The bacterial transcription termination factor Rho coordinates Mg²⁺ homeostasis with translational signals

Michelle A. Kriner[1,2] and **Eduardo A. Groisman**[1,2,3,*]

[1]Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, CT 06536

[2]Yale Microbial Sciences Institute, Yale University, West Haven, CT 06516

[3]Howard Hughes Medical Institute

Abstract

The bacterial protein Rho triggers transcription termination at the ends of many operons and when transcription and translation become uncoupled. In addition to these genome-wide activities, Rho implements regulation of specific genes by dictating whether RNA polymerase terminates transcription within the 5' leader region or continues into the downstream coding region. Here, we report that the Mg^{2+} channel gene *corA* in *Salmonella enterica* serovar Typhimurium, which was previously thought to be constitutively expressed, is regulated by a Rho-dependent terminator located within its 5' leader region. We demonstrate that the unusually long and highly conserved *corA* leader mRNA can adopt two mutually exclusive conformations that determine whether or not Rho interacts with a Rho utilization (*rut*) site on the nascent RNA and thereby prevents transcription of the *corA* coding region. The RNA conformation that promotes Rho-dependent termination is favored by efficient translation of *corL*, a short open reading frame located within the *corA* leader. Thus, *corA* transcription is inversely coupled to *corL* translation. This mechanism resembles those governing expression of *Salmonella*'s other two Mg^{2+} transport genes, suggesting that Rho links Mg^{2+} uptake to translational signals.

Graphical Abstract
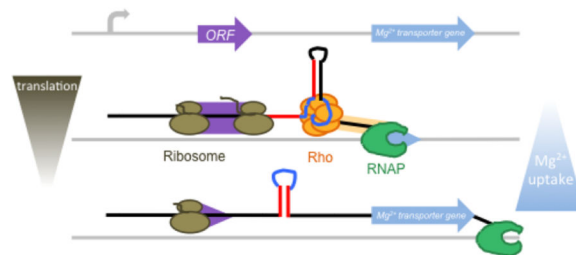
*To whom correspondence should be addressed. eduardo.groisman@yale.edu; Tel. (+ 1) 203-737-7940; Fax (+ 1) 203-737-2630.

INTRODUCTION

The transcription termination factor Rho is an ATP-dependent RNA helicase that performs critical roles in bacteria. For example, Rho defines the 3' end of many operons, suppresses antisense transcription, and implements transcriptional polarity by triggering RNA release from RNA polymerase (RNAP) when the normal coupling between transcription and translation is disrupted[1; 2; 3]. Whereas Rho executes these homeostatic functions genome-wide, it can also regulate expression of specific genes by controlling whether RNAP terminates transcription within a 5' leader region or continues into the associated coding region[4]. In such cases, the efficiency of Rho-dependent termination is modulated by particular physiological signals. Here, we establish that the mRNA leader of the Mg^{2+} channel gene *corA* in the bacterium *Salmonella enterica* serovar Typhimurium harbors a Rho-dependent terminator whose efficiency is inversely correlated with translation of a short open reading frame (ORF) in the *corA* leader.

Rho-dependent transcription termination is a multi-step process. First, a primary binding site (PBS) located on an exposed face of the Rho hexamer must recognize a Rho utilization (*rut*) site on a nascent RNA[5; 6]. The PBS contains six clefts (one from each monomer) that can each accommodate a YC (CC or UC) dinucleotide[7; 8]. Correspondingly, *rut* sites tend to be pyrimidine-rich, at least 70 nt in length and relatively unstructured[9; 10; 11]. Once the PBS binds an RNA substrate, transient opening of the ring-shaped protein allows a downstream RNA segment to thread into the central channel and contact a secondary binding site[8; 12]. This secondary interaction activates Rho's ATPase and helicase activities to initiate translocation along the RNA in a 5' to 3' direction[13; 14]. When a translocating Rho molecule catches up to a paused RNAP, it induces dissociation of the elongation complex[2; 15].

The accessibility of *rut* sites within 5' mRNA leaders can be modulated by a variety of intracellular signals including ions[16], small molecule metabolites[16], amino acid availability[17], sRNAs[18] and proteins[19]. These disparate signals all act via one of two mechanisms. In one mechanism, sequences required for recognition by Rho are sequestered or exposed by selective RNA secondary structure formation. For example, binding of the protein CsrA to the *pgaA* leader or the small molecule flavin mononucleotide to the *ribB* leader in *Escherichia coli* both unfold RNA secondary structures that normally sequester a *rut* site[16; 19]. Because transcription and translation occur at the same time and location in bacteria, the presence or absence of translating ribosomes can also control *rut* site accessibility. For instance, excess tryptophan promotes ribosome stalling during translation of a short ORF in the leader region of the *E. coli tnaA* mRNA, thereby exposing a *rut* site that would normally be occluded by a ribosome[17]. Similarly, the *Salmonella* sRNA ChiX can base pair with the *chiPQ* transcript at the *chiP* translation start site to uncouple transcription and translation, thus allowing Rho to abort transcription of *chiQ*[18].

The leader mRNA of the *Salmonella* Mg^{2+} transporter gene *mgtA* utilizes both of the above mechanisms to regulate Rho-dependent transcription termination. That is, the formation of RNA secondary structures controlling *rut* site availability is independently determined by direct binding of Mg^{2+} to the *mgtA* leader mRNA and by translation of *mgtL*, a short ORF in the *mgtA* leader[20]. High levels of cytoplasmic Mg^{2+} and efficient *mgtL* translation both promote an RNA conformation that facilitates Rho access to a *rut* site[16; 20]. By contrast, low cytoplasmic Mg^{2+} or a reduction in availability of proline-charged tRNA^{Pro} (which leads to ribosome stalling during translation of *mgtL*) favor an alternative RNA conformation that sequesters the *rut* site within a stem. Notably, the architecture of the *mgtA* leader region creates a situation opposite to canonical polarity; whereas Rho generally aborts transcription of incompletely translated messages, it instead attenuates transcription of *mgtA* when translation of *mgtL* is efficient[20]. A similar mechanism regulates another *Salmonella* Mg^{2+} transporter gene, *mgtB*; transcription of the *mgtB* coding region is inversely related to translation of a short ORF in its mRNA leader via a Rho-dependent mechanism[21; 22].

Here we report that the *corA* gene, which specifies the third and primary Mg^{2+} transporter operating in *Salmonella*, harbors a Rho-dependent transcription terminator in its 5' leader region. We determine that the *corA* leader mRNA is highly conserved and can adopt two mutually exclusive conformations. One conformation promotes premature transcription termination by exposing sequences required for recognition by Rho, whereas the alternative conformation sequesters these regions to allow read-through by RNAP. Whether the *corA* leader RNA adopts one conformation or the other is dictated by translation of the short ORF *corL*; efficient translation of *corL* favors the RNA conformation that presents an accessible *rut* site. Thus, expression of all three Mg^{2+} transporters in *Salmonella* is governed by Rho-mediated inverse coupling of transcription and translation, suggesting that this mechanism represents a general strategy for linking Mg^{2+} uptake to the translational status of the cell.

RESULTS

A long, highly conserved mRNA leader represses expression of the *corA* coding region

Given that 5' leader mRNAs in *Salmonella* have an average length of 20–65 nucleotides (nt)[23], the presence of a >400 nt intergenic region upstream of the *corA* start codon (Fig. 1A) raises the possibility that an unusually long mRNA leader precedes the *corA* coding region. To determine at which position *corA* transcription begins, we performed primer extension on total RNA harvested from *Salmonella* grown to mid-log phase in N-minimal medium. Reverse transcription yielded a single band corresponding to a position 238 nt upstream of the *corA* start codon (Fig. 1B). This transcription start site, hereafter denoted nucleotide +1, has near-consensus −10 and −35 promoter elements (Fig. 1A) and matches that previously identified via RNA-seq performed using a different *Salmonella* strain and growth condition[23]. Therefore, a 238 nt 5' leader precedes the coding region in the *corA* transcript.

The nucleotide sequence of the *corA* promoter and leader region is highly conserved among members of the family Enterobacteriaceae (Fig. 1A). In addition, all examined homologs of the *Salmonella corA* leader contain a short ORF, termed *corL*, whose existence was first reported in a ribosome profiling study in *E. coli*[24]. *corL*'s position, length (18 sense codons)

and use of a non-canonical UUG start codon are universally conserved, as is the deduced amino acid sequence of its first six codons (Fig. 1A,C).

If the *corA* leader mRNA is required for normal expression of the *corA* coding region, then deletion of the corresponding DNA sequence from the *Salmonella* chromosome should alter the abundance of the CorA protein. Indeed, replacement of the chromosomal *corA* leader with an unrelated 85 nt sequence increased CorA protein levels 10-fold compared to the wild-type (Fig. 1D). This result indicates that the *corA* leader represses expression of the *corA* coding region.

The *corA* leader mRNA contains a Rho-dependent transcription terminator

That the *corA* leader reduces CorA abundance raises the possibility that it attenuates transcription or translation of the coding region. To determine whether the *corA* leader affects transcription of the associated coding region, we compared the β-galactosidase activity of wild-type *Salmonella* harboring plasmid pYS1000, in which a constitutive promoter drives transcription of the *E. coli lacZ* gene, or its derivative pYS1040, in which the *corA* leader is inserted upstream of the *lacZ* Shine-Dalgarno (SD) sequence. The β-galactosidase activity originating from the pYS1040-containing strain was 54-fold lower than that of the isogenic strain harboring pYS1000, indicating that the *corA* leader represses transcription of its associated coding region (Fig. 2A).

Because the *corA* leader acts at the level of transcription (Fig. 2A) but contains no sequences resembling an intrinsic termination signal (Fig. 1A), we predicted that it might harbor a Rho-dependent transcription terminator. To test this hypothesis, we compared the β-galactosidase activity of wild-type *Salmonella* harboring plasmid pYS1040 in the presence or absence of the Rho-specific inhibitor bicyclomycin (BCM). Because Rho is essential to viability[25], cells were exposed for 30 min to a sub-lethal dose. The β-galactosidase activity was 5-fold higher in cells treated with BCM compared to the mock-treated control (Fig. 2B), suggesting the presence of a Rho-dependent terminator in the *corA* leader. This conclusion is supported by the observation of increased RNAP occupancy directly upstream of *corA* in *E. coli* treated with BCM[26].

To test directly for Rho-dependent termination in the *corA* leader region, we performed single-round *in vitro* transcription assays using purified *Salmonella* RNAP and a DNA template containing the *corA* leader and 49 nt of the coding region under the control of a constitutive promoter. In the absence of Rho, RNAP transcribed full-length RNA (Fig. 2C). Addition of purified *Salmonella* Rho protein resulted in accumulation of premature termination products and a significant decrease in the amount of run-off transcript, consistent with the presence of a Rho-dependent transcription terminator (Fig. 2C). Rho-dependent termination *in vitro* takes place at a series of sites beginning slightly upstream of the *corA* SD sequence and ending shortly after the *corA* start codon, resulting in truncated transcripts of ~230–260 nt in length (Fig. 2C). Cumulatively, these results indicate that the *corA* leader harbors a Rho-dependent transcription terminator.

The *corA* leader mRNA can adopt two mutually exclusive conformations

Based on its position within a leader region, we hypothesized that the function of the *corA* Rho-dependent terminator is regulated. To explore how the *corA* leader RNA might modulate the efficiency of Rho-dependent termination in the *corA* leader, we used the M-fold web server to analyze RNA secondary structures that the *corA* leader might adopt. We identified two possible RNA conformations whose regions of base pairing overlap such that they are mutually exclusive (Fig. 3A). The potential for alternative base pairing is conserved across species (Fig. 1A, compare colored bars to Fig. 3A). The previously hypothesized stem-loop C[27] is predicted to be thermodynamically favored over the competing stem-loop B structure ($\delta G_C = -20.5$ kcal/mol versus $\delta G_B = -16.8$ kcal/mol). However, the stem-loop A +B conformation has a kinetic advantage over the stem-loop C+D conformation because nucleotides G24-G30 and C80-C86 have the opportunity to base pair before nucleotides G101-G111 emerge from RNAP (Fig. 3A).

