

Overlapping transcription and bacterial RNA removal

Iñigo Lasa¹ and Maite Villanueva

Laboratory of Microbial Biofilms, Idab-Consejo Superior de Investigaciones Científicas-Universidad Pública de Navarra-Gobierno de Navarra and Departamento Producción Agraria, 31006 Pamplona, Spain

The precise understanding of the biology of a living cell requires the identification and quantification of the molecular components necessary to sustain life. One such element is RNA. Two independent high-throughput strategies are available to identify the entire collection of RNA molecules produced by a cell population, which is currently known as the transcriptome. One technique relies on microarray technology (tiling arrays), whereas the second one relies on sequencing the RNA pool (RNA-seq) (1). Both techniques offer the advantage that the identification of the RNA content is not biased by protein-based genome annotation. The application of these methods to the transcriptome analysis in bacteria has uncovered the

existence of a large amount of RNA molecules that overlap at least in some portion with protein-encoding RNA transcripts, generating perfect sense/antisense RNA duplexes (2–5). However, because transcriptome studies have been performed using microgram amounts of RNA purified from millions of bacterial cells instead of RNA purified from a single bacterium, the presence of overlapping sense/antisense RNAs from a genomic region does not necessarily mean that both sense and antisense transcripts are simultaneously present in the same bacteria. Hence, it might be possible that a subgroup in the bacterial population synthesized the sense transcript, another subgroup synthesized the antisense tran-

script, and consequently overlapping transcripts would never be together in the same cell. A report in PNAS by Lybecker et al. (6) provides clear evidences that both sense and antisense transcripts can be present simultaneously within the same bacterial cell. Using a monoclonal antibody that recognizes double-stranded RNA molecules (dsRNA) irrespectively of the nucleotide sequence, the authors perform immunoprecipitation assays to pull down dsRNA molecules (IP-dsRNA) from a total RNA sample extracted from *Escherichia coli*, and identified the purified dsRNA by RNA-seq.

Previous studies have identified examples of at least four different mechanisms to generate dsRNA duplexes in bacteria (2, 7): (i) the presence of bona fide noncoding antisense RNAs (asRNA); (ii) overlapping in the 5' region of mRNAs from contiguous genes that are transcribed in divergent directions; (iii) overlapping in the 3' regions of mRNAs from contiguous genes transcribed in convergent directions; and finally, (iv) genes that, being located in the middle of an operon, are transcribed in the opposite direction to the rest of the operon. According to these mechanisms, the entire mRNA molecule seems to be susceptible to be targeted by an overlapping transcript. However, with only a handful of transcriptomes available so far, it is too early to establish whether overlapping transcription preferentially locates in a specific region of the mRNA relative to the ORF. In this respect, Lybecker et al. find that the majority of IP-dsRNAs correspond to the 5' region of genes (50%), whereas only 0.5% of the IP-dsRNAs correspond to overlapping transcripts that affect the 3' region (6). The most common scenario is overlapping between long 5' UTRs of divergently transcribed genes, followed by overlapping caused by asRNAs transcribed opposite to the 5'/intergenic ends. This description of such a strong bias of dsRNAs toward the 5' region is unique. Whether differences in the size of overlapping regions might determine a more efficient

