

Stem-loops direct precise processing of 3' UTR-derived small RNA MicL

Taylor B. Updegrove^{1,†}, Andrew B. Kouse^{1,†}, Katarzyna J. Bandyra² and Gisela Storz^{1,*}

¹Division of Molecular and Cellular Biology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD 20892-5430, USA and ²Department of Biochemistry, University of Cambridge, Tennis Court road, Cambridge CB2 1GA, UK

Received June 21, 2018; Revised October 29, 2018; Editorial Decision October 30, 2018; Accepted November 05, 2018

ABSTRACT

Increasing numbers of 3'UTR-derived small, regulatory RNAs (sRNAs) are being discovered in bacteria, most generated by cleavage from longer transcripts. The enzyme required for these cleavages has been reported to be RNase E, the major endoribonuclease in enterica bacteria. Previous studies investigating RNase E have come to a range of different conclusions regarding the determinants for RNase E processing. To better understand the sequence and structure determinants for the precise processing of a 3' UTR-derived sRNA, we examined the cleavage of multiple mutant and chimeric derivatives of the 3' UTR-derived MicL sRNA *in vivo* and *in vitro*. Our results revealed that tandem stem-loops 3' to the cleavage site define optimal, correctly-positioned cleavage of MicL and probably other sRNAs. Moreover, our assays of MicL, ArcZ and CpxQ showed that sRNAs exhibit differential sensitivity to RNase E, likely a consequence of a hierarchy of sRNA features recognized by the endonuclease.

INTRODUCTION

While bacterial small RNAs (sRNAs) that act by limited base pairing to increase or repress synthesis from target mRNAs initially were found to be encoded as independent genes in intergenic regions, more and more base pairing sRNAs derived from the 3'UTRs (untranslated regions) of mRNAs are being discovered (1–5). Some of these sRNAs, such as *Salmonella enterica* CpxQ, arise from the cleavage of the longer mRNA transcript, while others, such as *Escherichia coli* MicL, are transcribed from promoters internal to the protein coding sequence. Even the sRNAs transcribed from the internal promoters can be cleaved to give rise to a shorter product, which, in the example of MicL, contains the region for target base pairing (4).

The cleavage of the 3'UTR-derived sRNAs characterized thus far has been observed to be well defined, often at or very near the stop codon of the corresponding upstream gene, raising the question of how this specific cleavage occurs. For the *S. enterica* CpxQ sRNA, which is cleaved from the *cpxP* mRNA, the cleavage was found to be dependent on the conserved endoribonuclease RNase E (1). In addition, recent RNA-Seq analysis comparing the 5' ends of transcripts with and without inactivation of a temperature-sensitive RNase E mutant in *S. enterica* indicated that the majority of 3'UTR-derived sRNAs are generated by RNase E (6).

Tetrameric RNase E (encoded by *rne*), the major endoribonuclease in enteric bacteria, forms the core of the degradosome (reviewed in (7,8)). Given the central role of RNase E and the degradosome in the processing of mRNA, rRNA, tRNA and sRNA (reviewed in (9)), the determinants of RNase E-dependent cleavage have been studied extensively. Multiple studies have shown that the site of cleavage generally is single stranded and AU-rich (6,10). Initial characterization of the RNase E cleavage site in the bacteriophage T4 mRNA led to a proposed recognition sequence of (G,A)AUU(U,A) (11). *In vitro* assays examining the cleavage of poly(A) or poly(U) oligonucleotides with substitutions at specific positions further showed that while A- or U-rich sequences are uniformly cleaved, specific nucleotides near the cleavage site impact the position of cleavage (12,13). The recent genome-wide analysis of RNase E cleavage sites in *S. enterica* further led to the proposal of a RN₁WUU core motif (R = A or G, N = any nucleotide and W = A or U), in which the location of a uridine residue two nucleotides downstream of the cleavage site is most critical (6).

Considering the degenerate nature of the proposed cleavage motifs, it is not surprising that additional factors have been proposed to impact RNase E-dependent cleavage including the status of the 5' nucleotide, secondary structure and proteins bound to the RNA. A number of studies have shown that a 5' monophosphate stimulates RNase

*To whom correspondence should be addressed. Tel: +1 301 402 0968; Fax: +1 301 480 0975; Email: storzg@mail.nih.gov

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Present address: Taylor B. Updegrove, Laboratory of Molecular Biology, National Cancer Institute, Bethesda, MD 20892 USA.

E-dependent cleavage at sites near this end (14). However, transcriptome-wide comparisons of total RNA in the presence and absence of the RppH RNA pyrophosphohydrolase, which generates the 5' monophosphate ends showed that a large percentage of cleavage sites are not impacted by the 5' end but rather are 'direct entry' sites for RNase E (15). This study also suggested that the presence of multiple single stranded regions enhanced direct entry by RNase E. Another genome-wide characterization of the secondary structure of the *E. coli* transcriptome via parallel analysis of RNA structure (PARS) coupled to deep sequencing revealed sequences 5' to 1800 known RNase E cleavage sites were significantly structured (10). For at least one RNA, the structure was shown to be critical for RNase E cleavage (11). Finally, the binding of proteins can impact the position and extent of cleavage both positively and negatively. This is illustrated by RNase E cleavage of the GlmZ sRNA, which is blocked by the binding of the Hfq RNA chaperone protein but is stimulated by the binding of the RapZ protein (16). In this example, it was found that RNase E cleaves at a site 6 or 7 nucleotides downstream of the stem-loop RapZ binding site, and conversion of the single stranded region from AU rich to GC rich did not prevent RNase E cleavage, while removal of the upstream stem-loop binding site did.