To determine whether the *corA* leader mRNA can adopt the predicted secondary structures, in-line probing assays were performed by incubating *in vitro*-transcribed, 5'-radiolabeled *corA* leader mRNA at room temperature for 40 h and analyzing degradation products by denaturing gel electrophoresis. Base paired nucleotides generally cannot sample the in-line conformation that facilitates spontaneous cleavage of the RNA sugar-phosphate backbone and thus appear with reduced intensity on a gel[28]. Regions predicted to participate in secondary structures in both the A+B and C+D conformations, such as the right "ear" of stem-loop A ([A57-G62, C68-U73], Fig. 3A), displayed the expected reduction in band intensity (Fig. S1). In the case of regions predicted to base pair differentially in the A+B versus C+D conformations, the cleavage pattern suggests the presence of two RNA subpopulations, one existing in each conformation. For example, U108-C112 should be single-stranded in the A+B conformation and double-stranded in the C+D conformation, whereas the reverse is true for C92-U100 and U118-A120 (Fig. 3A), yet all three of these regions were strongly cleaved during the in-line probing reaction (Fig. S1).

To clearly define the two identified conformations, we performed in-line probing of RNAs harboring mutations anticipated to favor formation of one conformation or the other (Fig. S1). In-line probing performed with RNA containing a mutation predicted to disrupt stem-loop B (M3 [T100A], Fig. 3A) demonstrated formation of stem-loops C and D due to complete protection of nucleotides within their stems (C74-U78, C80-C90, G101-G111, G113-G117, G121-C125, G143-C147) as well as strong cleavage within their loops (C92-U100, C126-C140) and the linker between the stems (U118-A120) (Fig. S1). By contrast, in-line probing of an RNA harboring a mutation predicted to disrupt stem-loop C (M1 [G84C G85C C86G], Fig. 3A) demonstrated formation of stem-loop B due to increased cleavage at positions that should be single-stranded only in the A+B conformation (C74-C90, U108-C112) and decreased cleavage at positions that should be double-stranded only in the A+B conformation (U96-U100, U118-A120) (Fig. S1).

Cumulatively, the in-line probing results indicate that the *corA* leader can adopt the mutually exclusive conformations consisting of stem-loops A+B versus C+D (Fig. 3A). However, because the *corL* sequence is embedded in stem-loop A, *corL* translation is predicted to preclude formation of stem-loop A *in vivo* (Fig. 3B).

Two pyrimidine-rich regions flanking stem-loop B constitute the *rut* site

To understand how the structure of the *corA* leader mRNA might influence the function of the Rho-dependent terminator, we sought to characterize the *rut* site required for Rho-dependent termination in the *corA* leader. We identified two regions at positions U87-U97 and C125-C140 that have high pyrimidine content and thus might contribute to recognition of the *corA* leader as a Rho substrate. Mutations that reduce the pyrimidine content within these regions (M8 [C90T C91G C92 T93A T94G], M9 [T128A T129A C130G] and M10 [C138A C139G C140A], Fig. 3A) increased expression *in vivo* (Fig. 4A) and decreased Rho-dependent termination *in vitro* (Fig. 4B). The effect was largest when mutations on both sides of stem-loop B were combined (Fig. 4A,B), indicating that both pyrimidine-rich tracts are required for efficient Rho-dependent termination.

If the effects of mutations in the pyrimidine-rich regions are due to *rut* site disruption, then the corresponding RNAs should exhibit a decreased ability to stimulate Rho's ATPase activity *in vitro*. As hypothesized, RNAs harboring mutations that reduce the pyrimidine content of U87-U97 and C125-C140 stimulated Rho's ATPase activity to a much lesser extent than wild-type RNA (Fig. 4C). The pyrimidine-rich tract downstream of stem-loop B appears to play a larger role than the upstream one because mutation M9+M10 had a larger effect than M8, both *in vitro* and *in vivo* (Fig. 4A–C).

Rho is thought to require at least 70 nucleotides to successfully load onto a transcript[10]. However, the *rut* site nucleotides identified in the *corA* leader span a region of ~50 nt in length, from positions 87–140. Therefore, we examined whether other pyrimidines in the *corA* leader are specifically required for Rho-dependent transcription termination. Removal of pyrimidines from positions 66–68 or 74–77 actually decreased expression *in vivo* (Fig. S2A), which is inconsistent with participation of these nucleotides in Rho loading. Removal of pyrimidines from positions 78–80, 176–177, 186–188 or 194–196 slightly increase expression *in vivo* but did not reduce the ability of the corresponding RNAs to stimulate Rho's ATPase activity *in vitro* (Fig. S2). These results indicate that RNA contacts made by Rho outside of positions 87–140 may not be sequence-specific. Notably, the location of the two pyrimidine-rich regions relative to stem-loops B and C (Fig. 3A) suggests that the RNA conformation adopted by the *corA* leader might control Rho's ability to terminate transcription before the *corA* coding region.

Stem-loops C and D are required for transcription of the *corA* coding region

To determine the effect of the of *corA* leader mRNA conformation on transcription of the associated coding region, we measured the β-galactosidase activity originating from wild-type *Salmonella* harboring plasmid pYS1040 or derivatives containing point mutations expected to favor or disrupt a particular RNA structure. Because a single nucleotide substitution often disrupts more than one structure, both arms of a given stem-loop were disrupted independently. Moreover, by combining the mutations in both arms to restore base pairing, we could assess whether the structure itself, rather than the identity of the mutated nucleotides, was responsible for the observed behavior.

Stem-loop C promotes transcription elongation into the associated coding region because a mutation predicted to prevent formation of stem-loop C (M1, Fig. 3A) abolished β-galactosidase activity (Fig. 5A). Even though this mutation also compromises formation of stem-loop A, the observed effect is specifically due to disruption of stem-loop C because: (i) a mutation that disrupts the opposite arm of stem-loop A (M1* [G24C, C25G, C26G], Fig. 3A) did not abolish expression (Fig. 5A); (ii) a double mutation that restores stem-loop A base pairing potential while disrupting stem-loop C (M1+M1*, Fig. 3A) had the same phenotype as the M1 single mutant (Fig. 5A); and (iii) stem-loop A is unlikely to form *in vivo* due to translation of *corL* (Fig. 3A). A mutation that disrupts formation of both stem-loops B and C (M2 [G105C C106G C107G], Fig. 3A) increased expression slightly (Fig. 5B). By contrast, combination of the M1 and M2 mutations, which restores stem-loop C base pairing (and disrupts stem-loops A and B), dramatically increased expression (Fig. 5B), providing further evidence that stem-loop C favors transcription of the downstream coding region.

Stem-loop B hinders expression because mutations that disrupt base pairing within stem-loop B (M3 [T100A] and M4 [A120T], Fig. 3A) significantly increased expression, whereas restoration of stem-loop B base-pairing potential by combining M3 with M4 resulted in wild-type expression (Fig. 5C). That the M3+M4 mutant displayed higher expression than M1 (Fig. 5A and 5C) likely reflects that the former allows formation of stem-loop C whereas the latter does not (Fig. 3A).

Stem-loop C is expected to facilitate stem-loop D formation by preventing adoption of the alternative stem-loop B (Fig. 3A). Given that stem-loop C promotes transcription of the downstream coding region, we wondered whether this was also the case for stem-loop D. Indeed, mutations that disrupt base pairing within stem-loop D (M5 [G121T G123T] and M6 [C143A C145A], Fig. 3A) abolished expression (Fig. 5D). The effect of the M5 mutation is specifically due to disruption of stem-loop D and not B because other mutations predicted to prevent formation of stem-loop D without disrupting stem-loop B (M5* [G123A], M5+M7 [T97G C99G G121T G123T], Fig. 3) also eliminated expression (Fig. S3). Restoration of base pairing within stem-loop D via combination of the M5 and M6 mutations restored wild-type expression (Fig. 5D). Cumulatively, these data indicate that stem-loop D is necessary to promote transcription into the associated coding region.

Formation of stem-loop D alone is not sufficient to promote constitutive transcription elongation because a triple mutant that retains stem-loop D but cannot form stem-loop C (M1+M5+M6, Fig. 3A) showed no expression (Fig. 5D). Taken together, the above results indicate that stem-loops C and D are both required for expression of the downstream coding region.

Stem-loops C and D promote transcription elongation by reducing Rho-dependent termination

If the lack of expression observed when stem-loops C or D are disrupted is due to Rho-promoted transcription termination, then inhibition of Rho activity should restore expression to mutants that produce RNAs unable to adopt stem-loops C or D. In agreement with this notion, BCM restored β-galactosidase activity to wild-type *Salmonella* harboring derivatives

of plasmid pYS1040 with mutations that disrupt stem-loops C or D (M1 or M5, respectively, Fig. 5E). That BCM treatment of M1+M2 slightly increased expression (Fig. 5E) suggests that some Rho-dependent termination still occurs when the RNA adopts the stem-loop C+D conformation.

In vitro transcription experiments directly demonstrated that disruption of stem-loops C or D increases the efficiency of Rho-dependent termination (M1 and M5, Fig. 5F). By contrast, locking the RNA in the stem-loop C+D conformation reduced termination efficiency relative to the wild-type RNA (M1+M2, Fig. 5F). Together these results indicate that stem-loops C and D promote expression of the coding region by reducing the efficiency of Rho-dependent termination within the *corA* leader.

The stem-loop C+D conformation inhibits transcription termination by decreasing Rho's ability to access the *rut* site

The pyrimidine-rich tracts comprising the *rut* site are single-stranded in the stem-loop B conformation but appear to be structurally constrained when the *corA* leader adopts the stem-loop C+D conformation (Fig. 3A). Thus, we reasoned that stem-loops C and D might hinder Rho-dependent termination by reducing Rho's ability to access the *rut* site. In agreement with this notion, RNAs harboring mutations that disrupt stem-loops C or D (M1 or M5, respectively) stimulated Rho's ATPase activity to a much greater extent than wild-type RNA (Fig. 6A). Conversely, an RNA locked in the stem-loop C+D conformation (M1+M2) elicited similar ATPase activity to the wild-type leader RNA (Fig. 6A). This result suggests that the wild-type RNA adopts the stem-loop C+D conformation under the assay conditions used.

If the stem-loop C+D conformation hinders Rho-dependent termination by reducing *rut* accessibility, then mutations in the pyrimidine-rich tracts constituting the *rut* site should be epistatic to mutations that disrupt stem-loops C or D. As predicted, mutations in the *rut* site (M8, M9+M10) restored expression to mutants in which stem-loops C or D are disrupted (Fig. 6B). (In this experiment, the alternative mutations M1B (G84A G85A C88G) and M5* (G123A) were used to disrupt stem-loops C and D, respectively, because the M1 and M5 mutations introduce pyrimidines that could potentially substitute for the altered *rut* site nucleotides (Fig. 3A).) The restoration of gene expression can be specifically attributed to decreased Rho action because stimulation of ATPase activity by RNAs unable to adopt stem-loops C or D was also reversed when combined with a *rut* site mutation (Fig. 6C).

If the stem-loop C+D conformation hinders Rho access to the *rut* site, then its disruption is predicted to favor Rho binding to the corresponding RNA substrate. In accordance with this notion, the affinity of Rho for a mutant *corA* leader mRNA unable to form stem-loop C (M1) was significantly higher than that for the wild-type *corA* leader RNA (Fig. 6D). By contrast, Rho displayed reduced affinity for an RNA harboring a mutation in the *rut* site (M9+M10, Fig. 6D). These experiments indicate that the stem-loop C+D conformation inhibits Rho binding to the *corA* leader independently of other effects that this conformation may have on Rho translocation and associated ATP hydrolysis. We therefore conclude that stem-loops C and D favor expression of the *corA* coding region by sequestering a *rut* site, and that stem-loop B hinders expression by increasing accessibility to the *rut* site.

