Processing generates 3′ ends of RNA masking transcription termination events in prokaryotes

Xun Wang[a], Monford Paul Abishek N[b], Heung Jin Jeon[b], Yonho Lee[b], Jin He[a], Sankar Adhya[c,1], and Heon M. Lim[b,1]

[a]State Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, Wuhan, Hubei 430070, People's Republic of China; [b]Department of Biological Sciences, College of Biological Sciences and Biotechnology, Chungnam National University, 305-764 Daejeon, Republic of Korea; and [c]Laboratory of Molecular Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

Contributed by Sankar Adhya, December 10, 2018 (sent for review July 31, 2018; reviewed by Max E. Gottesman and Jeffrey W. Roberts)

Two kinds of signal-dependent transcription termination and RNA release mechanisms have been established in prokaryotes in vitro by: (*i*) binding of Rho to cytidine-rich nascent RNA [Rho-dependent termination (RDT)], and (*ii*) the formation of a hairpin structure in the nascent RNA, ending predominantly with uridine residues [Rho-independent termination (RIT)]. As shown here, the two signals act independently of each other and can be regulated (suppressed) by translation–transcription coupling in vivo. When not suppressed, both RIT- and RDT-mediated transcription termination do occur, but ribonucleolytic processing generates defined new 3′ ends in the terminated RNA molecules. The actual termination events at the end of transcription units are masked by generation of new processed 3′ RNA ends; thus the in vivo 3′ ends do not define termination sites. We predict generation of 3′ ends of mRNA by processing is a common phenomenon in prokaryotes as is the case in eukaryotes.

transcription termination | RNA processing | gal operon | translation–transcription coupling

Two types of transcription termination mechanisms have been documented in prokaryotic organisms in vitro: (*i*) Rho-dependent termination (RDT) facilitated by binding of Rho protein to a cytidine-rich (C-rich) segment in the nascent RNA followed by dissociation of the RNA; and (*ii*) intrinsic or Rho-independent termination (RIT) facilitated by formation of an RNA hairpin structure with five to seven uridine residues at the end that causes RNA release (1–7). We investigated the two types of termination events both in vivo and in vitro using the *gal* operon of *Escherichia coli* as a model system.

The *gal* operon is tetracistronic and is about 4 kb long. DNA sequence shows that the end of *gal* contains both RIT and RDT signals in that order (Fig. 1). Northern analysis showed that the full-length *gal* mRNA in vivo is about 4.3 kb long (termed mM1) and its 3′ end appears at the RIT (8). First, the RIT signal in *gal* has only three uridine residues and thus not expected to efficiently release RNA during termination. Second, the stop codon of the last ORF (*galM*) being only 6 bp away from the beginning of the RIT signal is too close to allow the required RNA hairpin structure formation because of translation–transcription coupling (see below). Thus, it is not clear whether the mM1 RNA really originates at RIT. There is also no evidence whether the RDT signal is playing any role in *gal*. In fact, the actual functioning of both RIT and RDT signals in vivo has seldom been investigated in any system. We addressed this issue by investigating what role the two types of signals play in generating the 3′ end of the mM1 mRNA in vivo. Since translation–transcription coupling influences transcription termination (3, 9, 10), we also studied the role of the coupling in 3′ end formation. We report the occurrence of transcription at the two types of signals, how they are regulated by translation–transcription coupling, and how the final 3′ end of the terminated RNA is actually generated by RNA processing, thus, masking the actual termination processes.

Results

Two Tandem Termination Signals at the End of *gal*. The end of the *gal* operon is depicted in Fig. 1. The stop codon (UAA) of the last ORF, *galM*, is successively followed by 6 bp, a dyad symmetry and three uridine residues in RNA, the latter two features becoming a hairpin structure and three uridine residues in RNA which comprise a potential RIT signal. Downstream of this is located a 78-bp C-rich region in the template strand which in RNA becomes a potential RDT signal (11–13).

In Vitro and in Vivo Transcription in *gal*. The 3′ ends of specific transcripts were identified and mapped both in vitro and in vivo by using 3′ RACE. For in vitro transcription termination assays, we used the plasmid p*gal-gpmA* as a DNA template containing the entire *gal* operon, as well as the next monocistronic operon *gpmA* with the same direction of transcription (*SI Appendix*, Figs. S1 and S2*A*). The 3′ RACE assays on the *gal* transcript generated in vitro identified two major 3′ ends at positions 4396 and 4421 (Fig. 2, lane 1) and two minor ends at 4313 and 4315. Since no obvious secondary structures or C richness are located upstream of the 3′ ends of the two major RNA, it is likely that these 3′ ends result from "elemental" transcription pausing that usually occurs preceding actual transcription termination events (14–17). Sequence analysis of these sites also suggested that they contain a consensus core sequence for elemental pausing (18, 19). Of the minor RNA species without Rho, the one ending at 4315 (Fig. 2, lane 1) is located seven nucleotides downstream from the foot of the stem of the terminator hairpin (Fig. 1*B*). The 3′ end at 4315 may be the result of the RIT signal which is functioning inefficiently (poor U richness). The transcription factors NusA and NusG, respectively, stimulated and reduced the production of

Significance

Transcription termination by RNA polymerase in prokaryotes is well understood in contrast to similar mechanisms in higher organisms. Despite the in vitro occurrence of two types of demonstrable transcription termination events in prokaryotes at the end of transcription units, they are obscured in vivo in two ways: suppression of termination by traversing of the RNA polymerase through the termination sites when coupled to translation, or by further processing of the actual terminated RNA 3′ ends by RNases, as in eukaryotes.

Author contributions: H.M.L. designed research; X.W., M.P.A.N., H.J.J., and Y.L. performed research; X.W., M.P.A.N., H.J.J., Y.L., J.H., S.A., and H.M.L. analyzed data; and S.A. and H.M.L. wrote the paper.