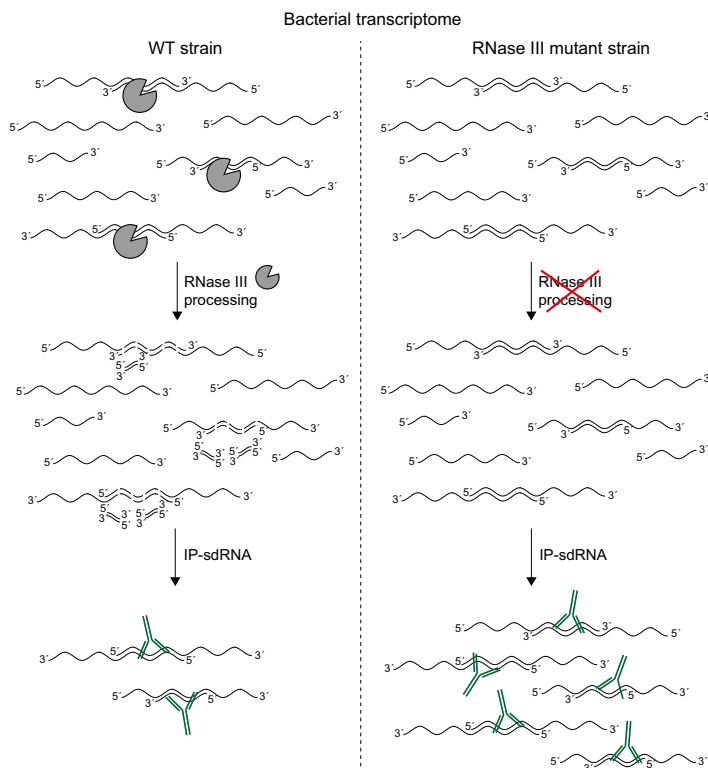


Fig. 1. Schematic description of the IP-dsRNA from a total RNA sample extracted from wild-type and RNase III mutant strains. RNase III-mediated digestion of overlapping transcripts reduces the amount of dsRNA in the bacterial transcriptome.

Author contributions: I.L. and M.V. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 3134.

¹To whom correspondence should be addressed. E-mail: ilasa@unavarra.es.

immunoprecipitation of dsRNA molecules at the 5' region cannot be excluded.

Pairing of overlapping transcripts provides dsRNA substrates that can be digested by specific RNases. RNase III is a dsRNA endoribonuclease, primarily known for its roles in rRNA maturation, mRNA degradation, and sRNA processing (8). First evidence that RNase III plays a critical role in the digestion of overlapping transcripts was obtained in a recent study devoted to analyzing the transcriptome of the human pathogen *Staphylococcus aureus* (5). This study found that RNase III digests overlapping transcripts producing a collection of short RNA fragments (20 nucleotides on average), providing the first evidence that both sense and antisense overlapping transcripts are present simultaneously in the same cell. Other evidence supporting a genome-wide role of RNase III in processing RNA molecules came from an independent study combining coimmunoprecipitation of a catalytically inactive version of RNase III enzyme with deep RNA-seq (9). This study revealed that RNase III is bound to many different antisense RNAs that cover 44% of annotated genes, including noncoding RNAs. In contrast, attempts to identify the collection of short RNA products generated by RNase III-mediated digestion of overlapping RNAs in *Salmonella* were unsuccessful, suggesting that either overlapping transcripts might be processed by a different mechanism or that the resulting short RNA molecules might be unstable in Gram-negative bacteria. The study of Lybecker et al. (6) shows that, indeed, RNase III mediates the digestion of overlapping transcripts in *E. coli*, indicating that this mechanism is well-conserved in bacteria. However, this study does not clarify which are the end products of RNase III activity because the antibody used is unable to bind dsRNAs shorter than 40 base pairs. Intuitively, if RNase III digests overlapping transcripts, one would expect that mutants in RNase III would accumulate higher levels of both sense and antisense transcripts. In agreement with this hypothesis, Lybecker et al. found that the dsRNA regions are stable and more abundant than the single-stranded regions of the same transcripts in the absence of an active RNase III (Fig. 1). The simplest explanation is that dsRNA regions remain protected, whereas single-stranded regions are degraded by the activity of other RNases.

Overlapping transcription can affect the expression of its complementary gene at different levels, including transcription, mRNA stability, or translation (10, 11). In this respect, Cossart's group has proposed a new paradigm of regulation based on overlapping transcription, termed "excludon" (12, 13). The exclu-

don concept describes the process by which the expression of a long mRNA transcript results in the repression of the expression of the overlapping transcript produced from the neighbor gene. The final consequence is that expression of both overlapping transcripts is mutually exclusive. However, the exact mechanism underlying the inhibitory effect of the overlapped transcripts in the excludon has not yet been determined. Based on the observations that RNase III mediates the digestion of overlapping transcripts in *S. aureus*, it was suggested that the selective degradation of the dsRNA that results from hybridization of overlapping sense and antisense transcripts could be a likely mechanism to explain the excludon paradigm (13). The Lybecker et al. (6) study strongly supports this idea.