Translation of the short ORF *corL* hinders formation of the stem-loop C+D conformation, favoring transcription termination within the *corA* leader region

The short ORF *corL* overlaps sequences that participate in mutually exclusive RNA secondary structure formation, indicating that *corL* translation may influence the *corA* leader conformation (Fig. 3). To determine whether *corL* is actually translated *in vivo*, we compared the β-galactosidase activity originating from wild-type *Salmonella* harboring either of three plasmids: pACYC-*corL*-'*lacZ*, with a constitutive promoter driving transcription of the *corA* leader through the last sense codon of *corL* fused in frame to the ninth codon of *lacZ*, a derivative of *pACYC-corL-'lacZ* mutated in the predicted *corL* SD sequence (M11 [G16C G17T]) and the empty vector, in which *lacZ* is present but lacks an SD sequence and start codon. The β-galactosidase activity from the pACYC-*corL*-'*lacZ*-carrying strain was 42-fold higher than that of the SD mutant and 5-fold higher than that of the vector control (Fig. 7A). These results demonstrate that *corL* is translated *in vivo* in *Salmonella*. This conclusion is supported by ribosome profiling experiments conducted in *E. coli*[24].

A ribosome occupying the *corL* ribosome binding site (RBS) or translating the *corL* open reading frame is expected to preclude stem-loop A formation, resulting in a kinetic competition between stem-loops B and C rather than stem-loops A and C. Because *corL* partially overlaps the left arm of stem-loop C, full translation of *corL* should prevent formation stem-loop C and favor stem-loop B (Fig. 3). Conversely, failure of a poised ribosome to initiate translation of *corL* should allow stem-loop C to dominate due to its thermodynamic stability and kinetic advantage over stem-loop B (Fig. 3). In accordance with this notion, replacement of the *corL* start codon with the UUA stop codon increased expression from a *corA-lacZ* transcriptional fusion (M12 [G29A], Fig. 7B). Mutation of the *corL* SD sequence de-repressed *corA-lacZ* transcription only 25% (M11, Fig. 7B) even though it prevented *corL* translation (M11, Fig. S4). These results confirm that ribosome occupancy of the *corL* ribosome binding site is necessary for preferential formation of stem-loop C, likely because stem-loop A can compete with stem-loop C in the absence of a ribosome at the RBS.

Because the *corL* start codon is a non-canonical UUG, we further explored the connection between *corL* translation initiation and *corA* transcription. Initiation of translation from UUG is significantly less efficient than from the standard AUG[29]. This is also the case with *corL* because the β-galactosidase activity from a strain harboring a *corL-lacZ* translational fusion was higher when the start codon was changed to AUG (Fig. S4). Conversely, alteration of the *corL* start codon from UUG to CUG dramatically reduced the efficiency of *corL* translation (Fig. S4). The AUG and CUG mutations slightly reduced and increased the β-galactosidase activity originating from strains with the *corA-lacZ* transcriptional fusion, respectively (Fig. 7D). Together, these results imply that *corA* transcription is inversely correlated with the efficiency of *corL* translation initiation.

The ribosome covers 12–15 nt of mRNA on both sides of its P site[30]. Therefore, a ribosome that stalls early on during translation of *corL* might not occlude the left arm of stem-loop C (Fig. 3B), leading to increased transcription of the *corA* coding region. To test this notion, we investigated the behavior of strains harboring derivatives of the *corA-lacZ* transcription

fusion in plasmid pYS1040 with stop codons at various positions within the *corL* ORF. Stop codons at positions at least 12 nucleotides from the base of stem-loop C (M15 and M16, Fig. 3) significantly increased expression (Fig. 7D). By contrast, a premature stop codon at a position from which the ribosome should still occlude the left arm of stem-loop C (M17, Fig. 3A) had no effect on *corA-lacZ* transcription (Fig. 7D). The phenotypes displayed by the strains with premature stop codons in *corL* are likely due to their effects on *corL* translation (as opposed to effects on the structure of the *corA* leader RNA) because the amber tRNA suppressor *supF* partially reversed the effect of a premature amber stop codon in *corL* (M16 (sup), Fig. S5A). As expected, the *supF*-expressing plasmid did not alter *corA-lacZ* transcription in strains with a wild-type *corL* sequence or with an ochre stop codon (Fig. S5A). A *corL*-specified peptide does not appear to regulate *corA-lacZ* transcription because expression of *corL* in *trans* failed to restore normal expression to a strain harboring a premature *corL* stop codon (Fig. S5B).

If *corL* translation exerts its effect by promoting formation of particular secondary structures in the *corA* leader RNA, then mutations forcing the RNA to adopt a certain conformation should be epistatic over those affecting *corL* translation. In accordance with this notion, a mutation that disrupts stem-loop C (M1, Fig 3A) is epistatic over mutations that reduce the efficiency of *corL* translation initiation or prevent elongation (M12 and M16, respectively, Fig. 3 and 7). In other words, preventing the ribosome from occluding stem-loop C has no effect on transcription of the associated coding region if the RNA cannot form stem-loop C. Taken together, the above results indicate that the efficiency of *corL* translation initiation and elongation influence the conformation adopted by the *corA* leader mRNA.

DISCUSSION

Expression of the *corA* gene was long thought to be constitutive due to its "housekeeping" function as the primary Mg^{2+} transporter in *Salmonella* and lack of observed Mg^{2+}-dependent regulation. However, this notion has recently been challenged by the observation that *corA* mRNA and protein levels do not correlate during the transition from exponential to stationary growth[31]. We have now established that *corA* is regulated at the level of transcription elongation by its 5' leader mRNA. Through selective secondary structure formation, the *corA* leader modulates the accessibility of a *rut* site to control the function of a Rho-dependent transcription terminator. Whether the RNA adopts one conformation or the other is dictated by translation of the short ORF *corL* located within the leader region. Under normal conditions, *corL* translation favors formation of stem-loop B due to occlusion of the left arm of stem-loop C by a translating ribosome (Fig. 3B). By contrast, a decrease in the efficiency of *corL* translation initiation or elongation promotes adoption of stem-loop C, in turn favoring transcription elongation into the *corA* coding region. In sum, there is an inverse correlation between *corL* translation and *corA* coding region transcription (Fig. 8). It is likely that *Salmonella* shares this regulatory mechanism with other enteric bacteria because the *corA* leader sequence is highly conserved within this family (Fig. 1A).

Surprisingly, RNAse III degrades *corA* mRNA in *E. coli* by specifically targeting stem-loop C[27]. Thus, adoption of the stem-loop C+D conformation is predicted to both increase transcription elongation into the coding region and decrease mRNA stability, resulting in

two opposing effects on overall *corA* mRNA levels. Conversely, the RNAse III cleavage site is not intact if the RNA adopts stem-loop B, so this conformation is predicted to be more stable even though it facilitates premature termination of transcription by Rho. The ultimate outcome of the interplay between these opposing forces likely depends on specific cellular conditions.

The *corA* and *mgtA* leaders utilize similar strategies to regulate *rut* accessibility

The mutually exclusive RNA conformations adopted by the *corA* leader (Fig. 3A) share a strong resemblance to those within the *mgtA* leader mRNA, which also modulates *rut* site accessibility via selective secondary structure formation[16], in spite of the fact that the corresponding nucleotide sequences are unrelated. In both the *corA* and *mgtA* leaders, the conformation that presents an accessible *rut* site contains an RNA hairpin (i.e., stem-loop B) that brings together two single-stranded pyrimidine-rich regions required for Rho loading. This function of stem-loop B is reminiscent of the ability of the *boxB* hairpin to clamp together the *rutA* and *rutB* sites of the λ tR1 Rho-dependent terminator[32]. In the *rut* site-accessible conformation, the *corA* and *mgtA* leaders also have the potential to form an additional three-helix structure (i.e., stem-loop A; Fig. 3A)[33], although in both cases, adoption of stem-loop A *in vivo* is likely precluded by translation of a short ORF that overlaps it (Fig. 3B)[20].

For both the *corA* and *mgtA* leader mRNAs, sequences within the right arm of stem-loop A and the left-arm of stem-loop B can alternatively base pair to form a structure (stem-loop C) that sequesters *rut* site nucleotides. In the case of *mgtA*, some of these nucleotides may participate in base pairing in the stem-loop C conformation[16], whereas the analogous nucleotides within the *corA* leader remain single-stranded but are structurally constrained (Fig. 3A).

In the *mgtA* leader mRNA, the upstream portion of the *rut* site (R1) plays a larger role than the downstream one (R2) in loading of Rho because R1 is required for the stem-loop B conformation to repress transcription elongation into the coding region whereas R2 is not[16]. Similarly, the upstream *rutA* is far more important than the downstream *rutB* for Rho-dependent termination at λ tR1[34]. By contrast, the downstream pyrimidine-rich tract in the *corA* leader is most important for Rho loading because mutation of this region (M9+M10) has larger effects than mutation of the upstream pyrimidine-rich tract (M8) on transcription *in vivo* (Fig. 4A), Rho-dependent termination *in vitro* (Fig. 4B) and Rho ATPase activity *in vitro* (Fig. 4C). The dominance of the downstream portion of the *rut* site in the *corA* leader may explain why the Rho-inaccessible conformation requires the additional stem-loop D (Fig. 3A).

Rho action establishes a regulatory link between Mg^{2+} homeostasis and translation

A canonical function of Rho is to implement transcriptional polarity. That is, when transcription and translation become uncoupled, cryptic *rut* sites normally occluded by translating ribosomes become accessible and Rho can terminate transcription[3]. Strikingly, the architectures of the *corA* (Fig. 3), *mgtA*[16] and *mgtB*[22] leaders reverse this relationship because Rho-dependent transcription termination within the leader is favored when

translation of short ORF(s) in the leader is efficient. Regulation of Mg^{2+} uptake in response to translational signals appears to be a universal mechanism because it was recently shown that expression of the eukaryotic Mg^{2+} channel TRPM7 is controlled by translation of two short ORFs within its 5' leader mRNA[35].

Why do enteric bacteria link Mg^{2+} homeostasis to Rho action? After all, it is possible to inversely couple translation of a short ORF within an mRNA leader to transcription of the associated coding region using an intrinsic (i.e., Rho-independent) transcription termination mechanism[36]. One possibility is that a Rho-dependent mechanism might facilitate sensing aspects of translation other than availability of particular amino acids, which is used to regulate amino acid biosynthetic operons[36]. In support of this notion, *corA* transcription can be modulated by altering the efficiency with which *corL* translation is initiated (Fig. 7B,C). Intriguingly, both *corL* and *mgtM* use the inefficient start codon UUG, which could be especially sensitive to global changes in translation, such as ribosome availability[37]. Because the majority of cytosolic Mg^{2+} is devoted to maintenance of ribosome structure and activity, an increase in Mg^{2+} uptake may help the cell cope with translational stressors[38]. Therefore, Rho's involvement in the control of Mg^{2+} transporter genes may promote Mg^{2+} uptake when mRNA translation is impaired.

MATERIALS AND METHODS

Bacterial strains and growth conditions

All experiments were carried out with wild-type *Salmonella enterica* serovar Typhimurium strain 14028s[39] or its derivative MK112, in which has the sequence corresponding to the *corA* leader is replaced by an 85 nt scar sequence. MK112 was constructed via the one-step gene disruption method[40] by transforming a PCR product generated from plasmid pKD4 using primers W332 and W333 into 14028s harboring pKD46 and selecting for KanR colonies. Following P22 transduction of the mutation into a wild-type background, the Kan[R] cassette was excised using plasmid pCP20, which left behind the 85 nt scar sequence from pKD4. *E. coli* DH5α[41] or XL-1 blue (Agilent) were used for cloning and plasmid DNA preparation. Bacteria were grown at 37°C in LB broth (Becton, Dickinson & Co), SOB (Becton, Dickinson & Co) or N-minimal medium (pH 7.4)[42] supplemented with 0.1% casamino acids, 38 mM glycerol and either zero or 10 mM $MgCl_2$. For plasmid maintenance, 20 μg/ml chloramphenicol (Cm), 50 μg/ml kanamycin (Kan) and/or 50 μg/ml ampicillin (Ap) were used.