Reviewers: M.E.G., Columbia University; and J.W.R., Cornell University.

The authors declare no conflict of interest.

Published under the PNAS license.

[1]To whom correspondence may be addressed. Email: adhyas@mail.nih.gov or hmlim@cnu.ac.kr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1813181116/-/DCSupplemental.

Published online February 19, 2019.

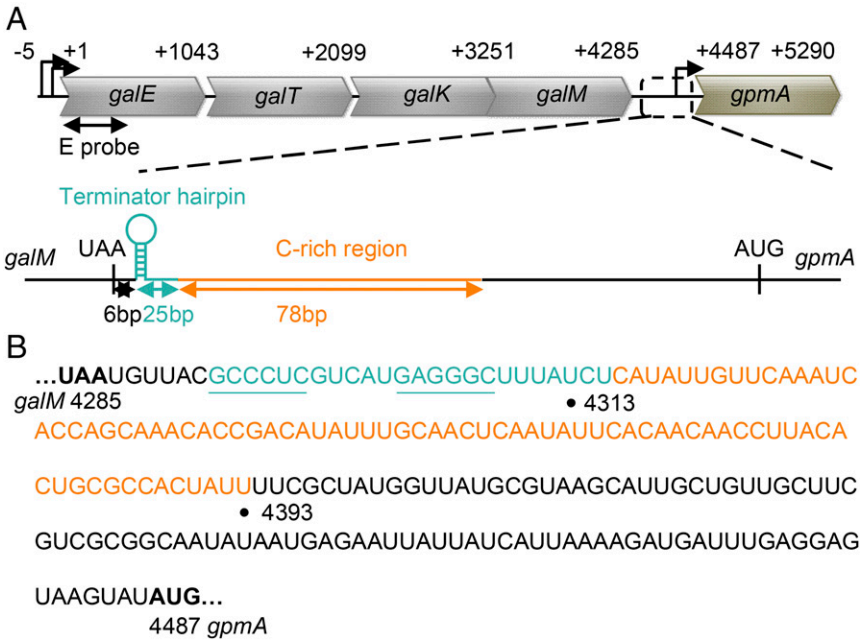

Fig. 1. (*A*) Schematic representation of the *gal* operon and the neighboring *gpmA* gene. Signals for the two tandem transcription terminations downstream of *galM* were shown. The RIT signal (the terminator hairpin) is presented as a cyan hairpin structure, and the RDT signal (C-rich region) is depicted as an orange line. Numbers indicate nucleotide position from the *gal* transcription initiation site, +1. E probe, which hybridizes to the first 500 nucleotides of *galE*, was used as the probe in Northern analyses throughout this study. (*B*) RNA sequences from the stop codon of *galM* to the start codon of *gpmA*. The terminator hairpin sequences are presented in cyan. The stem sequences are underlined. The C-rich region sequence is shown in orange. Numbers: 4285, the third nucleotide of the *galM* stop codon; 4313, the 3′ end of mM1 *gal* mRNA; 4393, the 3′ end of the C-rich region; and 4487, the first nucleotide of the *gpmA* start codon.

the RNA with 3′ end at 4315, suggesting that the 4315 species is the result of RIT (Fig. 2*A*, lanes 3 and 4, respectively). NusA and NusG have been previously established to stimulate and inhibit RIT functions in general (17, 20–23). We believe that the other minor band ending at 4313 (lane 1, Fig. 2*A*) mimicking in vivo *gal* mRNA mM1 and is generated by processing of the RNA ending at 4315 by RNase present as contaminants of RNA polymerase (see below). In the presence of Rho, however, a major RNA appeared with a 3′ end at 4409 (Fig. 2*A*, lane 2). Since 4409 is immediately downstream of the C-rich region (Fig. 1*B*), we conclude that Rho terminates transcription at 4409.

Total RNA extracted from wild-type (WT) *E. coli* cells was subjected to the 3′ end analysis. There was only one major *gal* mRNA in vivo whose 3′ end is at 4313, as was observed previously (8) (Fig. 2*A*, lane 5). We investigated whether this RNA is in any way related to the proximal RIT in *gal* that generated RNA with 3′ ends at 4315 in vitro. We surmised that in vivo the 3′ end at 4313 may be generated by exonucleolytic processing of either RIT or RDT generated RNA hairpin blocking further degradation at the 3′ end. The validity of this hypothesis is supported by the following experiments.

Functioning of RIT and RDT Signals. The putative RIT signal in *gal* having only three uridine residues in the U track may be at least partially defective and/or acts merely as a structural block to the 3′ to 5′ exonucleolytic digestion of 3′ ends generated downstream. We investigated this, by inserting a second *gal* RIT at position 4333 (marked RIT^2), which is 37 nucleotides downstream of the original RIT (*SI Appendix*, Fig. S2*B*) and introducing the construct into the plasmid p*gal-gpmA* for in vivo experiments in a Δ*gal* host. We reason that if the *gal* RIT is partially active in transcription termination, the RNA that stops at 4315 would be processed back to position 4313. The rest of the RNA that does not stop at RIT would stop at the downstream RDT but would be processed back to position 4313. Consistently the 3′ RACE assay results of the *gal* transcript from wild-type in vivo showed only a single RNA species ending at 4313. We note that the 4313 band is frequently a doublet; further investigation of the doublet bands revealed that they end at positions 4313 and 4314, suggesting that the nucleolytic process is sloppy. In the RIT^2-carrying strain, we expect that some of the read-through transcription from the first RIT would stop at the second RIT while others would go further, to be stopped at the RDT. In the latter two cases, RNA would be processed back to the stem and loop structure of the second RIT. The results showed that the RIT^2 variant produced RNA with the 3′ end at the 4313 position as expected, but in reduced amounts, but also made an RNA with the 3′ end at 4367 (Fig. 2*B*, lane 2). Reduction of the 4313 band to about 50% of WT and the appearance of about an equal amount of the 4367 RNA in RIT^2 suggest that 50% of transcription reaching the end of the *gal* operon terminates at the first RIT and the other 50% of transcription continues downstream. The 3′ end at 4367 is immediately downstream of the foot of the stem of the inserted second RNA hairpin, strongly suggesting that the 4367 RNA is a result of RNA processing blocked by the second RIT hairpin. We conclude that transcription terminates in *gal* at the RIT, albeit less efficiently in wild type, which is processed back from 4315 to 4313, while the remaining transcription passes through and terminates downstream by Rho. The latter is processed back, also creating the 4313 end, which is the major 3′ end of the *gal* operon mRNA.

