield, 1982), and is observed only in the absence of protein synthesis, consistent with the known incompatibility of Rho action with translation and with studies of Rho-dependent termination in the *hisG* gene of Salmonella (Ciampi *et al.*, 1982; Ciampi & Roth, 1988). According to the results of Guérin *et al.* (1998),

Rho-dependent termination seems to depend essentially on the Rho loading site.

Trans-acting factors in Rho-dependent termination

Rho factor. The hexameric protein Rho factor constitutes 0.1% of the total soluble protein of *E. coli* and *S. typhimurium* (Housley & Whitfield, 1982); it is an essential protein for the growth of *E. coli* and of its lambdoid phages (Das *et al.*, 1976), for *Rhodobacter sphaeroides* (Gomelsky & Kaplan, 1996) and for *Micrococcus luteus* (Nowatzke *et al.*, 1997); however, Rho is dispensable in *B. subtilis* (Quirk *et al.*, 1993) and in *Staphylococcus aureus* (Washburn *et al.*, 2001). Analysis of bacterial sequences suggests that *rho*-like homologues are ubiquitous in the bacteria and highly conserved (Opperman & Richardson, 1994; Washburn *et al.*, 2001).

Rho acts as a 5' to 3' ATP dependent RNA-DNA helicase, triggered by the interaction with the mRNA at the Rho loading site (reviewed by Richardson, 2003, and by Delagoutte & von Hippel, 2003). ATP hydrolysis provides the free energy for Rho translocation 5' to 3' along the message, until the protein encounters the transcription elongation complex stalled at a pause site. At these sites transcription is terminated and the RNA is dissociated from the transcription complex, by Rho's 5' to 3' helicase, which disrupts the transcript-template duplex, again powered by ATP hydrolysis.

Much of Rho's structure and function were known (reviewed by Richardson, 2002), prior to the crystallographic studies of Skordalakes & Berger (2003). Rho functions as a homohexameric ring (Finger & Richardson, 1982; Geiselmann *et al.*, 1992). Each Rho monomer bears two RNA-binding domains, the so-called primary and secondary sites (Richardson, 1982), located on the smaller N-terminal and larger C-terminal domains, respectively. Skordalakes and Berger's work shows that the two RNA-binding sites, the primary one, responsible for mRNA recognition, and the secondary one, required for translocation and unwinding, point toward the centre of the ring, and that the complexed hexameric ring is split open (likened to a 'lock washer'), a configuration likely required to allow Rho loading onto mRNA and translocation (reviewed by Richardson, 2003). Jeong *et al.* (2004), analysing the conformational changes of the Rho protein following nucleotide and nucleic acid binding, concluded that RNA binding changes the conformation of the Rho hexamer, resulting in opening of the subunit interfaces. ATP binding occurs and a new conformational change takes place, following RNA contact with the secondary sites lining the central channel of the Rho ring, which makes Rho capable of hydrolysing ATP and translocating.

NusG factor. Rho factor is sufficient to terminate transcription *in vitro* at most Rho-dependent terminators, whereas, *in vivo*, an additional factor, NusG, is required

by several Rho-dependent terminators. NusG is an essential, abundant protein in *E. coli* (Li *et al.*, 1992), whereas in *B. subtilis* NusG is not essential either for viability or for the autoregulation of Rho (Ingham *et al.*, 1999). NusG was originally isolated as a factor aiding the N protein in antitermination of transcription in bacteriophage λ (Li *et al.*, 1992); it is a transcription elongation factor, which stimulates the rate of transcription (Burova *et al.*, 1995) and enhances suppression of class II polymerase pause sites (Artsimovitch & Landick, 2000). Orthologue *nusG* genes exist in all the bacteria whose genomes have been sequenced, including *Mycoplasma genitalium*, which lacks other important transcription factors (Fraser *et al.*, 1995).

