High-resolution RNA 3′-ends mapping of bacterial Rho-dependent transcripts

Daniel Dar and Rotem Sorek[*]

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel

ABSTRACT

Transcription termination in bacteria can occur either via Rho-dependent or independent (intrinsic) mechanisms. Intrinsic terminators are composed of a stem-loop RNA structure followed by a uridine stretch and are known to terminate in a precise manner. In contrast, Rho-dependent terminators have more loosely defined characteristics and are thought to terminate in a diffuse manner. While transcripts ending in an intrinsic terminator are protected from 3′-5′ exonuclease digestion due to the stem-loop structure of the terminator, it remains unclear what protects Rho-dependent transcripts from being degraded. In this study, we mapped the exact steady-state RNA 3′ ends of hundreds of *Escherichia coli* genes terminated either by Rho-dependent or independent mechanisms. We found that transcripts generated from Rho-dependent termination have precise 3′-ends at steady state. These termini were localized immediately downstream of energetically stable stem-loop structures, which were not followed by uridine rich sequences. We provide evidence that these structures protect Rho-dependent transcripts from 3′-5′ exonucleases such as PNPase and RNase II, and present data localizing the Rho-utilization (*rut*) sites immediately downstream of these protective structures. This study represents the first extensive *in-vivo* map of exact RNA 3′-ends of Rho-dependent transcripts in *E. coli*.

INTRODUCTION

Transcription termination sets the gene 3′-end boundary and prevents leakage into downstream genes, enabling controlled expression in the densely packed genomes of bacteria (1). In addition, termination determines the exact 3′-terminal sequence of the mRNA, which often holds important information for transcript stability or for regulating gene expression (2,3). Bacteria rely on two major mechanisms for terminating the transcription reaction, which are defined according to their dependence on the proteinaceous termination factor, Rho (4).

Rho-independent terminators, also known as intrinsic terminators, efficiently destabilize the elongating RNA-polymerase (RNAP) complex in the absence of Rho and are encoded by a short ∼30nt DNA sequence downstream of the protein-coding region of the gene, as well as in some regulatory 5′ mRNA leaders (5,6). These terminators are composed of two main modules: a GC-rich sequence that folds into an energetically stable stem-loop RNA structure and a 7–8 nt uridine rich sequence (7–10). Termination initiates when the U-rich tract is transcribed and occupies the transcription bubble, causing the RNAP to pause and allowing the upstream stem-loop forming sequence to nucleate, which disrupts the transcription complex (8,9,11). Thus, the site of termination is precisely defined by the terminator sequence and results in mRNAs with stem-loop structures at their 3′-end, a feature that protects them from being digested by processive 3′-5′ exonucleases (8,12).

Transcripts regulated by Rho-dependent termination must associate with Rho during their transcription elongation stage. Rho forms a homohexameric ring complex and translocates on the nascent mRNA through its pore in an ATP-dependent manner from 5′ to 3′, resulting in a competition with the elongating RNAP (8). Rho preferentially binds the nascent mRNA at unstructured, C-rich, G-poor RNA sequences that activate its translocation activity and are generally defined as Rho-utilization (*rut*) sequences (13,14). In contrast to the compact encoding of intrinsic terminators, *rut* sequences can be encoded by ∼80–90 nucleotides and may be positioned in variable positions at over 100nt upstream to the actual site of termination (14–16). Furthermore, the efficiency of termination is dependent on kinetic coupling between the elongating RNAP and the translocating Rho that lags behind, resulting in 'diffuse' termination at multiple positions that are enriched around RNAP pause sites (17). While mRNAs terminated by the intrinsic mechanism are protected from degradation by their 3′ stem-loop structure, it remains unknown whether or how Rho-dependent genes are stabilized.

Despite the accumulation of insights into the Rho-dependent termination mechanism, *in-vivo* mapping of 3′ termini of Rho-dependent transcripts has largely been lim-

[*]To whom correspondence should be addressed. Tel: +972 8 934 6342; Email: rotem.sorek@weizmann.ac.il

ited to a few model cases. In this study, we utilize a combination of RNA-seq approaches to map the RNA 3′-ends of both intrinsic and Rho-dependent transcripts with single nucleotide precision and across hundreds of operons.

MATERIALS AND METHODS

Strains, growth conditions and RNA extraction

Escherichia coli BW25113 and the *pnp*, *rnb* and *rnr* single deletion strains from the keio collection (18) were cultured in LB media (10 g/l tryptone, 5 g/l yeast extract 5 g/l NaCl) under aerobic conditions at 37°C with shaking. Bacterial pellets were collected by centrifugation (4,000 rpm, 5 min, at 4°C), flash frozen and stored in –80°C until RNA extraction. For RNA isolation, frozen pellets were thoroughly resuspended and mixed in 100 μl lysozyme solution (1 mg/ml in 10 mM Tris–HCl and 1 mM EDTA) pre-warmed to 37°C and then incubated at 37°C for 1 min. The cells were then lysed by immediately adding 1 ml tri-reagent (Trizol) followed by vigorous vortexing for 10 s until solution was cleared. Following an incubation period of 5 min at room temp (RT), 200 μl chloroform was added and the sample was vortexed for another 10 s until homogeneous. The sample was incubated for 2–5 min at RT until visible phase separation was observed and then centrifuged at 12,000g for 10 min. The upper phase was gently collected (about 600μl) and mixed at a 1:1 ratio with 100% isopropanol and then mixed by vortexing for 2–3 s. The sample was incubated for 1 h at –20°C and then centrifuged (14,000 rpm, 30 min, at 4°C) to collect the RNA pellet. The solution was removed without disturbing the pellet, followed by two consecutive wash rounds using 750 μl 70% ethanol. The pellets were air dried for 5 min and then dissolved in nuclease free H_2O and incubated for 5 min at 50°C. All RNA samples were treated with TURBO deoxyribonuclease (DNase) (Life technologies, AM2238).