We investigated the proposed role of the C-rich region in RDT function and authenticated the RDT at the end of the *galM* gene. Three to six cytidine residues in this region could be grouped into six clusters of 13 nucleotides (Fig. 1*B*). Twenty-three cytidine residues occur within a total of 78 consecutive nucleotides from 4316 to 4393, yielding an ∼30% cytidine content. We replaced all 23 cytidines with guanine to generate the RDT° mutant (*SI Appendix*, Fig. S2*C*). If Rho-dependent termination fails at the *gal* RDT° site in vivo, transcription is expected to continue until the next termination signal at the end of the *gpmA* gene. We expect that in the RDT°, the *gal* RNA would be 1 kb longer compared with the full-length *gal* mRNA mM1 (the distance between the stop codons of *galM* and *gpmA* genes is about 1 kb). Northern blot analysis indeed showed that the RDT° mutant produced an RNA band (called mM2) that is 1 kb longer than the mM1 RNA (Fig. 3*A*, lane 2). Next, we changed 146 nucleotides downstream from the stop codon of *gpmA* in the RDT° mutant, as this region harbors a putative transcription terminator hairpin (RIT sequence) followed by a C-rich region (RDT sequence). This new mutant is referred to as RDT°* (*SI Appendix*, Fig. S2*D*). Northern assays revealed that the

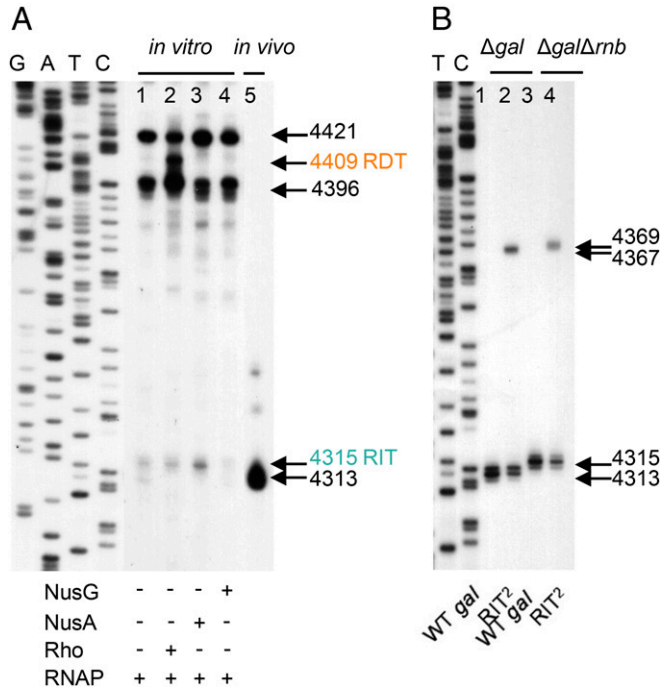

Fig. 2. (*A*) The 3′ RACE assay on transcripts generated from in vitro transcription of the entire *gal* operon (lanes 1–4) and from the *gal* transcripts generated in vivo (lane 5). The in vitro transcription was performed in the presence of Rho (100 nM), NusA (100 nM), or NusG (100 nM). The horizontal arrows and numbers indicate the position of the 3′ ends of the *gal* transcripts from the transcription initiation site (+1 in Fig. 1*A*). The RIT signal 4315 and RDT signal 4409 were shown in cyan and orange, respectively. See *SI Appendix, Supplementary Materials and Methods* for in vitro transcription reaction conditions. GATC are the DNA sequencing ladders. (*B*) The 3′ RACE assay on transcripts generated in WT and the RIT^2 mutant. The experiments were performed in Δ*gal* and Δ*gal*Δ*rnb* strains, respectively. In the *Δrnb* strain, the gene encoding RNase II is deleted.

mM2 RNA was produced in much reduced amount in the RDT°* mutant (Fig. 3*A*, lane 3). We also detected a smeared signal below the mM2 band as well as a more intense mM1 band. These observations show exonucleolytic digestions of the transcripts extended beyond the end of *gpmA*, suggesting that transcription that is not terminated by Rho-dependent termination at the end of the *gal* operon continues and terminates at the end of *gpmA*.

This conclusion was further confirmed in wild-type cells. We inhibited the Rho function in wild-type cells by exposing the growing cells to the Rho inhibitor bicyclomycin (BCM) for 10 min and RNA was analyzed by Northern blot (24–26). In these cells, mM2 RNA increased as the BCM concentration increased (Fig. 3*B*). At 80 μg/mL BCM, the mM2 species was in the equivalent amount to mM1. These results corroborated our findings that mM2 RNA is generated in the RDT° mutant due to the impairment of Rho-dependent termination at the end of the *gal* operon. Taken together, our experiments establish the role of the C-rich region in Rho-dependent termination and demonstrate that the RDT signal functions in the *gal* operon in vivo. The results of BCM mean that 50% of mM1 results from Rho-mediated termination at the downstream RDT that is then processed back to mM1.