The requirement for NusG in Rho-dependent termination was first studied by Sullivan & Gottesman (1992), who showed that cells depleted of NusG have reduced efficiency of transcription termination at certain Rho-dependent terminators. However, some sites are fully dependent on NusG: Rho-dependent termination in *galE*, the first gene of the *gal* operon, is entirely prevented upon NusG depletion (Sullivan & Gottesman, 1992). *In vitro*, NusG is strongly required at the *tiZ1* terminator and its requirement is context dependent at *tiZ2* (Burns *et al.*, 1999). NusG does not greatly affect *in vitro* termination at the *trp t'* (Nehrke *et al.*, 1993) nor at the λ *tR1* (Burns *et al.*, 1999) terminators; however, *in vivo*, the λ *tR1* terminator is strongly dependent on NusG (Sullivan & Gottesman, 1992). The discrepancy between *in vitro* and *in vivo* requirements at λ *tR1* may be due to the more-limiting context for Rho action *in vivo*, where the presence of ribosomes translating the message creates more pressure on Rho's interaction with the mRNA. In fact, at *tiZ2* and λ *tR1* terminators NusG is required to overcome kinetic limitations that may be imposed on Rho function, depending on the properties of the specific terminator and on the position of the terminator in the transcriptional unit (Burns *et al.*, 1999); this requirement may involve NusG-mediated modulation of the sensitivity of the transcription complex to Rho action. Alternatively, the NusG effect at λ *tR1* *in vivo* could be the consequence of its involvement in other events important for Rho activity, such as ribosome release. Indeed, *nusG* orthologues have been found in two organisms that lack Rho, *Mycoplasma genitalium* (Fraser *et al.*, 1995) and *Methanococcus jannaschii* (Bult *et al.*, 1996), which suggests a NusG role other than in Rho-dependent termination.

NusG has been suggested to connect Rho to the RNA polymerase–RNA complex, thereby facilitating recognition of the termination signals (summarized by Burns *et al.*, 1999; reviewed by Richardson, 2002). Besides, the interaction of NusG with the RNA polymerase may cause this enzyme to become more susceptible to Rho actions that mediate release of the transcripts, as proposed by Burns *et al.* (1999), who have shown that NusG enhances the rate of release of the RNA from the transcription elongation complex. The NusG effect on release may also be the cause of the requirement for NusG at some terminators.

NusA factor. NusA has been reported to stimulate Rho-dependent termination (Kainz & Gourse, 1998); however, in most cases it inhibits Rho (Kainz & Gourse, 1998), consistent with the involvement of NusA with *nut* utilization in λ antitermination (Li *et al.*, 1992) and with rRNA antitermination (Vogel & Jensen, 1997). NusA is an essential protein in *E. coli* (Kainz & Gourse, 1998), but it becomes dispensable in *rho* mutants with reduced termination efficiency (Zheng & Friedman, 1994), indicating a critical role in Rho-dependent termination. In *B. subtilis* NusA is essential; however, its requirement is independent of Rho (Ingham *et al.*, 1999), which suggests the involvement of the protein in some process other than Rho-dependent termination. As for NusG, NusA orthologues exist in all the bacteria whose chromosomes have been sequenced.

NusA binds both to Rho factor and to *E. coli* RNA polymerase, which provides a way of coupling Rho to the elongation complex (Schmidt & Chamberlin, 1984). The Rho–NusA interaction may modulate Rho action at certain terminators. NusA is also a transcription elongation factor, which, unlike NusG, slows transcript elongation, by enhancing RNA polymerase pausing at some pause sites (Lau *et al.*, 1983; Artsimovitch & Landick, 2000). Zheng & Friedman (1994) have proposed that the ability of NusA to enhance pausing, which results in tight coupling of transcription and translation, would interfere with the Rho–mRNA interactions, blocking termination of transcription as a consequence. NusA influences both pausing and Rho-dependent termination at λ *tR1*, by interaction with the C-terminal domain of the α subunit of *E. coli* RNA polymerase (Kainz & Gourse, 1998); however, a direct correlation has not been established, in this system, between the efficiency of Rho-dependent termination and the extent of pausing, two processes that appear to relate to each other in a complex manner.

Burns *et al.* (1998) tested the combinatorial effects of NusA and NusG on transcription elongation and on Rho-dependent termination at the *tiZ1* and *tiZ2* intragenic terminators, and it seems that the two factors act independently and at different sites: NusG enhances elongation and Rho-dependent termination, while NusA decreases both, and these opposing effects are balanced when both factors act. It would be interesting to know whether NusA and NusG affect the structural features of the transcription complex at a pause site.

Models for Rho-dependent termination

Models have been proposed to explain how Rho reaches the elongation complex stalled at a site downstream from the initial Rho loading site on the mRNA substrate, and how the nascent transcript is released. The possibility that Rho acts at a distance, simply by looping of the RNA, has been ruled out (Steinmetz *et al.*, 1990), demanding that Rho translocates from the loading site to the termination site.

A physical model for the mechanism of translocation of Rho 5' to 3' along the mRNA was proposed by Geiselmann *et al.* (1993). This model, defined as the 'pure' tracking model, is a modification of the simple tracking model (Richardson, 1982), and it predicts that Rho detaches from the initial *rut* site of interaction to track 5' to 3' along the RNA chain to downstream sites, where the nascent RNA is released by the 5' to 3' RNA-DNA helicase activity of Rho (Geiselmann *et al.*, 1993).