Library construction, deep-sequencing and read mapping

Term-seq and RNA-seq libraries were constructed as described by *Dar et al.* (19) and were sequenced on the Illumina Nextseq 500 platform in paired ends mode. The sequencing data has been deposited at Gene Expression Omnibus (GEO) under the GSE109766 accession code. Reads were then mapped to the reference genome using NovoAlign (Novocraft) V3.02.02 with default parameters, discarding reads that mapped to more than one genomic position. Paired reads in which the mapped insert length was greater than 500nt were discarded. Gene annotation and sequences were downloaded from Genbank (CP009273).

Identification of dominant RNA 3′-end

The number of 3′-end reads mapped to each genomic position was measured, and sites appearing in at least two of three replicates with an average coverage greater or equal to four reads were collected for downstream analysis. For each 3′-end site the average library insert length was calculated using the paired end read mapping positions. Sites were then associated with their respective genes, requiring that the average insert length allow at least 1nt overlap with the end of

the genes' coding region. Dominant 3′-ends were assigned as the most highly covered position associated with the 3′ end of a gene. In cases where equal coverage was shared between multiple sites, the position furthest downstream was selected.

Termini sequence and structure analysis

The upstream and downstream sequences relative to the exact termination sites were collected from the latest version of the reference genome (as described above). For the nucleotide usage enrichment analyses, the frequency of each base in a given position relative to the dominant 3′ ends (–1, +1, etc.) was calculated using the sequences above and compared to an identical calculation resulting from sampling 10,000 randomly chosen intergenic positions as described in (19). The $\log_2$ fold enrichment was calculated and plotted as in Figure 2. Predicted RNA structural stability analysis was performed by folding the 45nt long DNA sequence found upstream to the dominant 3′ ends using the RNAfold software (20). Distribution comparison was performed with the Wilcoxon rank-sum test R package using the two-sided hypothesis testing mode.

Measuring rho-dependence and exonuclease activity

The data from (21) was mapped against the *E. coli* genome (accession no.: CP009273) as described above and the number of reads per nucleotide was calculated for control and for Bcm-treated samples. The RNA-seq coverage immediately downstream and upstream of dominant termini was used to estimate the readthrough in the control (basal readthrough) and in the Bcm-treated sample, where Rho is inhibited. Specifically, the average RNA-seq read coverage over the downstream 75nt and the upstream as 250nt was calculated and the readthrough was measured as their ratio. Rho-dependence was measured as the difference in readthrough between the Bcm and control samples (e.g. for 5% readthrough in control and 40% readthrough in Bcm, the estimated rho-dependence was measured as 35%). Genes covered on average with less than 20 reads per nucleotide, or where the control readthrough (basal readthrough) was greater than 30%, were discarded from this analysis. The final set included 462 transcripts for which we were able to calculate a rho-dependence score based on the readthrough in Bcm treatment, as described above. The measured rho-dependence scores in this set were used to classify transcripts according to their termination strategy. Transcripts with a rho-dependence score of 5% or less were classified as rho-independent ($n = 174$) and transcripts with a score of 30% or more were classified as Rho-dependent ($n = 144$). Transcripts with values greater than 5% but smaller than 30% ($n = 144$) were not used in the analyses. We found that the results described in the manuscript were not sensitive to small changes in these cutoff values. The entire set of 462 transcripts considered in this analysis is available in Supplementary Table S2 along with their respective measured rho-dependence scores.

Exonuclease activity was measured in an identical manner to rho-dependence, but the Bcm samples were replaced with samples generated with exonuclease-deleted strains

taken from the keio collection (*pnp*, *rnb* and *rnr*) and compared to a control generated with the parental keio strain (WT) ([18]).

RESULTS

Mapping the 3′ termini of Rho-dependent and independent transcripts in *Escherichia coli*

Accurate and comprehensive mapping of RNA 3′-ends can shed important light on termination mechanisms in bacteria and archaea ([19],[22]). While *E. coli* is considered one of the best studied model microbes, a high-resolution map of RNA 3′-ends in this organism was still missing ([23]). We reasoned that such a map, in combination with a method for measuring rho-dependence, would enable a direct comparison of 3′ termini between transcripts terminated by Rho-dependent or independent mechanisms. We therefore applied the recently developed term-seq method ([24]) to comprehensively map the exposed RNA 3′-ends in three exponentially growing *E. coli* cultures, which resulted in over 5.5 million uniquely mapped paired-end sequencing reads representing the exact positions of *in-vivo* RNA 3′-ends (Methods). The dominant 3′ terminus, defined as the most highly covered, reproducibly occurring 3′-end present downstream of the gene (detected in at least two independent experiments with an average coverage of at least four reads across all three replicates), was detected for 1098 individual transcripts (Methods; Supplementary Table S1).