To establish the role of the hairpin RNA structure in the partially defective RIT at the end of *galM*, we removed one to four of the G:C base pairs from the bottom of the stem of the hairpin to generate a set of variants referred to as RIT°1, -2, -3, or -4, respectively, and tested them both in vitro and in vivo (*SI Appendix*, Fig. S2*E*). In vitro transcription of RIT°2 (with the bottom 2 G:C bp removed) in the presence of increasing amounts of NusA that stimulates RIT RNA (Fig. 2*A*) showed that the RNA with the 3′ end at 4315 increased as NusA concentration increased in wild-type as expected (Fig. 4*A*); however, the 4315 RNA in the RIT°2 mutant was absent even at the highest concentration of NusA tested (100 nM, Fig. 4*A*). These results demonstrate that the RIT at the end of *gal* giving rise to the 4315 RNA is an authentic albeit weak terminator; removal of the 2 G:C bp from the stem of the RNA hairpin totally impaired Rho-independent termination.

We also tested whether the RIT termination fails at the end of the *gal* operon and continues to the end of the next gene, *gpmA*, and terminates, generating the longer mM2 RNA in vivo by using the set of RIT°1, -2, -3, and -4 mutants in Northern blots. The RIT°1, -2, -3, and -4 mutants showed increasing levels of the mM2 RNA as more G:C bp were removed from the stem, authenticating the termination capacity of the original RIT (Fig. 4*B*). The Northern blot gel bands were scanned for quantification. The relative amounts of mM2 RNA in different lanes are reported in the legend of Fig. 4*B*. This analysis also revealed lower molecular weight smears (termed mM1*) that appeared slightly shorter than the full-length mM1 RNA band (Fig. 4*B*). The 3′ RACE assay showed that mM1* RNA have 3′ ends at 4208–4200 and 4194–4170 (Fig. 4*C*). We conclude that transcription through the terminator hairpin in these strains has been terminated by Rho, and the 3′ ends generated by Rho-dependent termination have been exonucleolytically digested to 4208–4200 and 4194–4171 due to the lack of hairpin structure that prevents such digestion in the RIT° mutants. The RNA processing may involve endonucleolytic activity of RNase E (27, 28). To test this idea, we performed Northern blot assays of the RNA in the

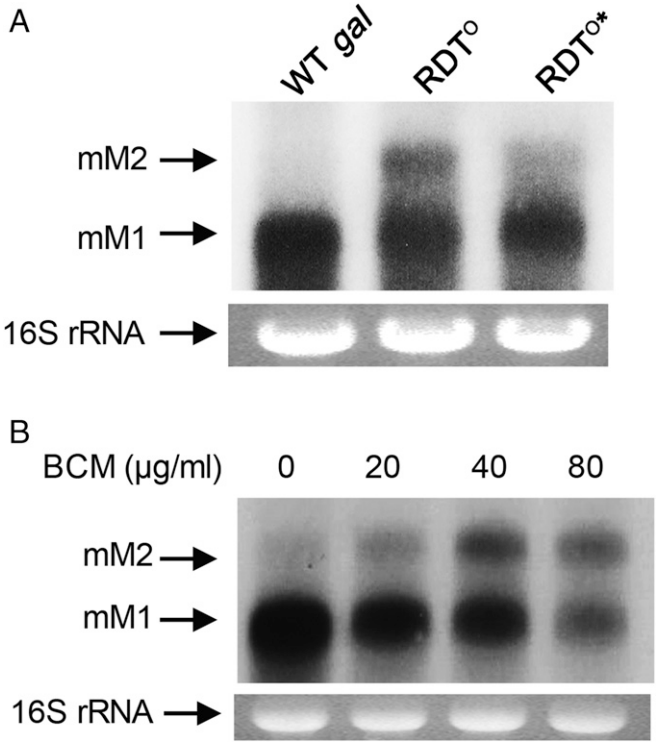

Fig. 3. (*A*) Northern analysis of mM1 and mM2 mRNA from WT, RDT°, and RDT°* mutants. mM1 is the full-length transcript of the *gal* operon. mM2 is a transcript that extends from mM1 to the end of the *gpmA* gene due to the failure of transcription termination at the end of *galM*. (*B*) Northern analysis of mM1 and mM2 from WT cells grown in different concentrations of the Rho inhibitor BCM for 10 min.