Later studies by Horiguchi *et al.* (1997) on the quaternary geometry of Rho factor highlighted structural resemblances between Rho and other helicases, such as RuvB and T7gp4. Thus, the finding that RuvB and T7gp4 interact with the RNA chain within the hole of the hexameric ring led Miwa *et al.* (1995) to speculate that the lining of the hole is a site of interaction with the RNA also for Rho, by analogy with the RuvB and T7gp4 helicases. This hypothesis was supported by the finding that *rho* mutations affecting the interaction of Rho with the RNA fall within the portion of the protein participating in the tertiary structure of the ATP-binding domain of Rho, and extending into the hole (Miwa *et al.*, 1995). This finding, together with the possibility that Rho binding to the mRNA occurs by wrapping of the RNA chain around the N-terminal RNA-binding domains, located on one side of the Rho hexameric ring, led Horiguchi *et al.* (1997) to propose that Rho hooks the RNA through its N-terminal domain, and that interaction of the RNA with the central hole allows traction of the RNA.

A similar model had been proposed by Richardson & Richardson (1996), who suggested that, after the initial interaction of the 5' end of the Rho utilization site of a nascent transcript with the RNA-binding domain surface of Rho, a dissociation and reassociation of subunits would drive the 3' tail of the *rut* region of the transcript into the hole. This would be made possible by a break in the ring, observed in electron micrographs of unbound Rho (Gogol *et al.*, 1991). After capturing of the RNA in the hole, Rho would move 5' to 3' along the RNA, with the aid of ATP hydrolysis, which provides the energy for dissociation of the transcript from the RNA polymerase.

The models proposed by Richardson & Richardson (1996) and by Horiguchi *et al.* (1997) assign a role to a part of Rho, the central hole, not previously implicated in translocation of the protein along the RNA, and they remind one of another major mode of translocation suggested by studies of Faus & Richardson (1990) and of Steinmetz & Platt (1994), and known as the 'tethered tracking' mechanism. In this model Rho is thought to remain bound to the *rut* site while translocating 5' to 3' along the mRNA, until it reaches the actual termination sites. The translocation would be mediated by interaction of the RNA with a secondary site – proposed by Richardson (1982) – separate from the primary *rut*-binding site.

The points of interaction of Rho with RNA were investigated by Burgess & Richardson (2001) and by Wei & Richardson

(2001). Burgess & Richardson (2001) made Rho proteins with single cysteine residues, which were photo-cross-linked to various bound derivatives of λ *cro* RNA. Wei & Richardson (2001) explored the sites of Rho-RNA binding, through cleavage protection studies in the presence of bound ATP, RNA or DNA. These studies verified the existence of a secondary RNA-binding site on the Q-loop of Rho's ATP-binding domain, inside the central hole of the Rho hexamer. The results of Burgess & Richardson (2001) and of Wei & Richardson (2001) are consistent with binding of the *rut* region of the RNA around the crown of the Rho hexamer and with interaction of the 3' side of *rut* with the central hole, for activation of ATP hydrolysis and hence translocation. More recently Skordalakes & Berger (2003) determined the crystal structure of Rho bound to an ssDNA recognition site mimic or to an ssRNA and an ATP analogue. These structures show that Rho forms hexameric rings, with the primary RNA-binding sites facing the inside of the ring and leading the RNA 3' side toward the hole of the hexamer. The structures also show that the Rho ring is stably open in the complexed state. These structural features clarify how Rho captures the 3' side of the mRNA in the interior of the ring, before translocation starts.

Conclusions

Rho-dependent transcription termination may be widespread, but most Rho-dependent terminators characterized occur in enteric bacteria. Known examples of genes regulated by Rho in Gram-positive bacteria are the *rho* and *trp* genes of *B. subtilis*. Rho-dependent terminators lie in regulatory regions preceding genes, within coding sequences and at the end of transcriptional units. Some Rho-dependent terminators are constitutive, but most identified so far are regulated. This suggests that Rho-dependent termination, being a complex process involving transcription and translation elongation, is subject to control in many ways, allowing mechanisms to evolve to regulate termination by Rho, in response to a variety of conditions.

Rho is an RNA-binding protein with ATPase and helicase activities. It is an essential protein in *E. coli* and inessential in *B. subtilis*, but the two factors function similarly; therefore, the wider range of *rho* mutations one could isolate in *B. subtilis* might illuminate Rho's function in *E. coli* and likely in *S. typhimurium* as well. Two additional factors, NusG and NusA, participate in Rho-dependent termination, by coupling Rho to the elongation complex. NusG enhances the rate of release of the RNA from the transcription elongation complex, and it may mediate Rho action on RNA polymerase, leading to termination of transcription. NusA enhances pausing of RNA polymerase, and it may participate in termination by lowering the rate of transcription elongation. NusA is essential in *B. subtilis* for reasons that seem to only partially apply to *E. coli*. The protein may be a key to understanding the differences in transcription elongation and Rho-dependent termination between these two bacteria, and ultimately to understanding the Rho-dependent termination mechanism. Indications of possible

effects of NusG and NusA on the structure of the transcription complex would illuminate how these factors mediate Rho action.