While term-seq provides high-resolution maps of the RNA 3′ ends associated with genes, more data is required to determine whether the genes are dependent on rho for their termination. In an earlier study, *Peters et al.* identified over a thousand regions where Rho-dependent termination takes place by exposing *E. coli* bacteria to a sub-lethal dosage of Bicyclomycin (Bcm), a potent Rho inhibitor, and then sequencing mRNAs using a strand-specific RNA-seq protocol ([21]). In Bcm treated bacteria, the activity of Rho is compromised and leads to leaky transcription beyond Rho-dependent termination sites ([21],[25],[26]). To complement the term-seq analysis, we used the *Peters et al.* ([21]) data to evaluate the rho-dependence of each of the transcripts for which we were able to assign a dominant 3′ end, and for which significant coverage was available in the *Peters et al.* ([21]) RNA-seq experiments. Specifically, we estimated rho-dependence as the Bcm-dependent 'read-through', i.e., the relative increase in RNA-seq coverage downstream of the dominant 3′ termini in Bcm-treated bacteria as compared with the control (Figure [1]; Materials and Methods).

We calculated a rho-dependence metric for 42% (462/1098) of the genes in our data, which were covered by sufficient amount of RNA-seq reads in the *Peters et al.* experiment and for which we were able to assign an dominant 3′ end using term-seq (Figure [1]A–C; Supplementary Table S2; Materials and Methods). This analysis divided our gene set into a clear group of rho-independent genes and a second group of genes showing a varied distribution of rho-dependence (Figure [1]C). Based on these data, we classified our genes into two groups: genes displaying <5% rho-dependence (i.e. <5% 'read through' downstream of the gene terminus in the Bcm-treated samples) were classified as rho-independent (Figure [1]A) and those presenting

Bcm-dependent leakage rates larger than 30% were marked as Rho-dependent (Figure [1]B). This classification resulted in a set of 174 independent and 144 Rho-dependent genes (Figure [1]A–C). This analysis suggests that >30% (144/462) of *E. coli* genes depend on rho for their termination, a figure which agrees with previous estimates of 20–50% ([25–27]). Whereas most of the gene-associated rho-dependent 3′ termini analyzed above were mapped to intergenic regions, we also identified 14 cases where the termini were localized over 30 bases into the downstream gene (Supplementary Table S2).

We find that rho-dependent and independent genes are encoded on both strands with similar frequency and are uniformly distributed across the genome. In addition, comparing the RNA-seq coverage levels of Rho-dependent and independent genes, we found that genes with intrinsic termination mechanisms had on average a 6-fold higher expression level than Rho-dependent genes ($P < 10^{-17}$, Wilcoxon rank-sum; Figure [1]D), suggesting that termination strategy might be linked with additional aspects of gene regulation.

Molecular determinants of RNA 3′-ends of Rho-dependent transcripts

The known molecular mechanism of Rho-dependent termination, in which the rho factor must 'catch up' with the transcribing RNAP complex to catalyze its dissociation, suggests that the termination sites of Rho-dependent genes ought to be dispersed across multiple positions downstream of the gene ([8]). Indeed, *in-vitro* Rho-dependent termination assays often show multiple or smeared termination patterns ([17]). In contrast, we find that the vast majority of the *in-vivo* RNA 3′-end reads associated with Rho-dependent transcripts are narrowly distributed within a space of 2–3nt relative to the dominant site, as is also seen (and expected) for rho-independent genes (Figure [2]A; Materials and Methods). These results demonstrate that in steady-state, the RNA 3′-ends of Rho-dependent transcripts are not dispersed but are rather locally precise, similarly to rho-independent termini.