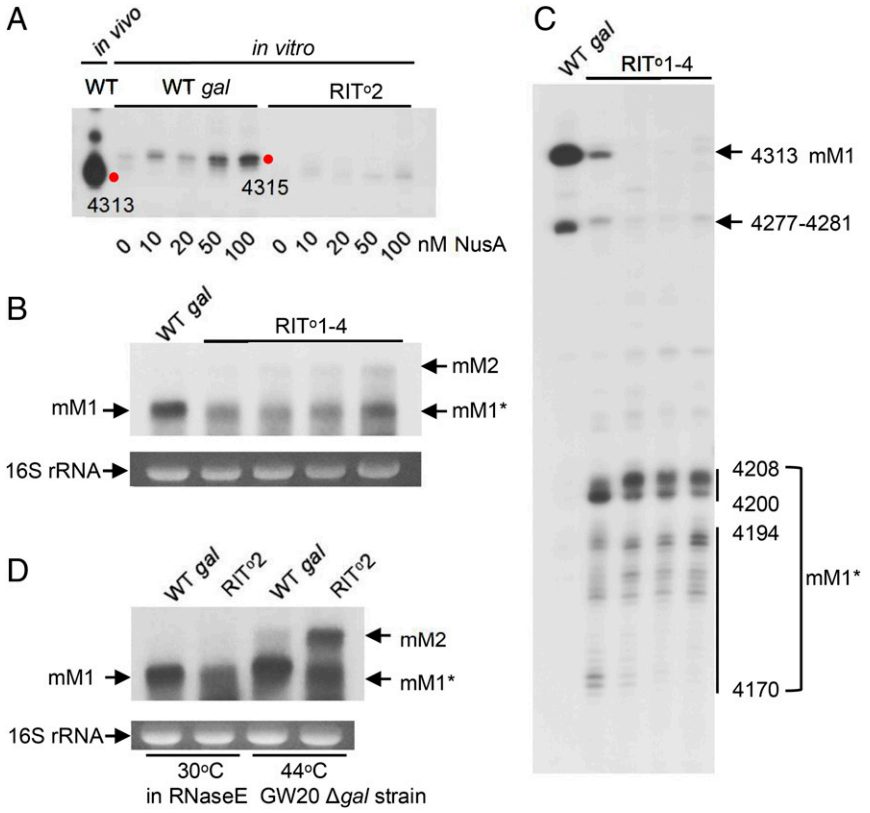

Fig. 4. (*A*) The 3′ RACE assay on transcripts generated from in vitro transcription of WT and RIT°2 mutant. In vitro transcription reactions were performed in the presence of different concentrations of NusA as indicated. (*B*) Northern analysis of the *gal* full-length transcripts generated in the RIT°1–4 series mutants. The mM1* band is about 100 nucleotides smaller than the mM1 band and was generated in all of the RIT°1–4 series mutants due to the impaired terminator hairpin structure. Note that as more base pairs from the foot of the stem of the terminator hairpin are removed, greater amounts of mM2 are generated. The mM2 RNA bands were quantified by scanning; the relative expression levels of mM2 for WT, RIT°1, RIT°2, RIT°3, and RIT°4 strains, were 0, 1.0, 6.4, 9.8, and 25.1, respectively. (*C*) The 3′ RACE assay on transcripts generated in the RIT°1–4 series mutants in (*B*). Numbers indicate the 3′ end position of the transcripts generated in vivo. (*D*) Northern analysis of the full-length transcripts generated from the WT and the RIT°2 mutant in GW20Δ*gal* (temperature-sensitive RNase E mutant) cells in which the entire *gal* operon is deleted from the chromosome. Analysis was performed on cells grown at the permissive temperature (30 °C) and at the nonpermissive temperature (44 °C).

RIT°2 variant in temperature-sensitive RNase E-mutant cells (Δ*gal rne*^{ts}) (29–31). At the permissive temperature (30 °C), the RIT°2 mutant generated mM1* but not mM2 (Fig. 4*D*). At the nonpermissive temperature (44 °C), however, the mutant generated both mM1* and mM2. Thus, in RIT°2 mutants, transcripts that extended beyond the terminator hairpin were cleaved by RNase E, confirming that Rho-independent termination at the end of the *gal* operon is impaired in the RIT°2 mutant. Why there is much mM2 RNA and not RDT-terminated RNA in the absence of processing is not clear.

Processing of Terminated RNA by RNase II. It was suggested that RNase II is involved in the exonucleolytic digestion of the 3′ RNA end generated by RDT termination in the *trp* operon of *E. coli* and in bacteriophage T3 (1, 32). We deleted the *rnb* gene encoding RNase II in the Δ*gal* strain and generated the Δ*gal*Δ*rnb* strain carrying the WT and the RIT^2 variant *gal* plasmid. Interestingly, the 3′ RACE assays of the *gal* transcripts in the two cases showed that the 3′ end at 4313 shifted to 4315 (RIT site) and the 3′ end at the 4367 site shifted to 4369 (second RIT site) (Fig. 2*B*, lane 3). If there were any RDT events downstream, an unknown RNase must have digested the Rho-terminated 3′ ends to 4315 and the RNase II removed the last two nucleotides from the 3′ end at 4315–4313.

RIT and RDT Work Independently in Vivo. Upon inhibition of Rho by BCM, the mM1 transcript was still formed (Fig. 3*B*), suggesting that the RIT element is functional. Also, generation of the mM2 band in RDT° (Fig. 3*A*) demonstrates that Rho-independent termination remains functional where Rho-dependent termination is impaired. Similarly, generation of processed mM1 (mM1*) RNA in RIT°2 mutant (Fig. 4*D*) confirmed the functioning of Rho-dependent termination and that Rho-dependent termination remains functional where Rho-independent termination is impaired. The mM1 and mM1* RNAs were formed in RIT°2 mutant (Fig. 4*D*) because the RDT RNA is processed back by ribonucleases.

Coupled Translation–Transcription Occludes Transcription Termination. In bacteria, transcription is tightly coupled with translation (33–35). We reasoned that if the stop codon of *galM* is changed to an amino acid coding triplet (UAA → AAA), transcription would remain coupled at both RIT and RDT segments and likely prevent RIT RNA hairpin formation as well as access of Rho to the C-rich region of RDT. Northern analysis in this *galM* stop^0 mutant (*SI Appendix*, Fig. S2*F*) showed a very small amount of mM1 (with the 3′ end at 4313) but a larger amount of mM2 RNA (Fig. 5). These results suggest the coupling largely suppresses both RIT- and RDT-mediated transcription termination; the mM2 production in *galM* stop° is most likely a result from the failure of both Rho-independent termination, as well as Rho-dependent termination. Thus, transcriptions that gave rise to mM2 are the ones that are coupled with translation. The small amount of mM1 produced in the *galM* stop° is likely the product of processing of residual (unsuppressed) termination at RIT and/or RDT. Any residual termination may be due to spontaneous dissociation of the leading ribosome from the elongating RNA polymerase allowing RNA to terminate by RIT or RDT. The trailing ribosomes do not get connected to a transcribing RNA polymerase for coupling. The suppression of transcription termination by coupling is analogous to suppression of intragenic termination signal by translation–transcription coupling (10, 11).