The RNA transcript plays a key role in Rho function, by mediating interaction between Rho and the transcription elongation complex. Rho binding to *rut* on the mRNA activates Rho's RNA-dependent ATPase, which drives Rho 5' to 3' along the message, and the ATP-dependent RNA-DNA helicase, which 'peels' the transcript from its template. Rho-dependent terminators show a high content of C residues, required for activation of the Rho ATPase, and few Gs, minimizing RNA secondary structure and permitting Rho loading onto the transcript. These general features may be the only elements required for Rho-RNA binding on naked RNAs of appropriate length, even though specific nucleotides have been identified at some Rho-dependent terminators. Other terminator-specific *cis*-acting sequences, termed 'auxiliary elements' in this review, may aid in Rho loading, contributing to the efficiency of termination.

Models have been proposed to explain the mode of Rho-RNA interaction leading to termination of transcription. Rho-dependent termination is accomplished by the initial binding of Rho to *rut*, through the primary RNA-binding sites of the N-terminal domains of the Rho hexameric ring; subsequent sliding of the RNA into the central hole of the hexamer would then occur through a gap in the ring. Traction would translocate Rho 5' to 3' toward the RNA termination and release sites, while maintaining association with the initial binding site, as predicted by the 'tethered tracking' model, or following release of the initial binding site. Future structural studies will be needed to distinguish between these two models.

Acknowledgements

I am grateful to Glenn Herrick for many helpful discussions and for useful criticisms on earlier versions of this manuscript. I thank an anonymous referee for constructive comments. I apologize for not citing all relevant primary papers; this was not possible, due to space limitations. This article is dedicated to John R. Roth, who inspired it.

References

- Alifano, P., Ciampi, M. S., Nappo, A. G. & Carlomagno, M. S. (1988). In vivo analysis of the mechanism responsible for strong transcriptional polarity in a 'sense' mutant within an intercistronic region. *Cell* **55**, 351–360.
- Alifano, P., Rivellini, F., Limauro, D., Bruni, C. B. & Carlomagno, M. S. (1991). A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell* **64**, 553–563.
- Artsimovitch, I. & Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. *Proc Natl Acad Sci U S A* **97**, 7090–7095.
- Bianco, R., Moramarco, G., Stella, A. & Ciampi, M. S. (1998). Relief of transcriptional polarity by a mutation that creates a promoter in the *hisG* gene of *Salmonella typhimurium* LT2. *Mol Gen Genet* **257**, 529–533.
- Briani, F., Ghisotti, D. & Dehò, G. (2000). Antisense RNA-dependent transcription termination sites that modulate lysogenic development of satellite phage P4. *Mol Microbiol* **36**, 1124–1134.
- Bult, C. J., White, O., Olsen, G. J. & 20 other authors (1996). Complete genome sequence of the methanogenic archaeon *Methanococcus jannaschii*. *Science* **273**, 1058–1072.
- Burgess, B. R. & Richardson, J. P. (2001). RNA passes through the hole of the protein hexamer in the complex with the *Escherichia coli* Rho factor. *J Biol Chem* **276**, 4182–4189.
- Burns, C. M., Richardson, L. V. & Richardson, J. P. (1998). Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in *Escherichia coli*. *J Mol Biol* **278**, 307–316.
- Burns, C. M., Nowatzke, W. L. & Richardson, J. P. (1999). Activation of Rho-dependent transcription termination by NusG. Dependence on terminator location and acceleration of RNA release. *J Biol Chem* **274**, 5245–5251.
- Burova, E., Hung, S. C., Sagitov, V., Stitt, B. L. & Gottesman, M. E. (1995). *Escherichia coli* NusG protein stimulates transcription elongation rates *in vivo* and *in vitro*. *J Bacteriol* **177**, 1388–1392.
- Calva, E. & Burgess, R. R. (1980). Characterization of a Rho-dependent termination site within the *cro* gene of bacteriophage lambda. *J Biol Chem* **255**, 11017–11022.
- Carlomagno, M. S. & Nappo, A. (2001). The antiterminator NusB enhances termination at a sub-optimal Rho Site. *J Mol Biol* **309**, 19–28.
- Carlomagno, M. S. & Nappo, A. (2003). NusA modulates intragenic termination by different pathways. *Gene* **308**, 115–128.
- Chen, C. Y. A. & Richardson, J. P. (1987). Sequence elements essential for r-dependent transcription termination at λ *tR1*. *J Biol Chem* **262**, 11292–11299.
- Ciampi, M. S. & Roth, J. R. (1988). Polarity effects in the *hisG* gene of *Salmonella* require a site within the coding sequence. *Genetics* **118**, 193–202.
- Ciampi, M. S., Schmid, M. B. & Roth, J. R. (1982). Transposon Tn10 provides a promoter for transcription of adjacent sequence. *Proc Natl Acad Sci U S A* **79**, 5016–5017.
- Ciampi, M. S., Alifano, P., Nappo, A. G., Bruni, C. B. & Carlomagno, M. S. (1989). Features of the Rho-dependent transcription termination polar element within the *hisG* cistron of *Salmonella typhimurium*. *J Bacteriol* **171**, 4472–4478.
- Colonna, B. & Hofnung, M. (1981). *rho* mutations restore *lamB* expression in *E. coli* K12 strains with an inactive *malB* region. *Mol Gen Genet* **184**, 479–483.
- Das, A. (1992). How the phage lambda *N* gene product suppresses transcription termination: communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA. *J Bacteriol* **174**, 6711–6716.
- Das, A., Court, D. & Adhya, S. (1976). Isolation and characterization of conditional lethal mutants of *Escherichia coli* defective in transcription termination factor Rho. *Proc Natl Acad Sci U S A* **73**, 1959–1963.
- De Crombrughe, B., Adhya, S., Gottesman, M. & Pastan, I. (1973). Effect of Rho on transcription of bacterial operons. *Nature New Biol* **241**, 260–264.
- Delagoutte, E. & von Hippel, P. H. (2003). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: integration of helicases into cellular processes. *Q Rev Biophys* **36**, 1–69.
- Ebina, Y. & Nakazawa, A. (1983). Cyclic AMP-dependent initiation and Rho-dependent termination of Colicin E1 gene transcription. *J Biol Chem* **258**, 7072–7078.