Plasmid construction

Plasmid pYS1040 was constructed by cloning a PCR fragment generated from 14028s genomic DNA using primers 8446 and 8447 between the SalI and PstI sites of pYS1000[33]. The resulting plasmid harbors a *corA-lacZ* transcriptional fusion in which the *plac1–6* promoter drives transcription of *corA* leader followed by a 16 nt linker (containing PstI and XhoI sites), the *lacZ* ribosome binding site, the *lacZ* coding region and the *hisT* transcription terminator. As a result of the cloning process, pYS1040 and its derivatives contain two extra nucleotides between the plac1–6 promoter and the true TSS. A strain harboring a derivative of this plasmid with the correct TSS, designated pYS1040+1, produced similar β-

galactosidase activity compared to the strain carrying pYS1040 (Fig. S6A). All mutations tested in the pYS1040+1 background show analogous effects to those observed in pYS1040 (Fig. S6A). Nucleotide positions are numbered relative to the true TSS. Plasmid pYS1040FL, which contains the entire *corA* 5' intergenic region and lacks the plac1–6 promoter, was constructed by cloning a PCR fragment generated from 14028s genomic DNA using primers 11992 and 11987 between the SalI and XhoI sites of pYS1040.

Because some of the *in vitro*-mapped Rho-dependent termination sites are downstream of the *corA* start codon, we constructed a derivative of plasmid pYS1040 termed pYS1040ext that contains 49 nt of the *corA* coding region upstream of the pYS1040 *lacZ* ribosome binding site. Plasmid pYS1040ext was constructed by cloning a PCR fragment generated from 14028s genomic DNA using primers 8447 and W1431 between the SalI and PstI sites of pYS1040. Basal expression from pYS1040ext-carrying *Salmonella* is two-fold higher than the pYS1040-carrying strain (Fig. S6B). All mutations tested in the pYS1040ext background show analogous effects to those observed in pYS1040 (e.g. Fig. S6B).

Plasmid pMK100 was constructed by cloning a PCR product amplified from 14028s genomic DNA using primers W823 and W947 between the SpeI and SphI sites of pKH100. This places the *corA* leader and the first 49 nt of the *corA* coding region downstream of the λ_{pR} promoter and 26 nt C-less initial transcribed sequence and upstream of the *hisT* terminator[43].

Plasmid pACYC-*corL*-'*lacZ* was constructed by cloning a PCR product generated from 14028s genomic DNA with primers 11903 and 11905 between the XmaI and XbaI sites of plasmid pACYC-'*lacZ*. This generates a translational fusion in which the constitutive $p_{lac1–6}$ promoter drives transcription from the beginning of the *corA* leader up to the last sense codon of *corL* fused in frame to the ninth codon of the *E. coli lacZ* gene.

Plasmid p*corL* was constructed by cloning a PCR product generated from 14028s genomic DNA with primers 11903 and W574 into the HindIII site of pBAD18[44]. Correct orientation of the insert was confirmed by sequencing a PCR product generated from candidate plasmid DNA using primers 10127 and 10128.

Nucleotide substitutions in all plasmids were generated using the QuikChange II site-directed mutagenesis kit (Agilent) and primer pairs listed in Table S2.

Primer extension

10 pmol of primer 11984 was 5'-radiolabeled using [γ-^{32}P]-ATP (Perkin Elmer) and T4 polynucleotide kinase (NEB) according to the manufacturer's instructions for 30 min at 37°C. Labeled primer was purified on a G-50 column (GE Healthcare) and 0.4 pmol was mixed with 8 μg of total RNA extracted from *Salmonella* grown to mid-log phase in N-minimal media (10 mM $MgCl_2$). The mixture was incubated at 65°C for 5 min and cooled on ice for 1 min to promote annealing. Reverse transcription was carried out using Superscript II (Life Technologies) according to the manufacturer's instructions. Products were ethanol precipitated and mixed with 2x loading buffer II (Ambion) prior to being run on 8% polyacrylamide sequencing gels (Sequagel system, National Diagnostics). Four

sequencing reactions were performed, each in the presence of one ddNTP, using Vent (exo-) DNA polymerase (NEB) according to the manufacturer's instructions with plasmid pYS1040FL as the template.

Western blotting

1 ml N-minimal medium (10 mM $MgCl_2$) overnight cultures were washed once and diluted 1:50 in 25 ml of the same medium. Following 2 h at 37°C with 250 rpm shaking, 1 ml was removed for measuring cell density (OD_{600}) and 1 ml was washed with 1x TBS before freezing the pellet at −20°C. The remaining cultures were washed 2x and re-suspended in in cold N-minimal medium (no $MgCl_2$). Incubation was continued at 37°C with 250 rpm shaking with samples collected for OD_{600} measurement and Western blotting every 1–2 h. Western blot samples were prepared by re-suspending cell pellets in a volume of SDS-PAGE loading buffer (2.3% v/v sodium dodecyl sulfate, 22% v/v glycerol, 0.015 g/ml Tris, 0.05 mg/ml bromophenol blue, 87 µL/ml β-mercaptoethanol) determined by the formula: loading buffer (µl) = OD_{600} x sample volume (µl) x 100. The samples were heated to 95°C for 5 min, diluted to 0.1X in 15 µl SDS-PAGE loading buffer and heated again to 95°C for 3 min. Samples were loaded onto 12% Bis-Tris polyacrylamide gels (Life Technologies) and run in 1x MOPS running buffer at 200 V for approximately 40 min.

Samples were transferred to nitrocellulose membranes using the iBlot system (Invitrogen). Membranes were rinsed briefly in 1X TBS and blocked with 5% skim milk in 1X TBS for 1 h. Following a brief rinse in 1X TBS-T (TBS with 0.1% Tween20), primary antibody treatment was carried out for 1 h. Primary antibodies were prepared in 1X TBS-T with 0.01% sodium azide. Rabbit anti-CorA antibody (a generous gift from Michael Maguire) was used at a dilution of 1:10,000. Mouse anti-DnaK (Enzo Life Sciences) was used at a dilution of 1:5000 as a loading control. Following one 15 min wash and three 5 min washes in 1X TBS-T, secondary antibody treatment was carried out for 1 h using horseradish peroxidase-linked anti-Rabbit or anti-Mouse IgG (GE Healthcare). Membranes were again washed as described above prior to detection using the SuperSignal ELISA Femto kit (Thermo Scientific) and visualization with LAS-4000 (GE Healthcare). Quantification was performed with ImageQuant software (GE Healthcare).

In-line probing

T7 promoter-driven transcription templates corresponding to the *corA* leader were generated by PCR using plasmid pYS1040 or mutant derivatives and primers 8445 and 8444. RNA was synthesized from these templates using the Megascript *T7* kit (Ambion) according to the manufacturer's instructions. 1 µg of template DNA was used and reactions were incubated at 37°C overnight. RNAs were phenol/chloroform extracted, purified on 6% TBE urea gels (Life Technologies) and ethanol precipitated. 50 pmol of RNA was dephosphorylated by treating with calf intestinal alkaline phosphatase (NEB) according to the manufacturer's instructions for 30 min at 37°C. Dephosphorylated RNA was labeled with [γ-^{32}P]-ATP (Perkin Elmer) using T4 polynucleotide kinase (NEB) according to the manufacturer's instructions for 30 min at 37°C. Labeled RNA was purified on a 6% TBE urea gel and ethanol precipitated.

In-line probing reactions were set up as described[45]. Briefly, 1–2 μl labeled RNA was incubated in in-line buffer (50 mM Tris-HCl (pH 8.3), 200 mM KCl, 40 mM MgCl$_2$) for 40 h at room temperature. Control RNase T1 and alkaline hydrolysis reactions were carried out using the RNase T1 (1U/μl) kit (Ambion) according to the manufacturer's instructions. Denaturing 7% polyacrylamide sequencing gels (Sequagel system, National Diagnostics) were pre-run at 65 W for 1.5 h or until the gel reached 55°C, and then 8 μl of each sample was loaded and run for 1.5 or 3 h. After drying for 2 h at 80°C, the gel was exposed to a phosphoscreen overnight, visualized with a Typhoon FLA 9000 laser scanner (GE Healthcare) and bands were quantified with ImageQuant software (GE Healthcare).

β-galactosidase assays

0.8 ml aliquots of LB/Cm were inoculated with single wild-type *Salmonella* (14028s) colonies harboring plasmid pYS1040 or mutant derivatives and incubated overnight. Cells were washed once in LB and used to inoculate fresh 1 ml LB/Cm cultures in test tubes at a 1:100 dilution. Following incubation at 37°C for 4 h with 250 rpm shaking, 50 μl chloroform and 20 μl 0.1% SDS were added and cultures were vortexed vigorously for 1 min. After allowing the debris to settle for 15 min, β-galactosidase activity was measured as described[46]. Briefly, two 100 μl aliquots of each culture were transferred to a 96-well plate and 100 μl of 1.32 mg/ml ONPG was added to each well. Absorbance at 405 nm was recorded every 30 sec for 20 min at 30°C and the Vmax during a linear increase in absorbance was divided by cell density (absorbance at 562 nm) to yield a value in Miller units. Sterile media was used as a blank. Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

For experiments testing the effect of BCM (a generous gift from Max Gottesman), 1.5 ml cultures were used. After 4 h, two 600 μl aliquots were taken and water or 20 μg/ml BCM was added prior to incubation for an additional 30 min before treating the cells with chloroform/SDS and measuring the β-galactosidase activity.

For experiments using plasmids pBAD18[44] or *pcorL*, cells were grown in LB/Cm/Ap/0.4% L-arabinose to induce *corL* expression in *trans*. For experiments using plasmids *pUHE21-2lacI*q[20] or *pUH-supF*, cells were grown in LB/Cm/Ap/1mM IPTG.

In vitro transcription termination assays

Transcription templates were prepared by performing PCR with plasmid pMK100 or mutant derivatives as template and primers IA17 and IA256. Purified *Salmonella* RNA polymerase (RNAP) core enzyme and σ70 (courtesy of Kerry Hollands and Anastasia Sevostyanova) were mixed at a 1:3 molar ratio in RNAP storage buffer (50% v/v glycerol, 10 mM Tris pH 7.9, 0.1 M NaCl, 0.1 mM EDTA, 0.1 mM DTT) and allowed to associate by incubating at 30°C for 30 min.

Single round, synchronized *in vitro* transcription reactions were carried out as described[43]. 5 pmol DNA template was mixed with 5.6 pmol σ70+RNAP in a volume of 100 μl containing 100 mM Tris HCl (pH 8.0), 100 mM KCl, 50 mM MgCl$_2$, 5 μM each of ATP and UTP, 1 μM GTP, 100 μM ApU (TriLink) and 1.5 μl [α-^{32}P]-GTP (3000 Ci/mmol, Perkin Elmer).

This mixture was incubated at 37°C for 15 min to generate halted transcription elongation complexes (hTEC). hTEC were divided into aliquots and incubated at 37°C for 3 min. To initiate synchronized transcription elongation, a 5X chase +/− Rho was added (100 mM Tris HCl (pH 8.0), 100 mM KCl, 0.1mg/ml rifampicin, 1mM NTPs, with 600 nM Rho or an equal volume of Rho buffer). Reactions were incubated at 37°C for 5 min, terminated by addition of 2X stop solution (80% v/v formamide, 50 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue) and placed on ice. After heating at 95°C for 3 min, transcription products were resolved on denaturing 6% polyacrylamide sequencing gels (Sequagel system, National Diagnostics) and visualized using a Typhoon FLA 9000 laser scanner (GE Healthcare).

To map transcription termination sites, a portion of hTEC were used for RNA sequencing reactions as described[43]. Band intensities were measured using ImageQuant software (GE Healthcare) and normalized by subtracting the intensity of a blank region of the same area. Termination efficiency was calculated using the formula: $100 - 100*$ (runoff band intensity/ total lane intensity).

Rho ATPase assays

Rho ATPase assays were performed as described[16] using the EnzChek Phosphate Assay kit (Invitrogen). RNA was synthesized as described for in-line probing. 400 nM RNA in 0.75, 1.5 or 3 mM $MgCl_2$ (as indicated in the figure legends) was denatured by heating to 95°C and allowed to re-fold at room temperature for 1 h. The RNA was diluted in the same concentration of $MgCl_2$ to a final concentration of 0, 2, 4, 8, 15, 30, 50 or 100 nM in a 45 μl volume. 49 μl of master mix (50mM TrisHCl pH7.5, 0.1mM sodium azide, 0.4 mM MESG, 0.02 U/ml purine nucleoside phosphorylase, 0.75, 1.5 or 3 mM $MgCl_2$, 10 nM Rho protein) was added. The reactions were briefly centrifuged and left at room temperature for 10 m before adding 5 μl of 20 mM ATP. A SpectraMax Plus microplate reader (Molecular Diagnostics) was used to measure absorbance at 360 nM every 15 sec for 100 min at 37°C. Vmax was taken to be the slope during a 10 min window with the largest linear change in absorbance. A standard curve was generated from reactions using known concentrations of inorganic phosphate and the slope was used to convert Vmax to nmol ATP/min/μg Rho. Data shown correspond to mean values from two independent experiments and error bars indicate standard deviation.