While the sequence and RNA-structure determinants of intrinsic terminators are well understood, it is still unknown whether termini of Rho-dependent transcripts share any unifying characteristics ([8],[14]). To assess the effectiveness of our classification according to rho-dependence, we analyzed the sequence motifs and predicted RNA structure surrounding the dominant 3′ ends of rho-independent and dependent genes. For each group, we calculated the nucleotide usage frequency at specific positions relative to the 3′ end and compared their relative enrichment compared to randomly selected intergenic regions (Materials and Methods). In agreement with the knowledge on the mechanism of intrinsic termination we found that 99% (172/174) of the rho-independent 3′ termini mapped in our analysis occur immediately downstream of stable hairpin structure, most of which are followed by a uridine rich sequence (Figure [2]B and C; Supplementary Table S2). Surprisingly, we find that all but one of the 3′ termini of Rho-dependent transcripts, including those localized within protein-coding regions of downstream genes, were also associated with stem-loop structures, which presented nearly identical stem and loop
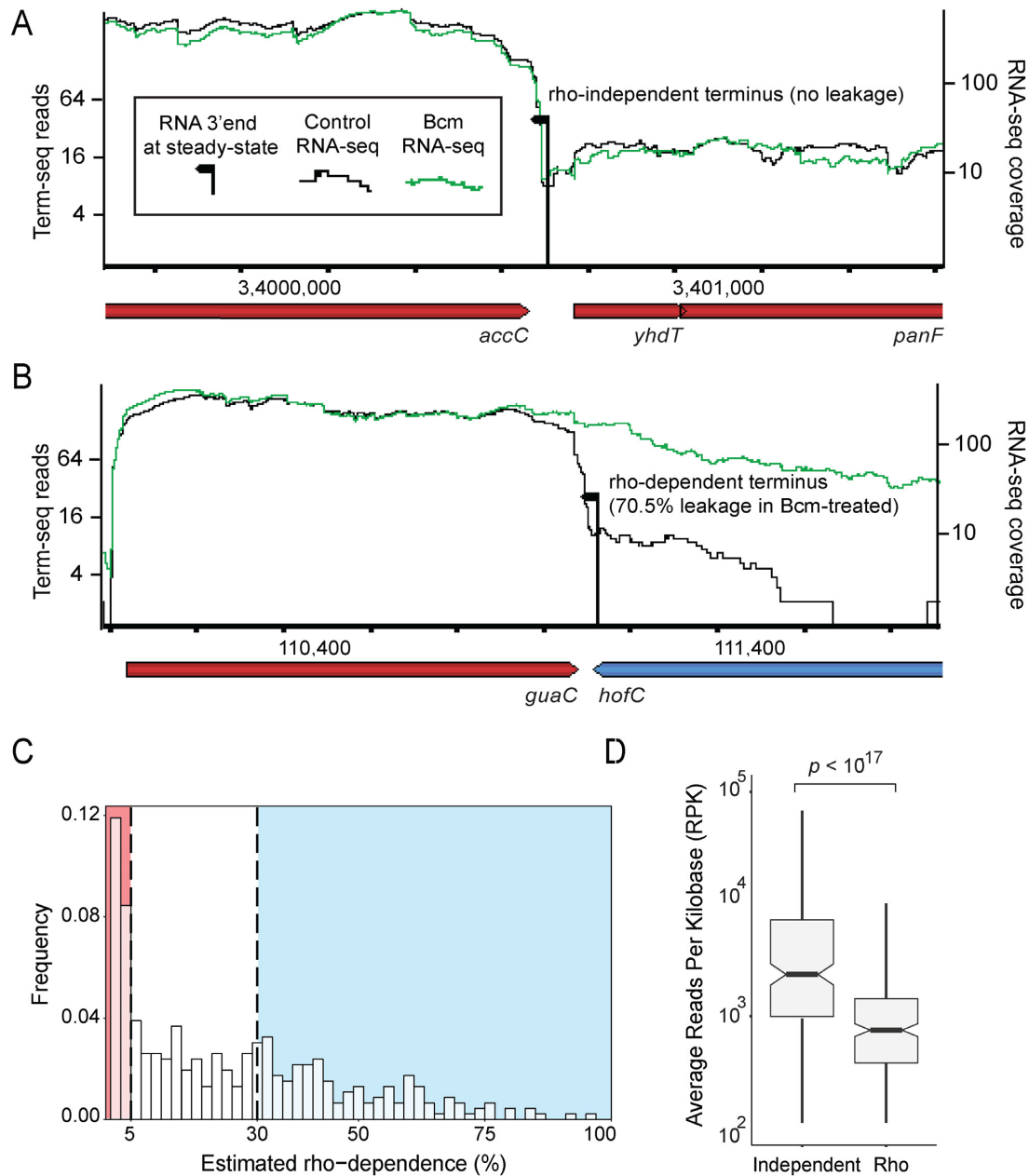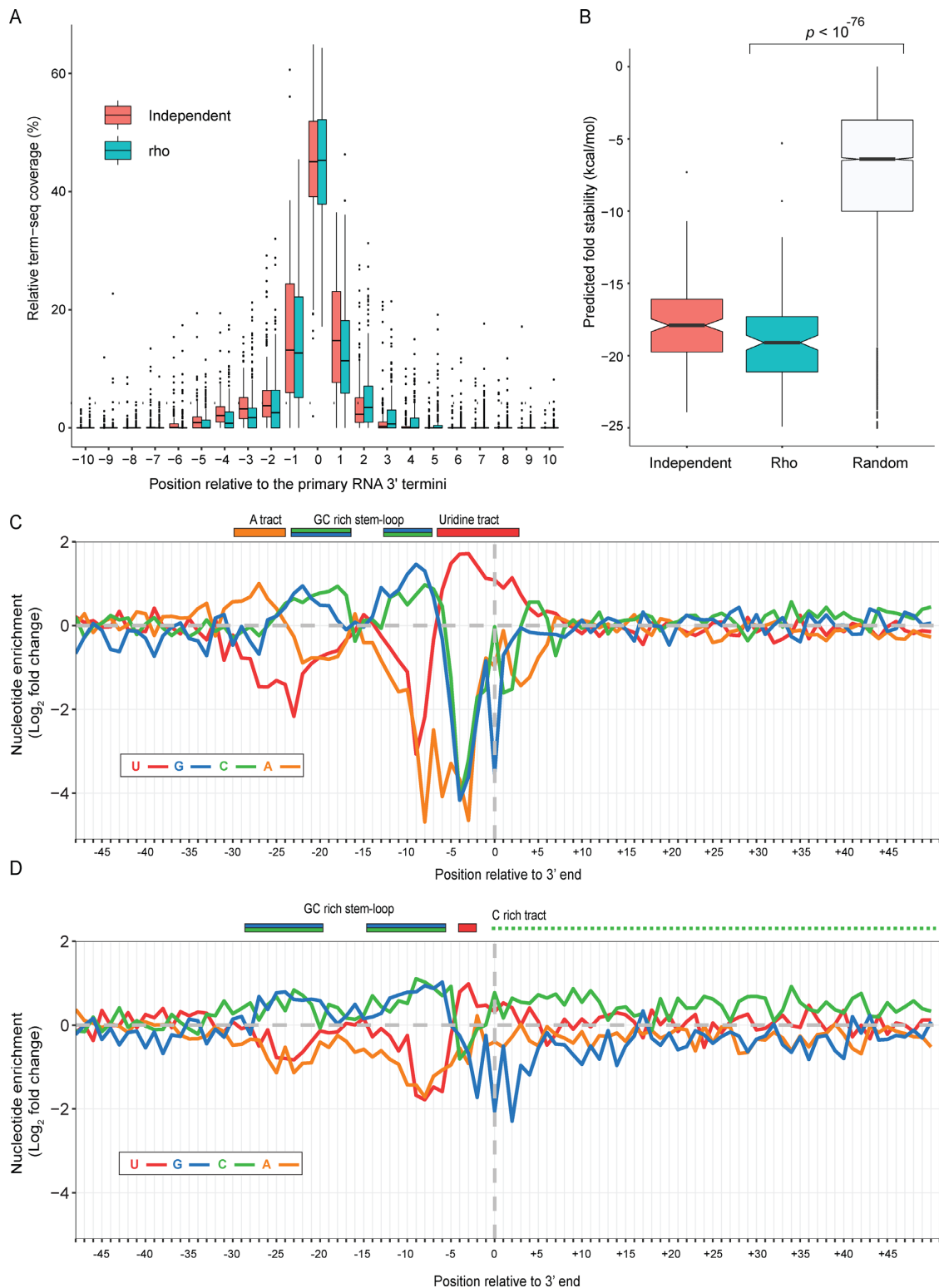
Figure 1. Mapping of Rho-dependent and independent RNA 3′ termini. Inhibition of the Rho termination factor using Bcm identifies Rho-dependent transcript elongation (read-through) beyond the major 3′ terminus. Dominant 3′-ends recorded with term-seq are depicted as black arrows and their height represents the average number of supporting reads for the position, as calculated from three replicates. RNA-seq coverage taken from *Peters et al.* (21) representing the control (black line) and Bcm treated (green) samples were normalized by the number of uniquely mapped reads in each library. Genes are shown below as red (sense strand direction) and blue (antisense strand) box arrows. (A and B) Shown are examples of rho-independent (**A**) and dependent (**B**) genes along with their recorded rho-dependence metric. (**C**) Distribution of Rho-dependent 'read-through' levels in Bcm-treated samples. The classification thresholds are shown by vertical dashed lines. (**D**) Average expression differences between Rho-dependent and independent genes. The distributions were compared using a two-sided Wilcoxon rank-sum test ($P < 10^{-17}$).