- Faus, I. & Richardson, J. P. (1990). Structural and functional properties of the segments of lambda *cro* mRNA that interact with transcription termination factor Rho. *J Mol Biol* **212**, 53–66.
- Finger, L. R. & Richardson, J. P. (1982). Stabilization of the hexameric form of *Escherichia coli* protein Rho under ATP hydrolysis conditions. *J Mol Biol* **156**, 203–219.
- Forti, F., Polo, S., Lane, K. B., Six, E. W., Sironi, G., Dehò, G. & Ghisotti, D. (1999). Translation of two nested genes in bacteriophage P4 controls immunity-specific transcription termination. *J Bacteriol* **181**, 5225–5233.
- Fraser, C. M., Gocayne, J. D., White, O. & 26 other authors (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**, 397–403.
- Friedman, D. I. & Olson, E. R. (1983). Evidence that a nucleotide sequence, 'boxA', is involved in the action of the NusA protein. *Cell* **34**, 143–149.
- Geiselmann, J., Yager, T. D., Gill, S. C., Calmettes, P. & von Hippel, P. H. (1992). Physical properties of the *Escherichia coli* transcription termination factor Rho. 1. Association states and geometry of the Rho hexamer. *Biochemistry* **31**, 111–121.
- Geiselmann, J., Wang, Y., Seifried, S. E. & von Hippel, P. H. (1993). A physical model for the translocation and helicase activities of *Escherichia coli* transcription termination protein Rho. *Proc Natl Acad Sci U S A* **90**, 7754–7758.
- Gogol, E. P., Seifried, S. E. & von Hippel, P. H. (1991). Structure and assembly of the *Escherichia coli* transcription termination factor Rho and its interactions with RNA. I. Cryoelectron microscopic studies. *J Mol Biol* **221**, 1127–1138.
- Gollnick, P. & Yanofsky, C. (1990). tRNA(Trp) translation of leader peptide codon 12 and other factors that regulate expression of the tryptophanase operon. *J Bacteriol* **172**, 3100–3107.
- Gomelsky, M. & Kaplan, S. (1996). The *Rhodobacter sphaeroides* 2.4.1 *rho* gene: expression and genetic analysis of structure and function. *J Bacteriol* **178**, 1946–1954.
- Gong, F. & Yanofsky, C. (2002). Analysis of tryptophanase operon expression in vitro: accumulation of TnaC-peptidyl-tRNA in a release factor 2-depleted S-30 extract prevents Rho factor action, simulating induction. *J Biol Chem* **277**, 17095–17100.
- Gong, F. & Yanofsky, C. (2003). Rho's role in transcription attenuation in the *tna* operon of *E. coli*. *Methods Enzymol* **371**, 383–391.
- Govantes, F. & Santero, E. (1996). Transcription termination within the regulatory *nifLA* operon of *Klebsiella pneumoniae*. *Mol Gen Genet* **250**, 447–454.
- Govantes, F., Andujar, E. & Santero, E. (1998). Mechanism of translational coupling in the *nifLA* operon of *Klebsiella pneumoniae*. *EMBO J* **17**, 2368–2377.
- Graham, J. E. (2004). Sequence-specific Rho-RNA interactions in transcription termination. *Nucleic Acids Res* **32**, 3093–3100.
- Graham, J. E. & Richardson, J. P. (1998). *rut* sites in the nascent transcript mediate Rho-dependent transcription termination *in vivo*. *J Biol Chem* **273**, 20764–20769.
- Guarente, L. (1979). Restoration of termination by RNA polymerase mutations is *rho* allele-specific. *J Mol Biol* **129**, 295–304.
- Guérin, M., Robichon, N., Geiselmann, J. & Rahmouni, A. R. (1998). A simple polypyrimidine repeat acts as an artificial Rho-dependent terminator *in vivo* and *in vitro*. *Nucleic Acids Res* **26**, 4895–4900.
- Hart, C. M. & Roberts, J. W. (1991). Rho-dependent transcription termination. Characterization of the requirement for cytidine in the nascent transcript. *J Biol Chem* **266**, 24140–24148.