Discussion

RNA 3′ End Formation in Prokaryotes and Eukaryotes. In eukaryotic transcription systems, definitive factor- or sequence-dependent transcription terminating mechanisms have not been clearly established. The 3′ end formation in these systems appears to be the result of RNA processing 3′–5′ exonucleases triggered by

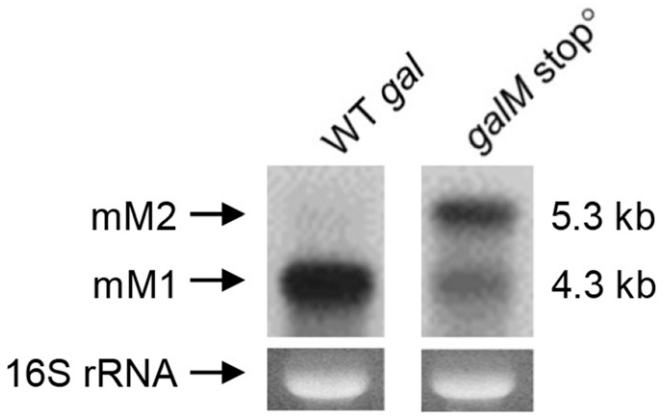

Fig. 5. Northern analysis of the full-length transcripts from the *gal* operons in WT *gal* and *galM* stop° strains.

components of the elongating RNA polymerase-associated factors at some critical points in nascent RNA (36–38). In prokaryotic systems, specific transcription termination mechanisms have been well established in vitro. Nevertheless, these signals appear to be silent when analyzed in vivo. In the *gal* operon, the factor-independent RIT and the Rho-dependent RDT signals clearly function both in vitro and in vivo. However, termination at these two sites are almost invisible in vivo. The major 3′ end of the in vivo message is different from the 3′ ends of RIT and RDT located downstream. The reasons for the invisibility of transcription termination in vivo are twofold.

Transcription–Translation Coupling. The effectiveness of both termination signals is dependent upon exposure of nascent RNA (*i*) to form a hairpin structure followed by several uridine residues in RIT, and (*ii*) with cytidine richness for Rho binding in RDT. In transcription–translation coupling, the leading ribosome on the nascent mRNA is physically linked to the elongating RNA polymerase through a bridge protein NusG (33). If coupled, transcription continues through the region of termination sequences, preventing RNA hairpin formation or access of Rho to cytidine residues in RNA, thus, suppressing transcription termination. This is common in transcription termination signals located within a gene; transcription terminates at intragenic signals only when transcription is decoupled by a mutational stop codon occurring at a site preceding the transcription termination (polarity in gene expression) (10, 12, 13, 39, 40). The distance between the active site of RNA polymerase and the aminoacylation site in ribosomes on RNA in a coupled elongating complex is 30 nucleotides (35). For transcription to terminate at termination signals, the stop codon at the end of the last gene (the decoupling point) must be located more than 30 nucleotides from the transcription termination point to allow proper exposure of the terminating RNA (hairpin formation or cytidine-rich RNA exposure) (9). A lesser distance would interfere in termination functions of the RNA. A search in the transcription terminator database (41) revealed that out of the 850 annotated ORF-containing operons, 688 of them contain recognizable termination signals: 23 contain RIT, 535 contain RDT, and 130 contain both (Fig. 6*A*). We analyzed the distance between the stop codon of the terminal cistron and RIT and also RDT in the operons. The distribution of the distance in nucleotides between stop codons and termination signals shows that 73% of RITs are within 30 nucleotides (Fig. 6*B*), suggesting that a large majority of RITs would be at least partially suppressed, allowing read-through transcription. Similar analysis of RDT signals shows that the frequency distribution of the distance between stop codons and RDT is random and do not have any preference (Fig. 6*C*); only 20% of the signals are suppressible. The physiological significance of the observation that RIT signals are mostly suppressible by translation–transcription coupling but the RDT signals are not is unknown.