The experiment shown in Fig. 6A was performed in 3 mM rather than 0.75 mM $MgCl_2$ because the M1 and M5 RNAs showed a decrease in ATPase activity at high RNA concentrations when the reactions were performed in 0.75 mM $MgCl_2$, suggesting that saturating activity was reached under this condition. Raising the $MgCl_2$ concentration to 3 mM reduced ATPase activity stimulated by all RNAs and returned the M1 and M5 mutants to an asymptotic curve, allowing for better resolution of differences between mutants. The experiment shown in Fig. 6C was performed in an intermediate $MgCl_2$ concentration of 1.5 mM to allow for resolution of mutants that both increase and decrease ATPase activity.

Filter binding assays

Rho-RNA binding was measured using the nitrocellulose binding method[47]. RNAs synthesized and 5' radiolabeled as described for in-line probing were diluted to 0.1 nM in binding buffer (50 mM KCl, 1 mM $MgCl_2$, 20 mM HEPES pH 7.9, 0.1 mM EDTA), heated at 80°C for 3 min, and allowed to re-nature at room temperature for 10 min. Yeast RNA was added to 0.05 mg/ml to prevent non-specific binding. 10 nM to 10 μM dilutions of Rho were prepared in Rho dilution buffer (10 mM Tris-HCL pH 7.9, 100mM KCL, 0.1 mM DTT, 0.1 mM EDTA, 50% v/v glycerol). 100 μl RNA aliquots (10 fmol) were mixed with 10 μl of each Rho dilution, or Rho dilution buffer, and incubated at room temperature for 10 min. Under vacuum, 0.45 μM nitrocellulose filters (Millipore) were washed with 5 ml binding buffer. Following application of the Rho-RNA samples, the filters were immediately washed with 5 ml binding buffer, air dried, and quantified using a Typhoon FLA 9000 laser scanner and ImageQuant software (GE Healthcare). To calculate the percentage of RNA bound, spot intensities were normalized to samples lacking Rho protein and then divided by the intensity of the highest Rho concentration used, which was a 10,000-fold excess and thus assumed to be saturating.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

ABBREVIATIONS

hTEC	halted transcription elongation complex
nt	nucleotide
ORF	open reading frame
PBS	primary binding site
RNAP	RNA polymerase
RT	reverse transcription
rut	Rho utilization
RBS	ribosome binding site
SD	Shine-Dalgarno
TSS	transcription start site

REFERENCES

1. Peters JM, Vangeloff AD, Landick R. Bacterial transcription terminators: the RNA 3'-end chronicles. J Mol Biol. 2011; 412:793–813. [PubMed: 21439297]

2. Boudvillain M, Figueroa-Bossi N, Bossi L. Terminator still moving forward: expanding roles for Rho factor. Curr Opin Microbiol. 2013; 16:118–124. [PubMed: 23347833]

3. Richardson JP, Grimley C, Lowery C. Transcription termination factor rho activity is altered in Escherichia coli with suA gene mutations. Proc Natl Acad Sci U S A. 1975; 72:1725–1728. [PubMed: 1098042]

4. Adhya S, Gottesman M. Control of transcription termination. Annu Rev Biochem. 1978; 47:967–996. [PubMed: 354508]

5. Modrak D, Richardson JP. The RNA-binding domain of transcription termination factor rho: isolation, characterization, and determination of sequence limits. Biochemistry. 1994; 33:8292–8299. [PubMed: 7518246]

6. Banerjee S, Chalissery J, Bandey I, Sen R. Rho-dependent transcription termination: more questions than answers. J Microbiol. 2006; 44:11–22. [PubMed: 16554712]

7. Bogden CE, Fass D, Bergman N, Nichols MD, Berger JM. The structural basis for terminator recognition by the Rho transcription termination factor. Mol Cell. 1999; 3:487–493. [PubMed: 10230401]

8. Skordalakes E, Berger JM. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. Cell. 2003; 114:135–146. [PubMed: 12859904]

9. Graham JE. Sequence-specific Rho-RNA interactions in transcription termination. Nucleic Acids Res. 2004; 32:3093–3100. [PubMed: 15181174]

10. Morgan WD, Bear DG, von Hippel PH. Rho-dependent termination of transcription. II. Kinetics of mRNA elongation during transcription from the bacteriophage lambda PR promoter. J Biol Chem. 1983; 258:9565–9574. [PubMed: 6223930]

11. Ciampi MS. Rho-dependent terminators and transcription termination. Microbiology. 2006; 152:2515–2528. [PubMed: 16946247]

12. Wei RR, Richardson JP. 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. 2001; 276:28380–28387. [PubMed: 11369775]

13. Kim DE, Patel SS. The kinetic pathway of RNA binding to the Escherichia coli transcription termination factor Rho. J Biol Chem. 2001; 276:13902–13910. [PubMed: 11278821]

14. Rabhi M, Gocheva V, Jacquinot F, Lee A, Margeat E, Boudvillain M. Mutagenesis-based evidence for an asymmetric configuration of the ring-shaped transcription termination factor Rho. J Mol Biol. 2011; 405:497–518. [PubMed: 21059356]

15. Epshtein V, Dutta D, Wade J, Nudler E. An allosteric mechanism of Rho-dependent transcription termination. Nature. 2010; 463:245–249. [PubMed: 20075920]

16. Hollands K, Proshkin S, Sklyarova S, Epshtein V, Mironov A, Nudler E, Groisman EA. Riboswitch control of Rho-dependent transcription termination. Proc Natl Acad Sci U S A. 2012; 109:5376–5381. [PubMed: 22431636]

17. Konan KV, Yanofsky C. Rho-dependent transcription termination in the tna operon of Escherichia coli: roles of the boxA sequence and the rut site. J Bacteriol. 2000; 182:3981–3988. [PubMed: 10869076]

18. Bossi L, Schwartz A, Guillemardet B, Boudvillain M, Figueroa-Bossi N. A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. Genes Dev. 2012; 26:1864–1873. [PubMed: 22895254]

19. Figueroa-Bossi N, Schwartz A, Guillemardet B, D'Heygère F, Bossi L, Boudvillain M. RNA remodeling by bacterial global regulator CsrA promotes Rho-dependent transcription termination. Genes Dev. 2014; 28:1239–1251. [PubMed: 24888591]

20. Park SY, Cromie MJ, Lee EJ, Groisman EA. A bacterial mRNA leader that employs different mechanisms to sense disparate intracellular signals. Cell. 2010; 142:737–748. [PubMed: 20813261]

21. Lee EJ, Groisman EA. Tandem attenuators control expression of the Salmonella mgtCBR virulence operon. Mol Microbiol. 2012; 86:212–224. [PubMed: 22857388]

22. Sevostyanova A, Groisman E. An RNA motif advances transcription by preventing Rho-dependent termination. Proc Natl Acad Sci U S A. In press.

23. Kröger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hébrard M, Händler K, Colgan A, Leekitcharoenphon P, Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ, Thomson NR, Vogel J, Hinton JC. The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. Proc Natl Acad Sci U S A. 2012; 109:E1277–E1286. [PubMed: 22538806]

24. Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, Nichols RJ, Typas A, Gross CA, Kramer G, Weissman JS, Bukau B. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell. 2011; 147:1295–1308. [PubMed: 22153074]

25. Das A, Court D, Adhya S. Isolation and characterization of conditional lethal mutants of Escherichia coli defective in transcription termination factor rho. Proc Natl Acad Sci U S A. 1976; 73:1959–1963. [PubMed: 132662]

26. Peters JM, Mooney RA, Kuan PF, Rowland JL, Keles S, Landick R. Rho directs widespread termination of intragenic and stable RNA transcription. Proc Natl Acad Sci U S A. 2009; 106:15406–15411. [PubMed: 19706412]

27. Lim B, Sim SH, Sim M, Kim K, Jeon CO, Lee Y, Ha NC, Lee K. RNase III controls the degradation of corA mRNA in Escherichia coli. J Bacteriol. 2012; 194:2214–2220. [PubMed: 22343302]

28. Mandal M, Breaker RR. Gene regulation by riboswitches. Nat Rev Mol Cell Biol. 2004; 5:451–463. [PubMed: 15173824]

29. Reddy P, Peterkofsky A, McKenney K. Translational efficiency of the Escherichia coli adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. Proc Natl Acad Sci U S A. 1985; 82:5656–5660. [PubMed: 3898067]

30. Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev. 2005; 69:101–123. [PubMed: 15755955]

31. Papp-Wallace KM, Maguire ME. Regulation of CorA Mg2+ channel function affects the virulence of Salmonella enterica serovar typhimurium. J Bacteriol. 2008; 190:6509–6516. [PubMed: 18676666]

32. Vieu E, Rahmouni AR. Dual role of boxB RNA motif in the mechanisms of termination/antitermination at the lambda tR1 terminator revealed in vivo. J Mol Biol. 2004; 339:1077–1087. [PubMed: 15178249]

33. Cromie MJ, Shi Y, Latifi T, Groisman EA. An RNA sensor for intracellular Mg(2+). Cell. 2006; 125:71–84. [PubMed: 16615891]

34. Graham JE, Richardson JP. rut Sites in the nascent transcript mediate Rho-dependent transcription termination in vivo. J Biol Chem. 1998; 273:20764–20769. [PubMed: 9694820]

35. Nikonorova IA, Kornakov NV, Dmitriev SE, Vassilenko KS, Ryazanov AG. Identification of a Mg2+-sensitive ORF in the 5'-leader of TRPM7 magnesium channel mRNA. Nucleic Acids Res. 2014; 42:12779–12788. [PubMed: 25326319]

36. Yanofsky C. Attenuation in the control of expression of bacterial operons. Nature. 1981; 289:751–758. [PubMed: 7007895]

37. Khudyakov, YuE; Neplyueva, VS.; Kalinina, TI.; Smirnov, VD. Effect of structure of the initiator codon on translation in E. coli. FEBS Lett. 1988; 232:369–371. [PubMed: 2967775]

38. Pontes MH, Sevostyanova A, Groisman EA. When too much ATP is bad for protein synthesis. J Mol Biol. 2015

39. Fields PI, Swanson RV, Haidaris CG, Heffron F. Mutants of Salmonella typhimurium that cannot survive within the macrophage are avirulent. Proc Natl Acad Sci U S A. 1986; 83:5189–5193. [PubMed: 3523484]

40. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A. 2000; 97:6640–6645. [PubMed: 10829079]

41. Hanahan D. Studies on transformation of Escherichia coli with plasmids. J Mol Biol. 1983; 166:557–580. [PubMed: 6345791]

42. Snavely MD, Gravina SA, Cheung TT, Miller CG, Maguire ME. Magnesium transport in Salmonella typhimurium. Regulation of mgtA and mgtB expression. J Biol Chem. 1991; 266:824–829. [PubMed: 1898738]

43. Artsimovitch I, Henkin TM. In vitro approaches to analysis of transcription termination. Methods. 2009; 47:37–43. [PubMed: 18948199]

44. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol. 1995; 177:4121–4130. [PubMed: 7608087]

45. Regulski EE, Breaker RR. In-line probing analysis of riboswitches. Methods Mol Biol. 2008; 419:53–67. [PubMed: 18369975]

46. Miller, JH. Experiments in Molecular Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1972. Assay of B-galactosidase.