length distributions and folding energies as intrinsic terminators (Figure 2B and D; Supplementary Table S2; Materials and Methods). However, these termini had substantially reduced uridine content and were not associated with the adenosine rich tracts commonly detected upstream of the stem-loops of intrinsic terminators (8,28) (Figure 2C–D). Instead, RNA termini of Rho-dependent transcripts were

followed by a mild enrichment of cytosine-over-guanosine usage in the downstream DNA sequence (Figure 2D).

3′ terminal stem-loop structures stabilize Rho-dependent mR-NAs

The genome-wide analysis of Rho-dependent transcripts by *Peters et al.* (21) showed that many estimated termination positions occurred within a distance of a few hundred

Figure 2. Determinants of RNA 3′ termini of Rho-dependent and independent transcripts. (**A**) Term-seq mapped 3′-ends are tightly distributed around the dominant 3′ termini position in both rho-independent (red boxes) and dependent (blue boxes) transcripts. The relative term-seq signal is calculated for the 20 nucleotide positions adjacent to the dominant terminator site (10nt on each side). Outliers are shown as black dots. (**B**) The dominant termini of both rho-independent and dependent transcripts occur immediately downstream of energetically stable stem-loop structures compared with random genomic sequences ($n = 10,000$). The distributions were compared using a two-sided Wilcoxon rank-sum test ($P < 10^{-76}$). Nucleotide frequency meta-analysis across all rho-independent (**C**) or dependent (**D**) dominant 3′ termini was measured by comparing the genomic sequences surrounding the termini sets to randomly selected intergenic positions ($n = 10,000$).

nucleotides of repetitive extragenic palindromic (REP) elements (21). Such REP elements commonly occur in the intergenic regions of *E. coli* and when transcribed, fold into stable secondary structures that resist exonuclease digestion (29–31). Peters and colleagues therefore suggested that Rho-dependent termination occurs far downstream of REP elements, which then act to protect Rho-dependent transcripts from premature transcript degradation, similarly to the RNA structures of rho-independent terminators (21). This hypothesis is also supported by the detection of higher C/G utilization ratios downstream of the REP elements, implying that the rut sequences, and therefore termination, occur downstream of the REPs (21). However, less than half of all of the termination sites detected by *Peters et al.* were found to be associated with REP elements. In addition, REP elements were also shown to sometime play direct roles in Rho-dependent termination (32). Our high-resolution 3′-end mapping now suggests that most Rho-dependent transcripts maintain energetically stable stem-loop structures that define the transcript terminus at steady state. However, only 24% (35/144) of these were associated with annotated REP elements, compared with 13% (22/174) of the intrinsic terminators ($P = 0.008$, Fisher's exact test; Materials and Methods).

Under the protection hypothesis, Rho-dependent termination should take place far downstream of the detected 3′ stem-loop structures, generating a long 3′ UTR tail which is then rapidly trimmed by 3′-5′ exonucleases up until the protective stem-loop (21). We reasoned that in this scenario, perturbing 3′-5′ exonuclease activity in the cell should result in stabilization of longer 3′ UTRs downstream of these stem-loops, revealing a signal that is difficult to detect in the WT steady state conditions. We therefore sequenced the mRNAs of mutant *E. coli* strains deleted for each of the three major 3′-5′ exonucleases: Polynucleotide Phosphorylase (PNPase), RNase II or RNase R (33) (Materials and Methods). While these exonucleases work on a redundant set of RNA substrates, double or triple deletion mutants are non-viable (34).