- Hart, C. M. & Roberts, J. W. (1994). Deletion analysis of the lambda *tR1* termination region: effect of sequences near the transcript release sites and the minimum length of Rho-dependent transcripts. *J Mol Biol* **237**, 255–265.
- Heinrich, T., Condon, C., Pfeiffer, T. & Hartmann, R. K. (1995). Point mutations in the leader *boxA* of a plasmid-encoded *Escherichia coli* *rrnB* operon cause defective antitermination *in vivo*. *J Bacteriol* **177**, 3793–3800.
- Horiguchi, T., Miwa, Y. & Shigesada, K. (1997). The quaternary geometry of transcription termination factor Rho: assignment by chemical cross-linking. *J Mol Biol* **269**, 514–528.
- Housley, P. R. & Whitfield, H. J. (1982). Transcription termination factor Rho from wild type and *rho-111* of *Salmonella typhimurium*. *J Biol Chem* **257**, 2569–2577.
- Ingham, C. J., Dennis, J. & Furneaux, P. A. (1999). Autogenous regulation of transcription termination factor Rho and the requirement for Nus factors in *Bacillus subtilis*. *Mol Microbiol* **31**, 651–663.
- Jeong, Y. J., Kim, D. E. & Patel, S. S. (2004). Nucleotide binding induces conformational changes in *Escherichia coli* transcription termination factor Rho. *J Biol Chem* **279**, 18370–18376.
- Kainz, M. & Gourse, R. L. (1998). The C-terminal domain of the alpha subunit of *Escherichia coli* RNA polymerase is required for efficient Rho-dependent transcription termination. *J Mol Biol* **284**, 1379–1390.
- Konan, K. V. & Yanofsky, C. (2000). Rho-dependent transcription termination in the *tna* operon of *Escherichia coli*: roles of the *boxA* sequence and the *rut* site. *J Bacteriol* **182**, 3981–3988.
- Korn, L. J. & Yanofsky, C. (1976). Polarity suppressors increase expression of the wild-type tryptophan operon of *Escherichia coli*. *J Mol Biol* **103**, 395–409.
- Küpper, H., Sekiya, T., Rosenberg, M., Egan, J. & Landy, A. (1978). A Rho-dependent termination site in the gene coding for tyrosine tRNA su3 of *Escherichia coli*. *Nature* **272**, 423–428.
- La Farina, M., Izzo, V., Costa, M. A., Barbier, R., Duro, G., Vitale, M. & Mutolo, V. (1990). Read-through transcription occurs at the Rho-dependent signal F1 TIV in suppressor cells. *Nucleic Acids Res* **18**, 865–870.
- Lau, L. F., Roberts, J. W. & Wu, R. (1983). RNA polymerase pausing and transcript release at the λ *tR1* terminator *in vitro*. *J Biol Chem* **258**, 9391–9397.
- Lau, L. F., Roberts, J. W., Wu, R., Georges, F. & Narang, S. A. (1984). A potential stem-loop structure and the sequence CAAUCAA in the transcript are insufficient to signal Rho-dependent transcription termination at lambda *tR1*. *Nucleic Acids Res* **12**, 1287–1299.
- Li, J., Horwitz, R., McCracken, S. & Greenblatt, J. (1992). NusG, a new *Escherichia coli* elongation factor involved in transcriptional antitermination by the N protein of phage λ . *J Biol Chem* **267**, 6012–6019.
- Madden, K. A. & Landy, A. (1989). Rho-dependent transcription termination in the *tyrT* operon of *Escherichia coli*. *Gene* **76**, 271–280.
- Matsumoto, Y., Shigesada, K., Hirano, M. & Imai, M. (1986). Auto-genous regulation of the gene for transcription termination factor Rho in *Escherichia coli*: localization and function of its attenuators. *J Bacteriol* **166**, 945–958.
- McSwiggen, J. A., Bear, D. G. & von Hippel, P. H. (1988). Interactions of *Escherichia coli* transcription termination factor Rho with RNA. I. Binding stoichiometries and free energies. *J Mol Biol* **199**, 609–622.
- Miwa, Y., Horiguchi, T. & Shigesada, K. (1995). Structural and functional dissections of transcription termination factor Rho by random mutagenesis. *J Mol Biol* **254**, 815–837.
- Mogridge, J., Mah, T. F. & Greenblatt, J. (1998). Involvement of *boxA* nucleotides in the formation of a stable ribonucleoprotein complex