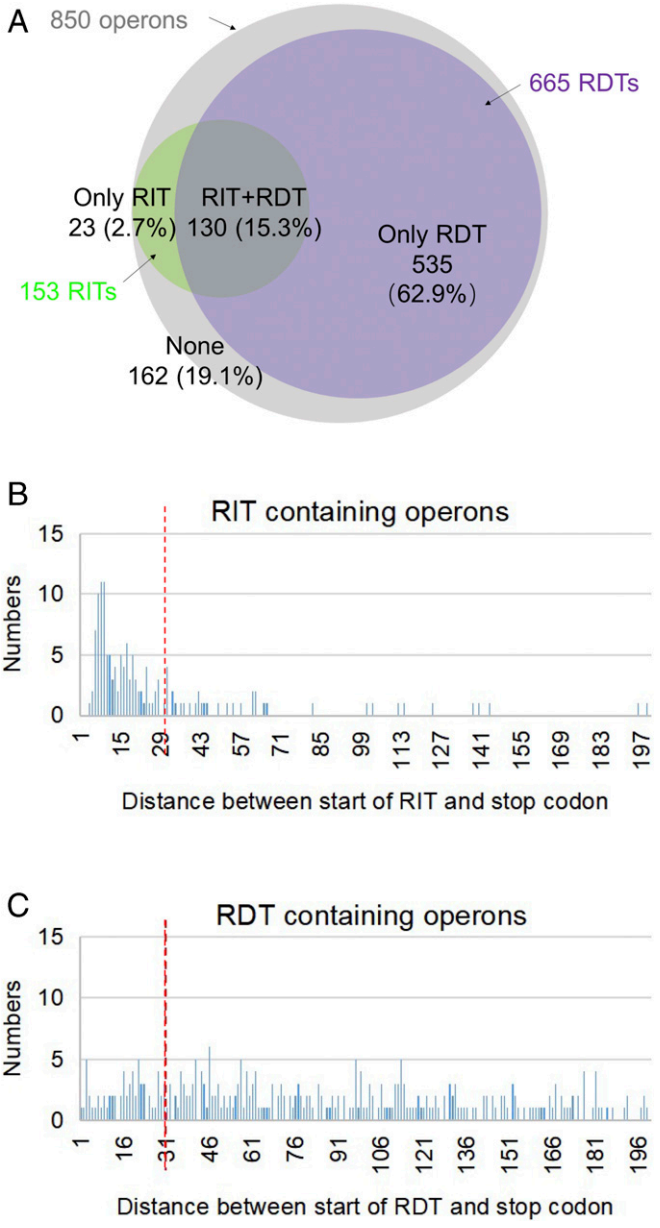

Fig. 6. In silico analyses of RITs and RDTs at the end of 850 operons in *E. coli*. (*A*) To get an overview of the genome distributions of terminators, RIT and RDT were predicted, and distance of terminators to termination stop codons was analyzed. First, we got the operon organization information from database DOOR (42), 850 operons with more than one ORF were annotated in the genome of *E. coli* MG1655. Next, RITs were collected from WebGeSTer DB (41). RDTs were predicted using the EMBOSS freak program (43). The specific parameter settings are described in *SI Appendix, Supplementary Materials and Methods*. Finally, the RITs and RDTs located downstream of 850 operons were collected and analyzed. The Venn diagram illustrates the occurrence of RITs (green) and RDTs (violet) at the end of 850 operons (gray). (*B*) Bar graph demonstrating the number of operons containing RIT signal and distance (in nucleotides) between the stop codon of the last cistron of the operon and the RIT signal. The vertical red broken line indicates the distance of 30 nucleotides. (*C*) Bar graph demonstrating the number of operons containing RDT signal and distance (in nucleotides) between the stop codon of the last cistron of the operon and the RDT signal. The vertical red broken line indicates the distance of 30 nucleotides.

RNA Processing. In the absence of coupling, transcription terminates at the RIT or RDT. But the terminated transcripts are not observed in vivo. When transcription terminated at RIT, the 3′ end of the terminated RNA was shorter in vivo. This is also true for RDT-terminated RNA. In both cases the terminated transcripts are processed by ribonucleases. For the RIT-terminated RNA in *gal* ending at the uridine residues after a hairpin, RNase II digests the 3′ terminated end further, making a new 3′ end. For RDT, the long RNA beyond the stop codon of *galM* appears to be processed by endoribonucleases like RNase E and by exoribnucleases like RNase II. The exoribonuclease-mediated processing stops when it encounters the RNA hairpin structure and ends up with a new 3′ end created by processing and not by RIT- or RDT-mediated transcription termination.

In summary, actual transcription termination by RNA polymerase in bacteria is well understood in contrast to similar mechanisms in higher organisms. Despite transcription termination, the actual 3′ ends in prokaryotes after all also appear to be generated by RNA processing in vivo, as in eukaryotes.

Materials and Methods

Extraction of RNA from *E. coli* Cells. Total RNA was prepared from 2×10^8 *E. coli* cells grown to OD_{600} of 0.6 in LB. Total RNA was purified from cleared cell lysates using the Direct-zol RNA MiniPrep kit (Zymo Research).

The 3′ RACE Assay. A RNA ligation reaction (15 μL volume containing 2.5 μg of total RNA, 5 U T4 RNA ligase, 2 nM synthetic RNA oligomer) was performed. One microgram of RNA in the ligation reaction was reverse transcribed using a DNA primer complementary to the synthetic RNA oligomer used in the RNA ligation reaction. The resulting *gal* cDNA was PCR amplified. The 3′ ends of the *gal* mRNAs were assayed by the primer-extension reaction using the amplified *gal* cDNA as a template.

In Vitro Transcription. The reaction was performed in a 50-μL volume containing the template DNA (plasmid p*gal-gpmA*), and *E. coli* σ^{70}-RNA polymerase.

Northern Blot Analysis. Ten micrograms of total RNA was subjected to agarose gel electrophoresis and blotted to a positively charged nylon membrane. The blot was baked and probed with a P^{32}-labeled 500-bp DNA fragment.

The *gal* mutants were constructed in the plasmid p*gal-gpmA* and assayed in MG1655Δ*gal*.

Protein Purification. The corresponding genes were cloned to pET-15b vector; proteins genetically fused His-tagged at its N terminus were overexpressed and purified by Ni-NTA affinity chromatography.

Detailed information on the experimental methods can be found in *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by the Basic Science Research Program of the National Research Foundation of Korea (Grant 2016 R1A2B1013515); the National Natural Science Foundation of China (Grant 31600061); and by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research (Project/Grant ZIA BC 010017).