We compared the RNA-seq coverage downstream of the dominant termini between the WT and exonuclease-deficient strains, allowing us to estimate the relative amount of transcript that is trimmed under each genetic background, which we defined here as 'exonuclease activity' (Figure 3A and B; Supplementary Table S3; Materials and Methods). We find that exonuclease dependent 3′-to-5′ trimming is significantly enriched downstream of Rho-dependent 3′-ends as compared to intrinsically terminated transcripts (Figure 3; Materials and Methods). These decay patterns were mainly dependent on PNPase (*pnp*) and RNase II (*rnb*) but not on RNase R (*rnr*) (Figure 3B). Termini trimmed by RNase II were generally also trimmed by PNPase (Figure 3C). These results support the model in which Rho-dependent transcripts terminate in a 'diffuse' manner downstream of the stem-loop structure and are then trimmed in a 3′-5′ direction, leaving stable RNA 3′ stem-loop structures to stabilize the mRNA in a manner similar to transcript terminated by rho-independent terminators.

DISCUSSION

In this study we mapped the positions of exposed RNA 3′-ends associated with Rho-dependent or independent transcripts in *E. coli*. This analysis was based on a combination of term-seq and RNA-seq experiments in control or Bcm treated bacteria, in which Rho is inhibited (21). We found that mRNAs that depend on Rho for their termination have locally precise termini that are directly downstream to an energetically stable stem-loop structure, essentially identical to that found in rho-independent terminators (but without the U-rich tract). While these structures could potentially hold direct roles in termination (35), as occurs in independent terminators, our data support a model in which they mainly serve as protective elements that resist 3′-5′ exonuclease digestion, validating the hypothesis put forth by Peters and colleagues (21).

Our experiments with exonuclease deficient strains strongly support the notion that termination occurs downstream of the measured termini and that the transcript is rapidly trimmed by housekeeping exonucleases such that in steady-state, the vast majority of mRNA isoforms have the processed, stable 3′ end.

The observation that almost all *E. coli* genes are protected by stem-loop RNA structures provides an improved understanding of the 'rules' for productive gene encoding and expression in bacteria. Moreover, the fact that these structures are energetically indistinguishable from those found in intrinsic terminators implies that terminators could easily switch between Rho-dependent and independent during evolution, simply by altering the uridine content downstream the stem-loop. Furthermore, the reduced and confined uridine enrichment downstream of stem-loop structures defining the steady-state termini of Rho-dependent transcripts could be indicative of cases where the gene is terminated by inefficient intrinsic mechanisms (because of lower uridine content) but is also terminated by Rho to achieve efficient control of expression. Such cases might be indicative of transition states between termination mechanisms.

Terminator evolution is also potentially interesting in light of our discovery that rho-dependence is anti-correlated with gene expression levels (Figure 1D). It stands to reason that Rho-dependent termination would be, on average, a less efficient mechanism than intrinsic termination, which naturally would be less robust to perturbations that influence Rho abundance or activity (e.g. Bcm sensitivity). Thus, one possible explanation for the observation that Rho-dependent genes have significantly lower expression is that higher expression of Rho-dependent genes may result in unwanted readthrough into neighboring genes, which might disrupt tightly regulated expression programs or even generate interference via antisense transcription (36). Another potential explanation is that the interaction with Rho also results in premature termination within the coding-region, which could limit the expression from such genes. In support of the later hypothesis, ChIP-seq experiments have shown that Rho can be detected across the protein-coding regions of its target mRNAs (37).