- containing the bacteriophage λ N protein. *J Biol Chem* **273**, 4143–4148.
- Moses, P. B. & Model, P. (1984). A Rho-dependent transcription termination signal in bacteriophage Φ 1. *J Mol Biol* **172**, 1–22.
- Nehrke, K. W., Zalatan, F. & Platt, T. (1993). NusG alters Rho-dependent termination of transcription *in vitro* independent of kinetic coupling. *Gene Expression* **3**, 119–133.
- Nowatzke, W. L., Keller, E., Koch, G. & Richardson, J. P. (1997). Transcription termination factor Rho is essential for *M. luteus*. *J Bacteriol* **179**, 5238–5240.
- Opperman, T. & Richardson, J. P. (1994). Phylogenetic analysis of sequences from diverse bacteria with homology to the *E. coli* rho gene. *J Bacteriol* **176**, 5033–5043.
- Platt, T. (1981). Termination of transcription and its regulation in the tryptophan operon of *E. coli*. *Cell* **24**, 10–23.
- Platt, T. (1986). Transcription termination and the regulation of gene expression. *Annu Rev Biochem* **55**, 339–372.
- Platt, T. & Richardson, J. P. (1992). *Escherichia coli* Rho factor: protein and enzyme of transcription termination. In *Transcriptional Regulation*, pp. 365–388. Edited by S. L. McKnight & K. R. Yamamoto. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Quirk, P. G., Dunkley, E. A., Lee, P. & Krulwich, T. A. (1993). Identification of a putative *B. subtilis* rho gene. *J Bacteriol* **175**, 647–654.
- Richardson, J. P. (1982). Activation of Rho protein ATPase requires simultaneous interaction at two kinds of nucleic acid-binding sites. *J Biol Chem* **257**, 5760–5766.
- Richardson, J. P. (1991). Preventing the synthesis of unused transcripts by Rho factor. *Cell* **64**, 1047–1049.
- Richardson, J. P. (2002). Rho-dependent termination and ATPases in transcript termination. *Biochim Biophys Acta* **1577**, 251–260.
- Richardson, J. P. (2003). Loading Rho to terminate transcription. *Cell* **114**, 157–159.
- Richardson, J. P. & Greenblatt, J. L. (1996). Control of RNA chain elongation and termination. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, pp. 822–848. Edited by F. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Richardson, L. V. & Richardson, J. P. (1996). Rho-dependent termination of transcription is governed primarily by the upstream Rho utilization (*rut*) sequences of a terminator. *J Biol Chem* **271**, 21597–21603.
- Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. (1978). The relationship between function and DNA sequence in an intercistronic regulatory region in phage lambda. *Nature* **272**, 414–423.
- Rossi, J., Egan, J., Hudson, L. & Landy, A. (1981). The *tyrT* locus: termination and processing of a complex transcript. *Cell* **26**, 305–314.
- Ruteshouser, E. C. & Richardson, J. P. (1989). Identification and characterization of transcription termination sites in the *Escherichia coli* *lacZ* gene. *J Mol Biol* **208**, 23–43.
- Schmidt, M. C. & Chamberlin, M. J. (1984). Binding of Rho factor to *E. coli* RNA polymerase mediated by NusA protein. *J Biol Chem* **259**, 15000–15002.
- Schneider, D., Gold, L. & Platt, T. (1993). Selective enrichment of RNA species for tight binding to *Escherichia coli* Rho factor. *FASEB J* **7**, 201–207.
- Skordalakes, E. & Berger, J. M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* **114**, 135–146.
- Stanssens, P., Remaut, E. & Fiers, W. (1986). Inefficient translation initiation causes premature transcription termination in the *lac Z* gene. *Cell* **44**, 711–718.
- Steinmetz, E. J. & Platt, T. (1994). Evidence supporting a tethered tracking model for helicase activity of *Escherichia coli* Rho factor. *Proc Natl Acad Sci U S A* **91**, 1401–1405.
- Steinmetz, E. J., Brennen, C. A. & Platt, T. (1990). A short intervening structure can block Rho factor helicase action at a distance. *J Biol Chem* **265**, 18408–18413.