47. Carey J, Cameron V, de Haseth PL, Uhlenbeck OC. Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. Biochemistry. 1983; 22:2601–2610. [PubMed: 6347247]

Highlights

- The widespread Mg^{2+} channel CorA was thought to be constitutively expressed

- *Salmonella corA* leader mRNA harbors a Rho-dependent transcription terminator

- The *corA* leader can adopt two conformations that control Rho access to the mRNA

- Translation of a short ORF in the *corA* leader dictates the mRNA conformation

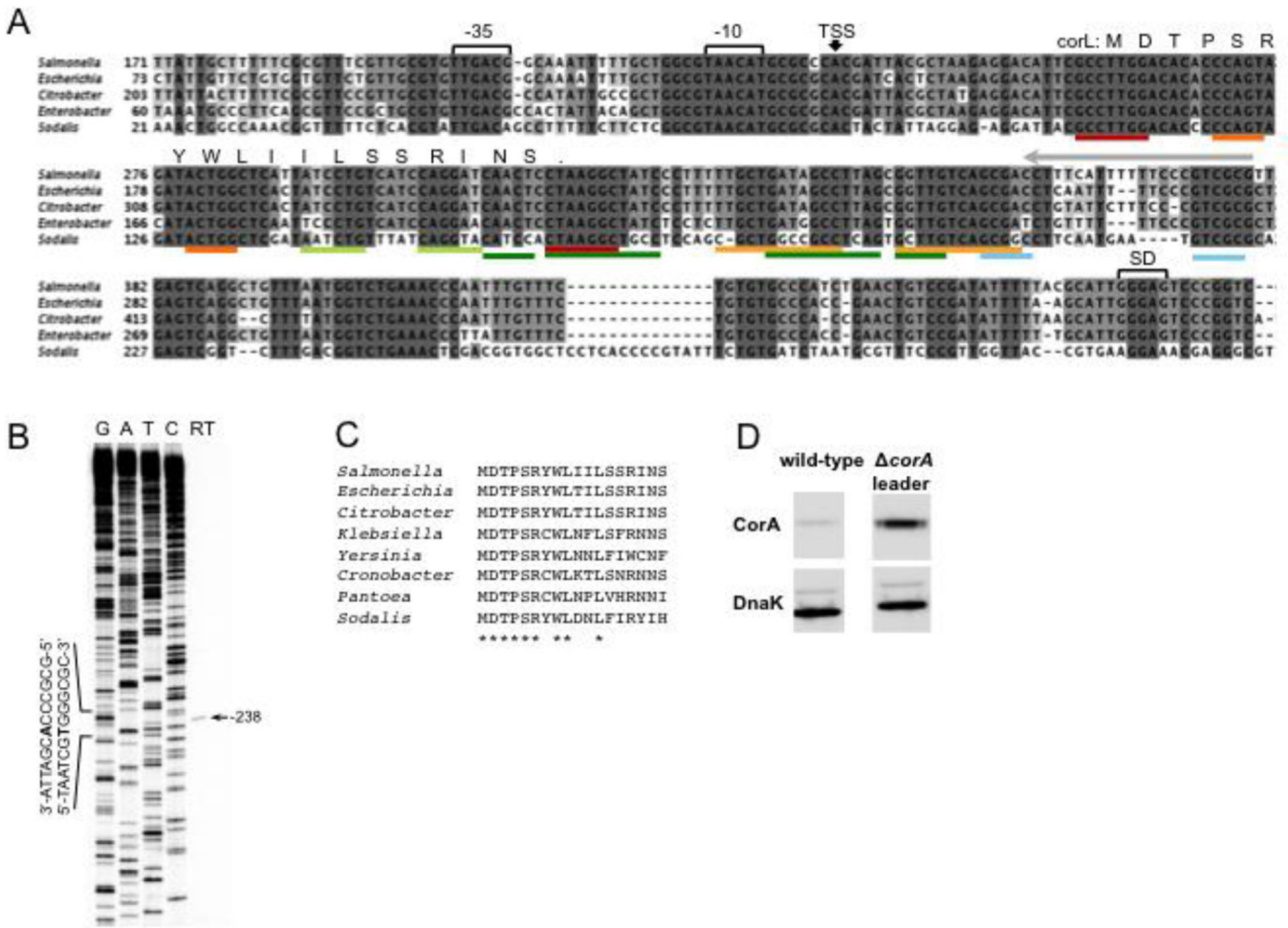- Rho links expression of all *Salmonella* Mg^{2+} transporters to translational signals

Fig. 1. The *corA* transcript includes a long, highly conserved 5' leader region that functions as a repressor

(A) Nucleotide alignment of the *corA* 5' intergenic region from selected enteric bacterial genomes. The location of promoter elements, RNA secondary structures (colored bars, see Fig. 3A), the short ORF *corL* and the *corA* Shine-Dalgarno (SD) sequence are shown. The transcription start site (TSS) mapped in (B) is denoted with a black arrow.

(B) Primer extension of total RNA isolated from wild-type *Salmonella* (14028s) grown to mid-log phase in N-minimal medium (10 mM MgCl$_2$) using a 5' radiolabeled oligonucleotide that anneals to nucleotides –93 to –114 relative to the *corA* start codon (Fig. 1A, grey arrow). The band obtained from reverse transcription (RT) was mapped to position –238 relative to the *corA* start codon using a DNA sequencing ladder generated from plasmid pYS1040FL (Materials and Methods).

(C) Alignment of the deduced amino acid sequence of *corL* from selected enteric bacterial genomes. Asterisk indicates positions conserved in all eight listed species.

(D) Western blot of cell lysates prepared from wild-type (14028s) *Salmonella* or a mutant strain in which the *corA* leader was replaced by a scar sequence from plasmid pKD4 (MK112). Bacteria were grown in N-minimal medium (10mM MgCl$_2$) for 2 h prior to switching to N-minimal media without MgCl$_2$ for 2 h.
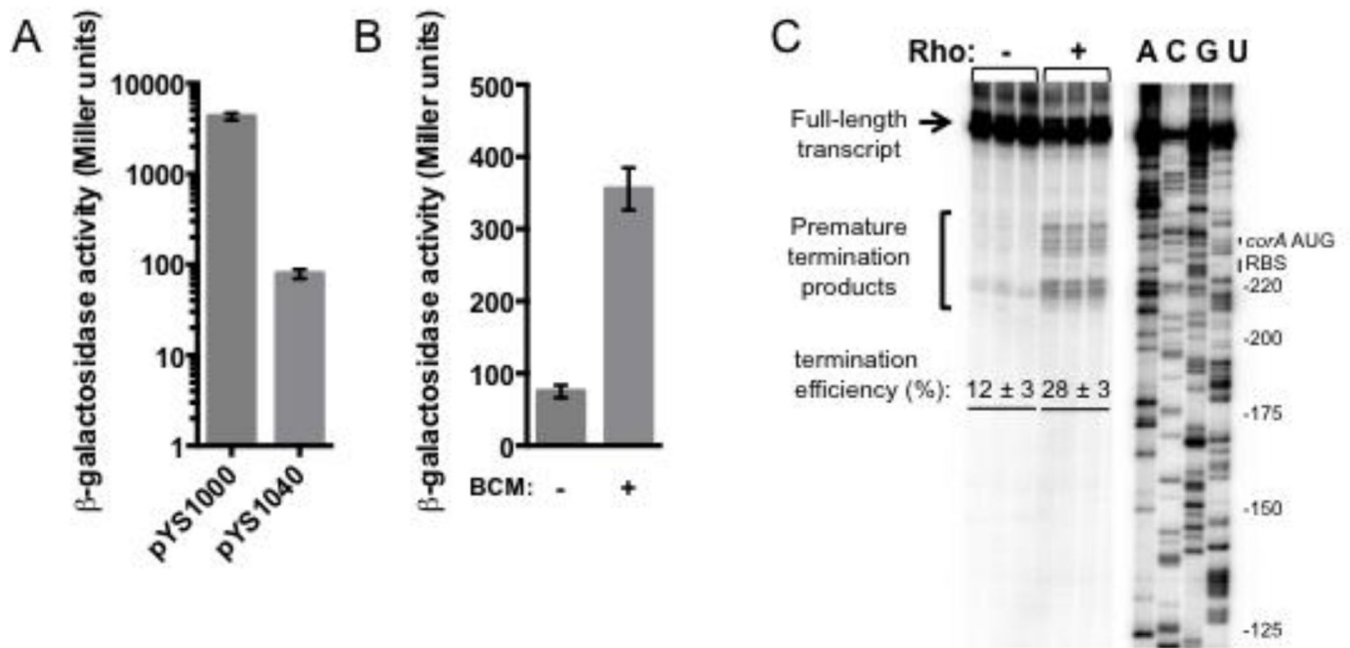
Fig. 2. The *corA* leader mRNA contains a Rho-dependent transcription terminator

(A) β-galactosidase activity (Miller units) from wild-type *Salmonella* (14028s) harboring plasmid pYS1000, in which the $p_{lac1\text{-}6}$ promoter drives expression of *lacZ*, or the derivative pYS1040, in which the *Salmonella corA* leader sequence is present between $p_{lac1\text{-}6}$ and the *lacZ* SD sequence. Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(B) β-galactosidase activity (Miller units) from wild-type *Salmonella* (14028s) harboring plasmid pYS1040 in the presence or absence of BCM. Bacteria were grown in LB/Cm medium for 4 h, followed by incubation in the presence of BCM or mock for 30 min. Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(C) A representative 6% denaturing gel of single-round, synchronized *in vitro* transcription reactions performed using a DNA template generated from plasmid pMK100. The assay shown was performed in 5 mM $MgCl_2$. Reactions were performed in triplicate, and only the relevant portion of the gel is shown. Calculation of termination efficiency and generation of an RNA sequencing ladder were performed as described in Materials and Methods.
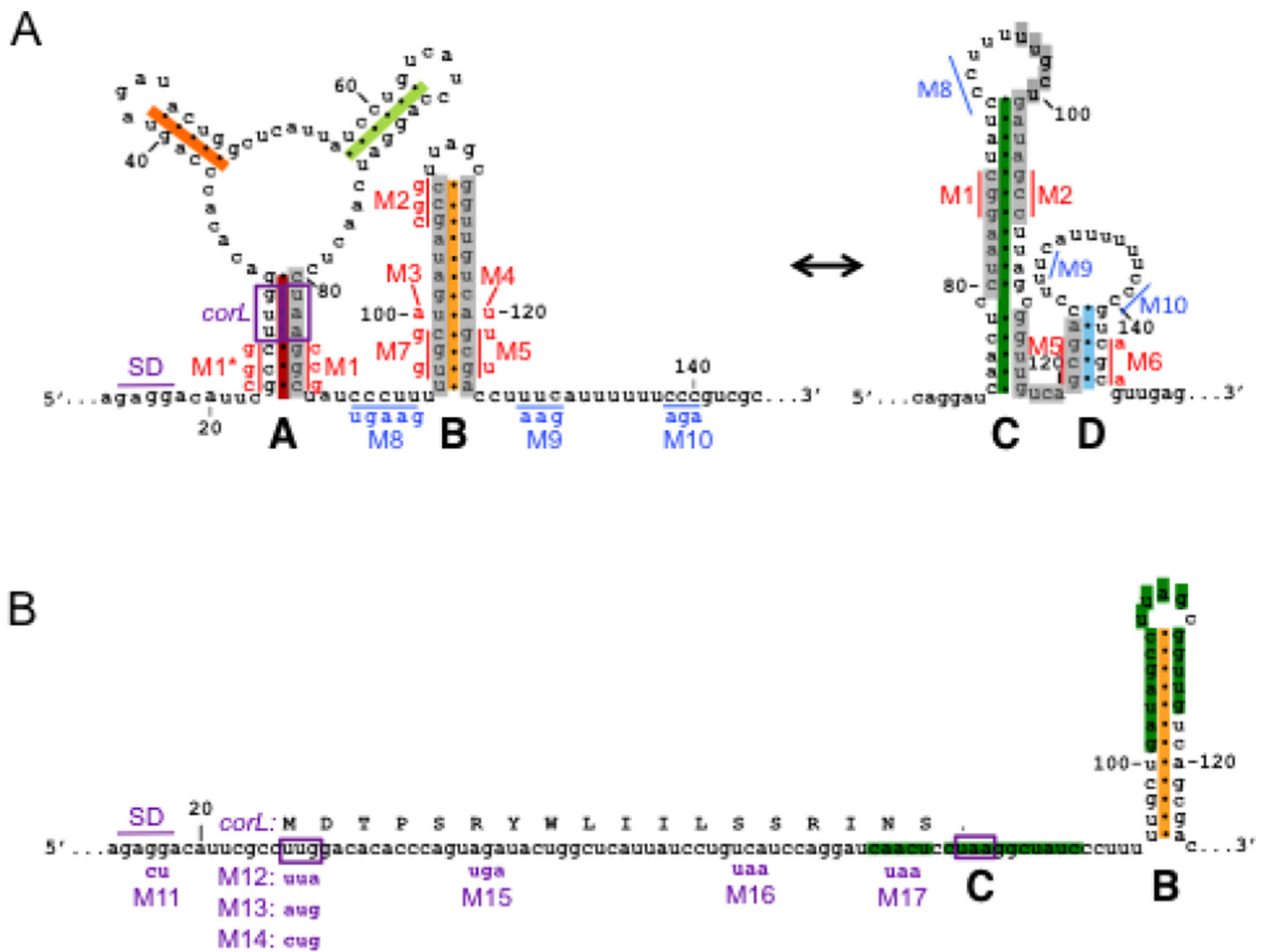
Fig. 3. The *corA* leader can adopt two mutually exclusive conformations that differentially regulate expression of the associated coding region