Our study provides insights into the characteristics of 3′ termini of Rho-dependent transcripts and suggests that es-
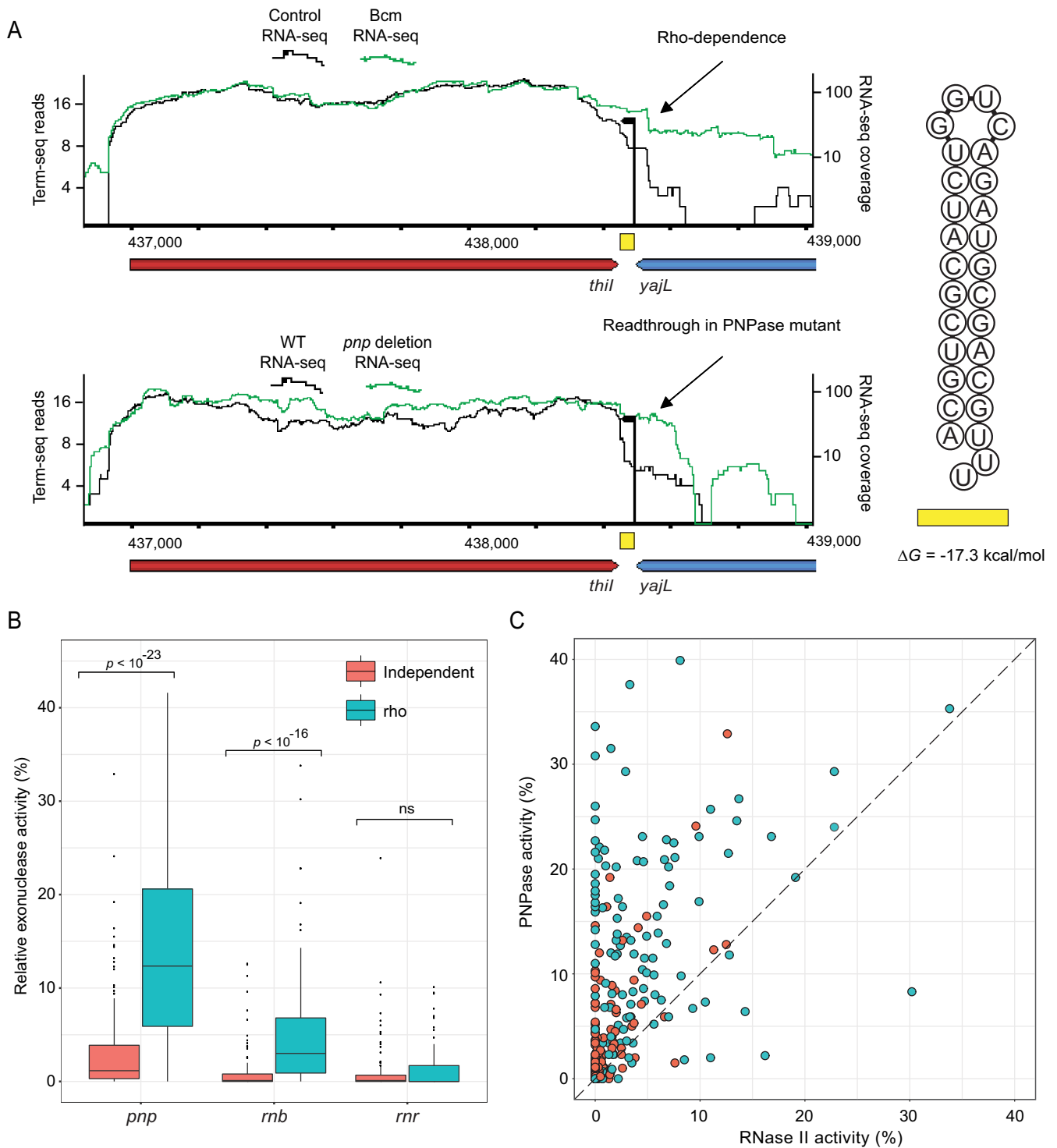
Figure 3. Stem-loop mediated protection from 3′-5′ exonucleases in Rho-dependent mRNAs. Exonuclease activity at specific loci is estimated by relative transcript elongation in exonuclease deletion strains (*pnp*, *rnb* and *rnr*). (**A**) An example of a terminus of a Rho-dependent transcript. RNA-seq coverage data is taken from *Peters et al.* (21) representing the control (black line) and Bcm treated (green) samples. The dominant 3′ terminus is shown as a black arrow and the RNA structure associated with the terminus is shown on the right as predicted by RNAfold (20). Substantial PNPase (*pnp*) dependent transcript read-through at the same locus indicates that the 3′ end generated by Rho-dependent termination is trimmed by PNPase until reaching the stem-loop structure. (**B**) Comparison of the specific exonuclease activities at rho-independent (red; *n* = 154) and dependent (blue; *n* = 123) transcripts covered by sufficient RNA-seq reads (Materials and Methods). The distributions were compared using a two-sided Wilcoxon rank-sum test and the *P*-values are shown above. (**C**) Scatter plot showing *pnp*-dependent versus *rnb*-dependent exonuclease activities. Blue and red circles represent Rho-dependent and independent genes, respectively.

sentially all protein-coding genes in *E. coli* must maintain energetically stable RNA structures at their termini to remain protected from 3′-5′ exonucleases. The comprehensive mapping of transcript 3′ ends in *E. coli* provides a new layer of information that expands the understanding of gene expression in one of the most studied model bacteria.