An interesting question regarding the RNase III processing of overlapping transcripts is whether the digestion occurs before or after mRNA translation. In general, it is assumed that the transcription and translation processes are coupled in bacteria. In this scenario, RNase III would be processing RNA molecules that had already been translated, and pairing between overlapping transcripts would represent another mechanism of RNA decay. Alternatively, if the digestion of RNAs by RNase III occurs before translation, this mechanism would provide an additional level of posttranscriptional regulation to adjust mRNA levels and to remove any transcript produced because of leaky transcription initiation. The former point is very important because uncontrolled transcription might be toxic if all mRNAs were translated at the same time. Depending on the expression levels of each overlapping transcript, Lybecker et al. (6) identified two classes of dsRNAs. In class I, both overlapping RNAs exhibit different expression levels, whereas in class II both transcripts have similar expression levels. Complementary proteomic studies will be necessary to deter-

mine whether the transcripts of class I that are produced at lower levels are indeed translated to proteins. On the other hand, it is important to highlight that, coinciding with previous observations by Lioliou et al. (9), a significant number of antisense RNAs transcribe opposite to noncoding regulatory RNAs, indicating that the RNase III-mediated digestion of overlapping transcripts may impact the functionality of the RNA molecules regardless of the protein translation process.

The Lybecker et al. (6) study provides a method to examine the genome-wide process of overlapping transcription in bacteria. This technology has provided a tool to demonstrate that sense and antisense transcripts exist simultaneously in the cytoplasm of *E. coli*. Identical conclusions were previously obtained in *S. aureus* using a completely different strategy, indicating that degradation of overlapping transcripts by RNase III is a highly conserved process in bacteria (5). Of course, many intriguing questions about the process remain. What are the end products of the RNase III digestion process in *E. coli*? Are these end products functional molecules with a specific role in gene regulation? When does the pairing between overlapping transcripts occur, before or after translation? What are the kinetics of the RNase III processing reaction? Are there specific proteins governing the pairing between overlapping transcripts? Which of the phenotypes associated to RNase III deficiency are because of the accumulation and translation of sense and antisense transcripts in the same cell? The initial skepticism about the biological relevance of the genome-wide overlapping transcription has been followed by the discovery that bacteria have a simple and efficient mechanism to remove the double-stranded sense/antisense pairs. Because this process can alter the levels of functional RNAs, the time has come to include overlapping transcription as another player of bacterial gene regulation.

- 1 Sorek R, Cossart P (2010) Prokaryotic transcriptomics: A new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 11(1):9–16.
- 2 Toledo-Arana A, et al. (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459(7249):950–956.
- 3 Sharma CM, et al. (2010) The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464(7286):250–255.
- 4 Georg J, et al. (2009) Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol Syst Biol* 5:305.
- 5 Lasa I, et al. (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci USA* 108(50):20172–20177.
- 6 Lybecker M, Zimmermann B, Bilusic I, Tukhtubaeva N, Schroeder R (2014) The double-stranded transcriptome of *Escherichia coli*. *Proc Natl Acad Sci USA* 111:3134–3139.
- 7 Lasa I, Toledo-Arana A, Gingeras TR (2012) An effort to make sense of antisense transcription in bacteria. *RNA Biol* 9(8):1039–1044.

- 8 Arraiano CM, et al. (2010) The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev* 34(5):883–923.
- 9 Lioliou E, et al. (2012) Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* 8(6):e1002782.
- 10 Thomason MK, Storz G (2010) Bacterial antisense RNAs: How many are there, and what are they doing? *Annu Rev Genet* 44:167–188.
- 11 Georg J, Hess WR (2011) *cis*-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev* 75(2):286–300.
- 12 Wurtzel O, et al. (2012) Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol Syst Biol* 8:583.
- 13 Sesto N, Wurtzel O, Archambaud C, Sorek R, Cossart P (2013) The excludon: A new concept in bacterial antisense RNA-mediated gene regulation. *Nat Rev Microbiol* 11(2):75–82.