(A) Schematic of predicted secondary structures within the *corA* leader mRNA. The two possible conformations, termed stem-loops A+B and stem-loops C+D, are mutually exclusive (grey bars). Positions are numbered relative to the *corA* TSS. Mutations used to disrupt base pairing (red) or Rho loading (blue) are denoted next to the mutated nucleotides. Structure prediction and mutation design were aided by the M-fold Web Server (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form). (B) Schematic denoting the position of the short ORF *corL* relative to the RNA secondary structures. The *corL* SD sequence, start codon and stop codon are shown in purple. Translation of *corL in vivo* is expected to favor stem-loop B by preventing formation of stem-loops A and C. Mutations used to alter *corL* translation initiation or elongation are denoted below the nucleotide sequence.
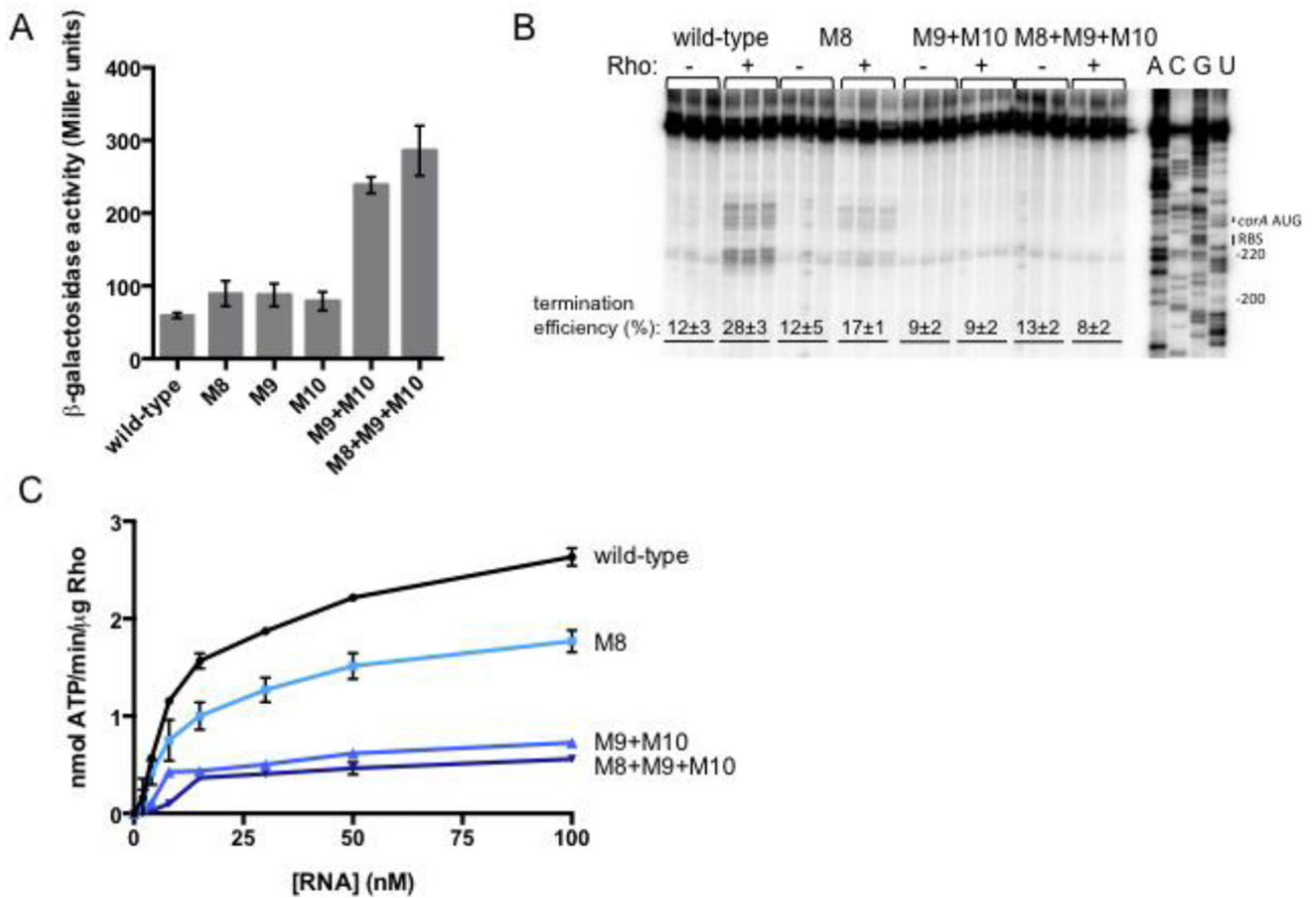
Fig. 4. Two pyrimidine-rich regions flanking stem-loop B constitute the *rut* site

(A) β-galactosidase activity exhibited by wild-type *Salmonella* (14028s) harboring plasmid pYS1040 or derivatives containing mutations that change the indicated pyrimidines to purines (see Fig. 3A). Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(B) A representative 6% denaturing gel of single-round, synchronized *in vitro* transcription reactions performed using templates generated from plasmid pMK100 or derivatives containing mutations in the pyrimidine-rich regions flanking stem-loop B (Fig. 3A). The assay shown was performed in 5 mM MgCl$_2$. Reactions were performed in triplicate, and only the relevant portion of the gel is shown. Calculation of termination efficiency and generation of an RNA sequencing ladder were performed as described in Materials and Methods.

(C) The accumulation of free phosphate was monitored to quantify Rho's ATPase activity. Purified Rho protein was incubated with RNA corresponding to the wild-type *corA* leader mRNA or variants harboring mutations in the pyrimidine-rich regions. The experiment shown was performed in 0.75 mM MgCl$_2$. Data shown correspond to mean values from two independent experiments, and error bars indicate standard deviation.
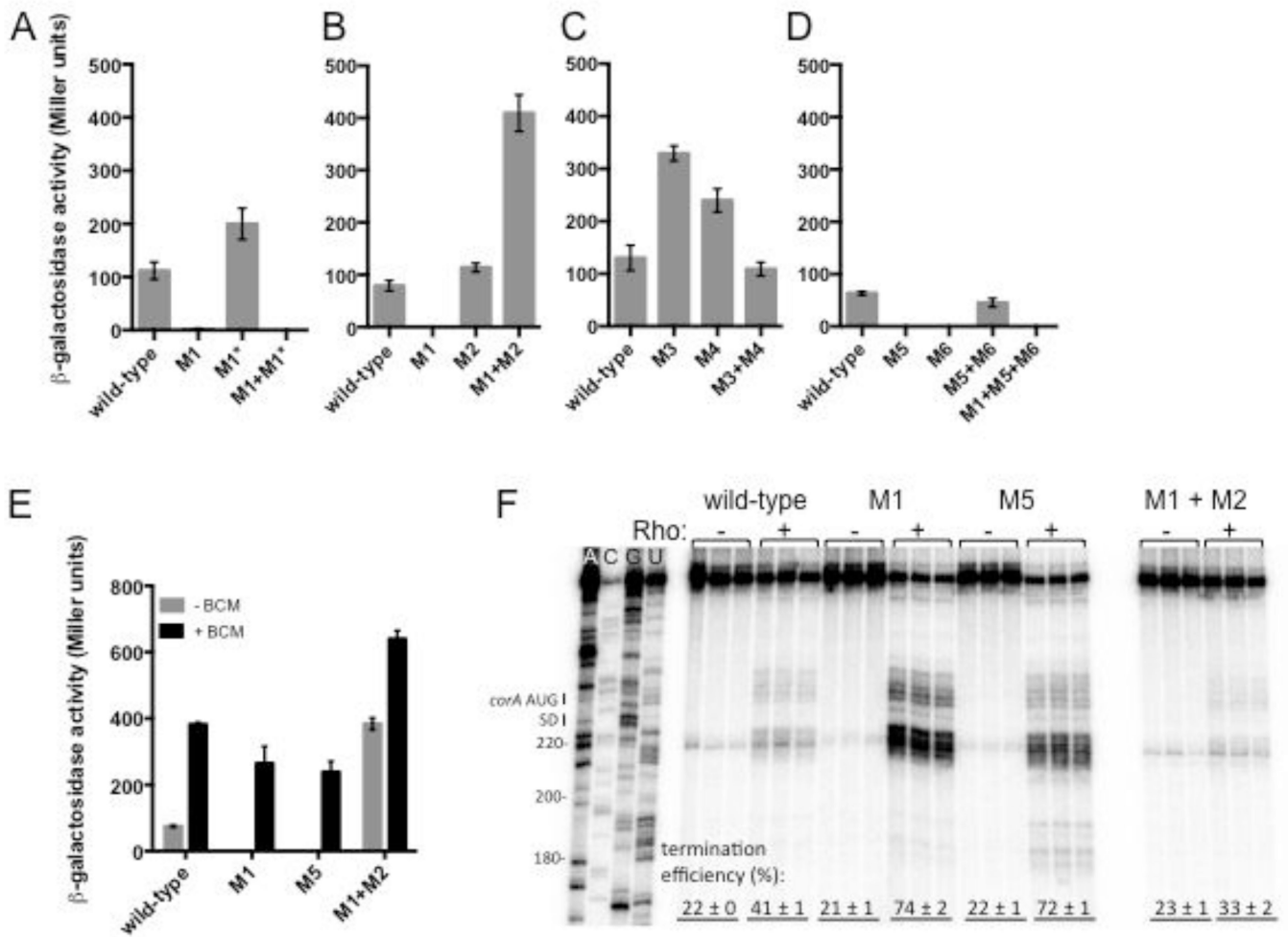
Fig. 5. Stem-loops C and D promote transcription of the associated coding region by hindering Rho-dependent termination