- Stewart, V., Landick, R. & Yanofsky, C. (1986). Rho-dependent transcription termination in the tryptophanase operon leader region of *Escherichia coli* K-12. *J Bacteriol* **166**, 217–223.
- Sullivan, S. L. & Gottesman, M. E. (1992). Requirement for *E. coli* NusG protein in factor-dependent transcription termination. *Cell* **68**, 989–994.
- Thirion, J. P. & Hofnung, M. (1972). On some genetic aspects of phage lambda resistance in *Escherichia coli* K12. *Genetics* **71**, 207–216.
- Tsurushita, N., Shigesada, K. & Imai, M. (1989). Mutant Rho factors with increased transcription termination activities. I. Functional correlations of the primary and secondary polynucleotide binding sites with the efficiency and site-selectivity of Rho-dependent termination. *J Mol Biol* **210**, 23–37.
- Vogel, U. & Jensen, K. F. (1995). Effects of the antiterminator *boxA* on transcription elongation kinetics and ppGpp inhibition of transcription elongation in *Escherichia coli*. *J Biol Chem* **270**, 18335–18340.
- Vogel, U. & Jensen, K. F. (1997). NusA is required for ribosomal anti-termination and for modulation of the transcription elongation rate of both antiterminated RNA and mRNA. *J Biol Chem* **272**, 12265–12271.
- Wang, A. & Roth, J. R. (1988). Activation of silent genes by transposons Tn5 and Tn10. *Genetics* **120**, 875–885.
- Wang, Y. & von Hippel, P. H. (1993). *Escherichia coli* transcription termination factor Rho. I. ATPase activation by oligonucleotide cofactors. *J Biol Chem* **268**, 13940–13946.
- Washburn, R. S., Marra, A., Bryant, A. P., Rosenberg, M. & Gentry, D. R. (2001). *rho* is not essential for viability or virulence in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **45**, 1099–1103.
- Wei, R. R. & Richardson, J. P. (2001). Identification of an RNA-binding site in the ATP binding domain of *Escherichia coli* Rho by H₂O₂/Fe-EDTA cleavage protection studies. *J Biol Chem* **276**, 28380–28387.
- Wek, R. C., Sameshima, J. H. & Hatfield, G. W. (1987). Rho-dependent transcriptional polarity in the *ilvGMEDA* operon of wild-type *Escherichia coli* K-12. *J Biol Chem* **262**, 15256–15261.
- Wu, A. M., Christie, G. E. & Platt, T. (1981). Tandem termination sites in the tryptophan operon of *Escherichia coli*. *Proc Natl Acad Sci U S A* **78**, 2913–2917.
- Yager, T. D. & von Hippel, P. H. (1987). Transcript elongation and termination in *Escherichia coli*. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, pp. 1241–1275. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Yakhnin, H., Babiarz, J. E., Yakhnin, A. V. & Babitzke, P. (2001). Expression of the *Bacillus subtilis* *trpEDCFBA* operon is influenced by translational coupling and Rho termination factor. *J Bacteriol* **183**, 5918–5926.
- Zalatan, F. & Platt, T. (1992). Effects of decreased cytosine content on Rho interaction with the Rho-dependent terminator *trp t'* in *Escherichia coli*. *J Biol Chem* **267**, 19082–19088.
- Zalatan, F., Galloway-Salvo, J. & Platt, T. (1993). Deletion analysis of the *Escherichia coli* Rho-dependent transcription terminator *trp t'*. *J Biol Chem* **268**, 17051–17056.
- Zheng, C. H. & Friedman, D. I. (1994). Reduced Rho-dependent termination permits NusA independent growth of *E. coli*. *Proc Natl Acad Sci U S A* **91**, 7543–7547.

Zhu, A. Q. & von Hippel, P. H. (1998a). Rho-dependent termination within the *trp* *t'* terminator. I. Effects of Rho loading and template sequence. *Biochemistry* **37**, 11202–11214.

Zhu, A. Q. & von Hippel, P. H. (1998b). Rho-dependent termination within the *trp* *t'* terminator. II. Effects of kinetic competition and Rho processivity. *Biochemistry* **37**, 11215–11222.