1. Adhya S, Sarkar P, Valenzuela D, Maitra U (1979) Termination of transcription by Escherichia coli RNA polymerase: Influence of secondary structure of RNA transcripts on rho-independent and rho-dependent termination. *Proc Natl Acad Sci USA* 76:1613–1617.
2. Richardson JP (1991) Preventing the synthesis of unused transcripts by Rho factor. *Cell* 64:1047–1049.
3. Adhya S, Gottesman M (1978) Control of transcription termination. *Annu Rev Biochem* 47:967–996.
4. Roberts JW (1969) Termination factor for RNA synthesis. *Nature* 224:1168–1174.
5. Brennan CA, Dombroski AJ, Platt T (1987) Transcription termination factor rho is an RNA-DNA helicase. *Cell* 48:945–952.
6. Gusarov I, Nudler E (1999) The mechanism of intrinsic transcription termination. *Mol Cell* 3:495–504.
7. Yarnell WS, Roberts JW (1999) Mechanism of intrinsic transcription termination and antitermination. *Science* 284:611–615.
8. Lee HJ, Jeon HJ, Ji SC, Yun SH, Lim HM (2008) Establishment of an mRNA gradient depends on the promoter: An investigation of polarity in gene expression. *J Mol Biol* 378:318–327.
9. Li R, Zhang Q, Li J, Shi H (2016) Effects of cooperation between translating ribosome and RNA polymerase on termination efficiency of the Rho-independent terminator. *Nucleic Acids Res* 44:2554–2563.
10. Das A, Court D, Gottesman M, Adhya S (1977) *Polarity of Insertion Mutations Is Caused by Rho-Mediated Termination of Transcription* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY), pp 93–97.
11. Ray-Soni A, Bellecourt MJ, Landick R (2016) Mechanisms of bacterial transcription termination: All good things must end. *Annu Rev Biochem* 85:319–347.
12. Alifano P, Rivellini F, Limauro D, Bruni CB, Carlomagno MS (1991) A consensus motif common to all Rho-dependent prokaryotic transcription terminators. *Cell* 64:553–563.
13. Alifano P, Ciampi MS, Nappo AG, Bruni CB, Carlomagno MS (1988) In vivo analysis of the mechanisms responsible for strong transcriptional polarity in a "sense" mutant within an intercistronic region. *Cell* 55:351–360.
14. Wang D, et al. (1995) Discontinuous movements of DNA and RNA in RNA polymerase accompany formation of a paused transcription complex. *Cell* 81:341–350.
15. Neidhardt F, Richardson J, Greenblatt J (1996) *Escherichia coli and Salmonella: Cellular and Molecular Biology* (ASM Press, Washington, DC).
16. Richardson LV, Richardson JP (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.
17. Kassavetis GA, Chamberlin MJ (1981) Pausing and termination of transcription within the early region of bacteriophage T7 DNA in vitro. *J Biol Chem* 256:2777–2786.
18. Larson MH, et al. (2014) A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* 344:1042–1047.
19. Vvedenskaya IO, et al. (2014) Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science* 344:1285–1289.
20. Farnham PJ, Greenblatt J, Platt T (1982) Effects of NusA protein on transcription termination in the tryptophan operon of Escherichia coli. *Cell* 29:945–951.
21. Lau LF, Roberts JW, Wu R (1983) RNA polymerase pausing and transcript release at the lambda tR1 terminator in vitro. *J Biol Chem* 258:9391–9397.
22. Burns CM, Richardson LV, Richardson JP (1998) Combinatorial effects of NusA and NusG on transcription elongation and Rho-dependent termination in Escherichia coli. *J Mol Biol* 278:307–316.
23. Burova E, Hung SC, Sagitov V, Stitt BL, Gottesman ME (1995) Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. *J Bacteriol* 177:1388–1392.
24. Peters JM, et al. (2009) Rho directs widespread termination of intragenic and stable RNA transcription. *Proc Natl Acad Sci USA* 106:15406–15411.
25. Cardinale CJ, et al. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science* 320:935–938.
26. Zwiefka A, Kohn H, Widger WR (1993) Transcription termination factor rho: The site of bicyclomycin inhibition in Escherichia coli. *Biochemistry* 32:3564–3570.
27. Carpousis AJ (2007) The RNA degradosome of Escherichia coli: An mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* 61:71–87.
28. Mackie GA (2013) RNase E: At the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* 11:45–57.
29. Wang X, et al. (2014) Expression of each cistron in the gal operon can be regulated by transcription termination and generation of a galk-specific mRNA, mK2. *J Bacteriol* 196:2598–2606.
30. Kuwano M, et al. (1977) Gene affecting longevity of messenger RNA: A mutant of Escherichia coli with altered mRNA stability. *Mol Gen Genet* 154:279–285.
31. Ono M, Kuwano M (1980) Chromosomal location of a gene for chemical longevity of messenger ribonculeic acid in a temperature-sensitive mutant of Escherichia coli. *J Bacteriol* 142:325–326.
32. Mott JE, Galloway JL, Platt T (1985) Maturation of Escherichia coli tryptophan operon mRNA: Evidence for 3′ exonucleolytic processing after rho-dependent termination. *EMBO J* 4:1887–1891.
33. Burmann BM, et al. (2010) A NusE:NusG complex links transcription and translation. *Science* 328:501–504.
34. Proshkin S, Rahmouni AR, Mironov A, Nudler E (2010) Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* 328:504–508.
35. Kohler R, Mooney RA, Mills DJ, Landick R, Cramer P (2017) Architecture of a transcribing-translating expressome. *Science* 356:194–197.
36. Porrua O, Libri D (2015) Transcription termination and the control of the transcriptome: Why, where and how to stop. *Nat Rev Mol Cell Biol* 16:190–202.
37. Grzechnik P, Gdula MR, Proudfoot NJ (2015) Pcf11 orchestrates transcription termination pathways in yeast. *Genes Dev* 29:849–861.
38. Shandilya J, Roberts SG (2012) The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. *Biochim Biophys Acta* 1819:391–400.
39. De Crombrugghe B, Adhya S, Gottesman M, Pastan I (1973) Effect of Rho on transcription of bacterial operons. *Nat New Biol* 241:260–264.
40. Morse DE, Yanofsky C (1969) Polarity and the degradation of mRNA. *Nature* 224: 329–331.
41. Mitra A, Kesarwani AK, Pal D, Nagaraja V (2011) WebGeSTer DB–A transcription terminator database. *Nucleic Acids Res* 39:D129–D135.
42. Mao X, et al. (2014) DOOR 2.0: Presenting operons and their functions through dynamic and integrated views. *Nucleic Acids Res* 42:D654–D659.
43. Bossi L, Schwartz A, Guillemardet B, Boudvillain M, Figueroa-Bossi N (2012) A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes Dev* 26: 1864–1873.