(A–D) β-galactosidase activity (Miller units) exhibited by wild-type *Salmonella* (14028s) harboring plasmid pYS1040 or derivatives containing mutations that disrupt base pairing in stem-loops A, B, C and/or D. Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(E) β-galactosidase activity (Miller units) exhibited by wild-type *Salmonella* (14028s) harboring plasmid pYS1040 or derivatives harboring mutations in stem-loops A, C and/or D, in the presence or absence of the Rho-specific inhibitor bicyclomycin (BCM). Following growth in LB/Cm for 4 h, BCM (to a final concentration of 20 μg/ml) or an equal volume of water was added and bacteria were incubated for an additional 30 min prior to measurement of β-galactosidase activity. Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(F) A representative 6% denaturing gel of single-round, synchronized *in vitro* transcription reactions performed using templates generated from plasmid pMK100 or derivatives containing mutations in stem-loops A, C and/or D. The assay shown was performed in 1 mM MgCl$_2$. Reactions were performed in triplicate, and only the relevant portion of the gel

is shown. Calculation of termination efficiency and generation of an RNA sequencing ladder were performed as described in Materials and Methods.
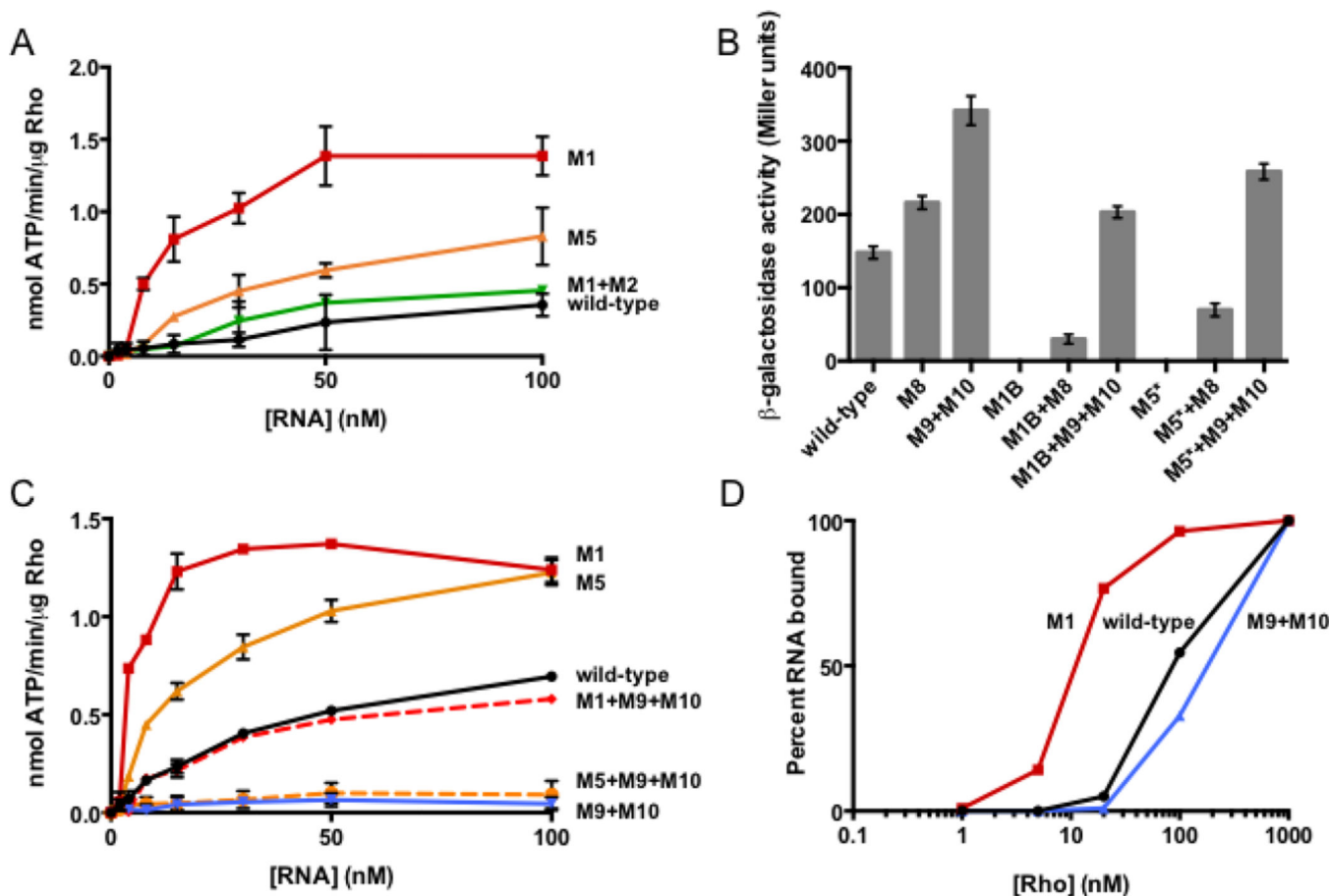
Fig. 6. Stem-loops C and D hinder termination by decreasing accessibility of the *rut* site to Rho
(A) The accumulation of free phosphate was monitored to quantify Rho's ATPase activity. Purified Rho protein was incubated with RNA corresponding to wild-type *corA* leader mRNA or variants harboring mutations in stem-loops A-C. The experiment shown was performed in 3 mM MgCl$_2$. Data shown correspond to mean values from two independent experiments, and error bars indicate standard deviation

(B) β-galactosidase activity exhibited by wild-type *Salmonella* (14028s) harboring plasmid pYS1040 or derivatives containing mutations that disrupt stem-loop C or D, reduce the pyrimidine content of the downstream pyrimidine-rich tract, or both. Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(C) The accumulation of free phosphate was monitored to quantify Rho's ATPase activity. Purified Rho protein was incubated with RNA corresponding to wild-type *corA* leader mRNA or variants harboring mutations that alter the RNA secondary structure, the downstream pyrimidine-rich tract, or both. The experiment shown was performed in 1.5 mM MgCl$_2$. Data shown correspond to mean values from two independent experiments, and error bars indicate standard deviation

(D) Percentage of 5' radiolabeled wild-type *corA* leader RNA (black) or mutant derivatives (M1, red and M9+M10, blue) bound by Rho and retained following vacuum filtration. Values were calculated assuming saturation at 1 mM Rho, as described in Materials and

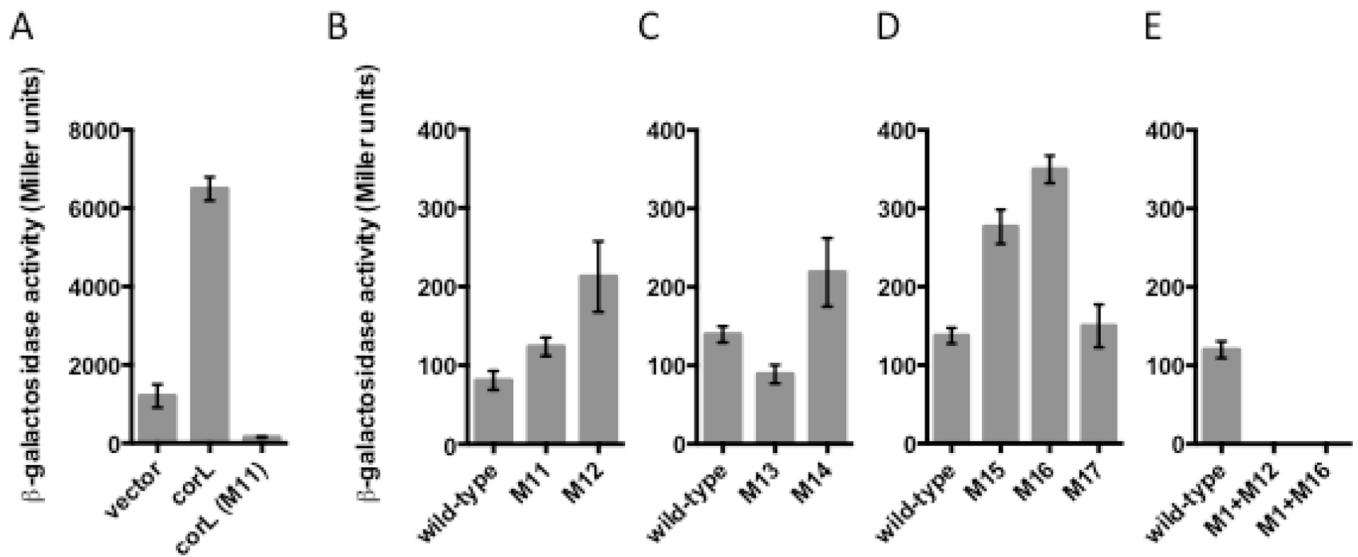Methods. Data shown are from a single experiment representative of three independent experiments.

Fig. 7. Rho-dependent termination within the *corA* leader is inversely coupled to translation of the short ORF *corL*

(A) β-galactosidase activity (Miller units) exhibited by wild-type *Salmonella* (14028s) harboring plasmid pACYC-'*ZacZ* (vector), which lacks translation initiation signals, pACYC-*corL*-'*lacZ* (*corL*), a translational fusion of a *corA* leader fragment (from the TSS to the last sense codon of *corL*) to the ninth codon of the *E. coli lacZ* gene, or a derivative containing a mutation in the *corL* SD sequence (*corL*(M11)). Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.

(B–E) β-galactosidase activity (Miller units) exhibited by wild-type *Salmonella* (14028s) harboring plasmid pYS1040 or derivatives containing mutations that alter the efficiency of *corL* translation initiation (M11-M14) or introduce a premature stop codon into *corL* (M15-M17). Data shown correspond to mean values from at least three independent experiments performed in duplicate, and error bars indicate standard deviation.
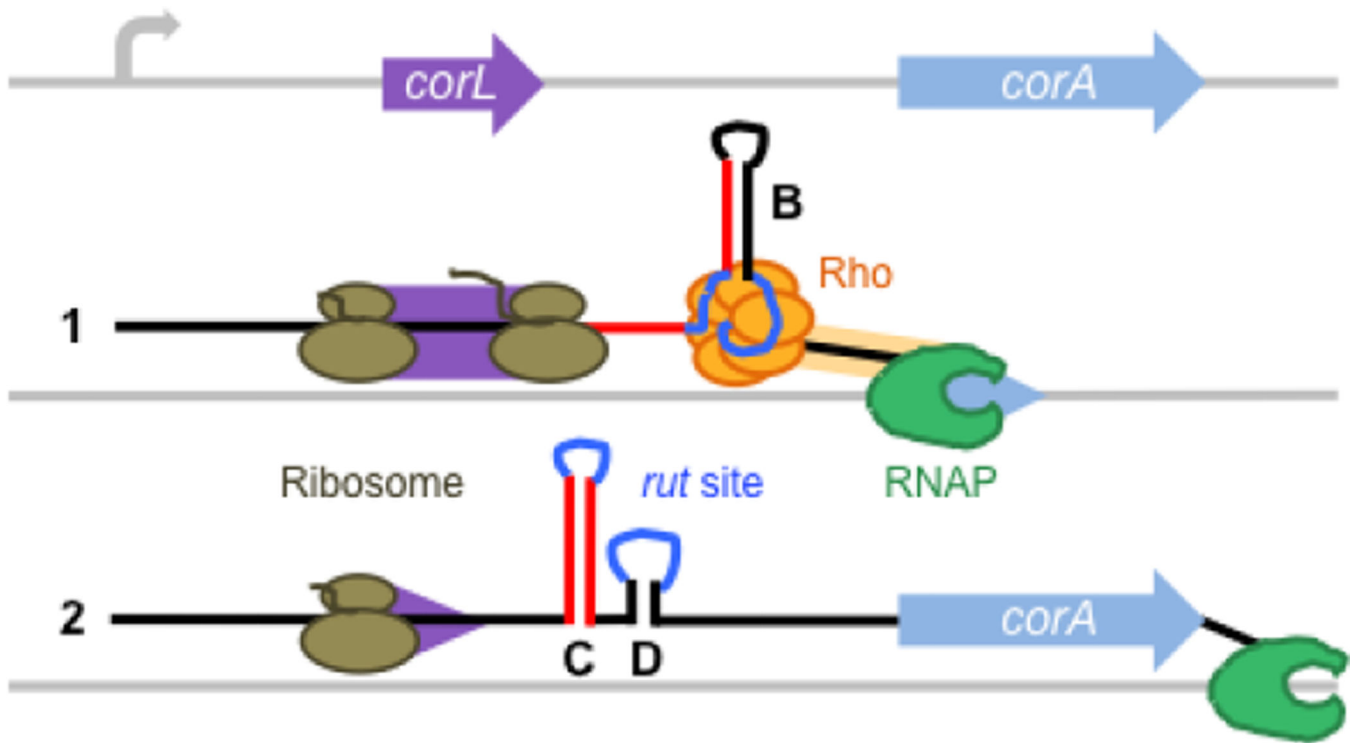
Fig. 8. Model for control of Rho-dependent transcription termination in the *corA* leader by translation of the short ORF *corL*

When *corL* (purple) is translated efficiently (1), the ribosome (brown) occludes the left arm of stem-loop C, thus promoting stem-loop B formation. This RNA conformation presents an accessible *rut* site (dark blue), allowing Rho (orange) to load onto the *corA* mRNA leader and trigger RNA release by RNAP (green) prior to transcription of the *corA* coding region (light blue). When translation of *corL* is impaired (2), either by inefficient translation initiation or elongation, the kinetic and thermodynamic advantages of stem-loop C over stem-loop B lead to formation of the stem-loops C+D conformation, which sequesters the *rut* site and allows RNAP to continue into the *corA* coding region.