DATA AVAILABILITY

The sequencing data has been deposited at Gene Expression Omnibus (GEO) under the GSE109766 accession code. An interactive online web browser containing the data used in this study is available at http://www.weizmann.ac.il/molgen/Sorek/ecoli_termseq/.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Madigan,M.T., Martinko,J.M., Bender,K.S., Buckley,D.H., Stahl,D.A. and Brock,T. (2014) *Brock Biology of Microorganisms.* 14th edn. Pearson.
2. Garst,A.D., Edwards,A.L. and Batey,R.T. (2011) Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.*, **3**, a003533.
3. Serganov,A. and Nudler,E. (2013) A decade of riboswitches. *Cell*, **152**, 17–24.
4. Roberts,J.W. (1969) Termination factor for RNA synthesis. *Nature*, **224**, 1168–1174.
5. Breaker,R.R. (2011) Prospects for riboswitch discovery and analysis. *Mol. Cell*, **43**, 867–879.
6. Peters,J.M., Vangeloff,A.D. and Landick,R. (2011) Bacterial transcription terminators: the RNA 3′-end chronicles. *J. Mol. Biol.*, **412**, 793–813.
7. d'Aubenton Carafa,Y., Brody,E. and Thermes,C. (1990) Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *J. Mol. Biol.*, **216**, 835–858.
8. Ray-Soni,A., Bellecourt,M.J. and Landick,R. (2016) Mechanisms of bacterial transcription termination: all good things must end. *Annu. Rev. Biochem.*, **85**, 319–347.
9. Gusarov,I. and Nudler,E. (1999) The mechanism of intrinsic transcription termination. *Mol. Cell*, **3**, 495–504.
10. Gardner,P.P., Barquist,L., Bateman,A., Nawrocki,E.P. and Weinberg,Z. (2011) RNIE: genome-wide prediction of bacterial intrinsic terminators. *Nucleic Acids Res.*, **39**, 5845–5852.
11. Epshtein,V., Cardinale,C.J., Ruckenstein,A.E., Borukhov,S. and Nudler,E. (2007) An allosteric path to transcription termination. *Mol. Cell*, **28**, 991–1001.
12. Hui,M.P., Foley,P.L. and Belasco,J.G. (2014) Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet.*, **48**, 537–559.
13. Lowery-Goldhammer,C. and Richardson,J.P. (1974) An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor. *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 2003–2007.
14. Ciampi,M.S. (2006) Rho-dependent terminators and transcription termination. *Microbiology*, **152**, 2515–2528.
15. Chen,C.Y. and Richardson,J.P. (1987) Sequence elements essential for rho-dependent transcription termination at lambda tR1. *J. Biol. Chem.*, **262**, 11292–11299.
16. Zalatan,F. and Platt,T. (1992) Effects of decreased cytosine content on rho interaction with the rho-dependent terminator trp t' in Escherichia coli. *J. Biol. Chem.*, **267**, 19082–19088.
17. Jin,D.J., Burgess,R.R., Richardson,J.P. and Gross,C.A. (1992) Termination efficiency at rho-dependent terminators depends on kinetic coupling between RNA polymerase and rho. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 1453–1457.
18. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
19. Dar,D., Prasse,D., Schmitz,R.A. and Sorek,R. (2016) Widespread formation of alternative 3′ UTR isoforms via transcription termination in archaea. *Nat. Microbiol.*, **1**, 16143.
20. Lorenz,R., Bernhart,S.H., Siederdissen,CHZ, Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
21. Peters,J.M., Mooney,R.A., Grass,J.A., Jessen,E.D., Tran,F. and Landick,R. (2012) Rho and NusG suppress pervasive antisense transcription in Escherichia coli. *Genes Dev.*, **26**, 2621–2633.
22. Mondal,S., Yakhnin,A. V, Sebastian,A., Albert,I. and Babitzke,P. (2016) NusA-dependent transcription termination prevents misregulation of global gene expression. *Nat. Microbiol.*, **1**, 15007.
23. Conway,T., Creecy,J.P., Maddox,S.M., Grissom,J.E., Conkle,T.L., Shadid,T.M. and Teramoto,J. (2014) Unprecedented High-Resolution view of bacterial operon architecture revealed by RNA sequencing. *MBio*, **5**, 1–12.
24. Dar,D., Shamir,M., Mellin,J.R., Koutero,M., Stern-Ginossar,N., Cossart,P. and Sorek,R. (2016) Term-seq reveals abundant Ribo-Regulation of antibiotics resistance in bacteria. *Science*, **352**, aad9822.
25. Peters,J.M., Mooney,R.A, Kuan,P.F., Rowland,J.L., Keles,S. and Landick,R. (2009) Rho directs widespread termination of intragenic and stable RNA transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 15406–15411.
26. Cardinale,C.J., Washburn,R.S., Tadigotla,V.R., Brown,L.M., Gottesman,M.E. and Nudler,E. (2008) Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in E. coli. *Science*, **320**, 935–938.
27. Hollands,K., Proshkin,S., Sklyarova,S., Epshtein,V., Mironov,A., Nudler,E. and Groisman,E.A. (2012) Riboswitch control of Rho-dependent transcription termination. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 5376–5381.
28. Chen,Y.-J., Liu,P., Nielsen,A.A., Brophy,J.A., Clancy,K., Peterson,T. and Voigt,C.A. (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods*, **10**, 659–664.
29. Stern,M.J., Ames,G.F.L., Smith,N.H., Robinson,E.C. and Higgins,C.F. (1984) Repetitive extragenic palindromic sequences: a major component of the bacterial genome. *Cell*, **37**, 1015–1026.
30. Newbury,S.F., Smith,N.H., Robinson,E.C., Hiles,I.D. and Higgins,C.F. (1987) Stabilization of translationally active mRNA by prokaryotic REP sequences. *Cell*, **48**, 297–310.
31. Khemici,V. and Carpousis,A.J. (2004) The RNA degradosome and poly(A) polymerase of Escherichia coli are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Mol. Microbiol.*, **51**, 777–790.
32. Espéli,O., Moulin,L. and Boccard,F. (2001) Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J. Mol. Biol.*, **314**, 375–386.

33. Hui,M.P., Foley,P.L. and Belasco,J.G. (2014) Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet.*, **48**, 537–559.

34. Donovan,W.P. and Kushner,S.R. (1983) Amplification of ribonuclease II(mb) activity in Escherichia coli K-12. *Nucleic Acids Res.*, **11**, 265–276.

35. Espéli,O., Moulin,L. and Boccard,F. (2001) Transcription attenuation associated with bacterial repetitive extragenic BIME elements. *J. Mol. Biol.*, **314**, 375–386.

36. Georg,J. and Hess,W.R. (2011) cis-Antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.*, **75**, 286–300.

37. Mooney,R.A., Davis,S.E., Peters,J.M., Rowland,J.L., Ansari,A.Z. and Landick,R. (2009) Regulator trafficking on bacterial transcription units in vivo. *Mol. Cell*, **33**, 97–108.