In contrast to the *S. enterica* CpxQ sRNA, the levels of the truncated *E. coli* MicL sRNA were not strongly reduced in a temperature-sensitive *rne* mutant strain, though levels of the full length transcript increased (4). Given the prevalence of 3'UTR-derived sRNAs, the precise cleavage observed, and our previous, ambiguous results regarding the ribonucleases acting on the 3'UTR-derived MicL sRNA, we set out to define the determinants for MicL cleavage.

MATERIALS AND METHODS

Bacterial strains and plasmids

The bacterial strains and plasmids used in the study are listed in Supplementary Tables S1 and S2, respectively. For plasmid construction, the desired gene fragments were either cloned into the vectors using the Gibson Assembly Cloning Kit (New England Biolabs) or were generated by PCR amplification using MG1655 genomic DNA as a template and, after digestion with restriction enzymes, were cloned into the corresponding sites of the indicated vectors. All oligonucleotides used are listed in Supplementary Table S3. pBR* (pMSG14, (4)) is a derivative of the pBR322-derived pBRplac vector (17) in which the ampicillin cassette was replaced by the kanamycin cassette. All cloning was performed using *E. coli* TOP10 cells (Invitrogen), and all mutations and plasmid inserts were confirmed by sequencing. Sequences of all inserts are given in Supplementary Table S4. Mfold (18) was used to predict the structures of the RNAs expressed from the constructs.

Growth conditions

Unless indicated otherwise, strains were grown aerobically at 37°C in LB (10 g tryptone, 5 g yeast extract, 5 g NaCl/l) to an OD₆₀₀ ~1.0. The following compounds were added at the following final concentrations where indicated; IPTG at 1 mM, ampicillin at 100 µg/ml, kanamycin at 30

µg/ml, chloramphenicol at 25 µg/ml and tetracycline at 12.5 µg/ml.

To deplete RNase E, *E. coli* strain KSL2000, with the pBAD-RNE plasmid (19), was grown overnight at 37°C in LB supplemented with ampicillin, chloramphenicol, tetracycline and 0.1% arabinose. Strains without arabinose failed to grow. The overnight culture (100 µl) was diluted into 30 ml of LB supplemented with ampicillin, chloramphenicol, tetracycline and 0.1% arabinose and grown at 37°C to OD₆₀₀ ~0.5. The culture was then washed 3× with 40 ml of LB media and resuspended in LB with ampicillin, chloramphenicol, tetracycline to an OD₆₀₀ ~0.2. The culture was then split into two 15 ml-cultures with one culture having 0.1% arabinose and the other 0.1% glucose. The two cultures were then grown at 37°C to OD₆₀₀ ~1 and ~1.2, and 1 ml of each culture was harvested and subject to total RNA extraction and northern analysis as described below. As expected, the culture with glucose grew slower than the culture with arabinose.

Northern analysis

For northern analysis, total RNA was extracted with TRIzol Reagent as described (20). Briefly, 1–5 ml cells grown to OD₆₀₀ ~1.0 (or indicated otherwise) were collected and resuspended in 1 ml of TRIzol™ Reagent (Thermo Fisher Scientific) with repetitive pipetting to lyse the cells. The mixture was incubated at room temperature for 5 min, 0.2 ml of chloroform was added, and the sample was vortexed vigorously. The sample was then centrifuged for 15 min at 4°C at 11 000 rpm. The top ~0.6 ml of the aqueous phase was transferred to a new Eppendorf tube and 0.5 ml of isopropyl alcohol added. After 10 min incubation at room temperature the sample was centrifuged at 11 000 rpm for 5 min at 4°C. The precipitated RNA pellet was washed with 1 ml of 75% ethanol and finally air-dried and dissolved in DEPC treated dH₂O. Total RNA concentration was determined based on OD₂₆₀.

Northern blots were performed as described (20). Briefly, 10 µg of total RNA was separated on an 8% polyacrylamide–7M urea gel (USB Corporation) in 1× TBE and transferred to Zeta-Probe membrane (Bio-Rad) overnight at 20 V in 0.5× TBE. Oligonucleotides were end-labeled with γ-³²P-ATP by T4 polynucleotide kinase (New England Biolabs). Membranes were UV cross-linked and hybridized overnight with the labeled probe at 45°C in UltraHyb (Ambion) hybridization buffer. Following hybridization, membranes were washed once with 2× SSC + 0.1% SDS followed by a 10 min incubation at 45°C with 2× SSC + 0.1% SDS. Membranes were subsequently washed 5× with 0.2× SSC + 0.1% SDS, allowed to air dry for 5 min, and exposed to KODAK Biomax X-ray film at –80°C. Where indicated, band intensities were quantitated using ImageJ 1.50i software.

In vitro RNA synthesis

The MicL (308 nt), Δ222 (86 nt), Δ222-stemAΔ2 (72 nt), Δ222+10 (96 nt), Δ222+15 (101 nt), 9S (243 nt) and ArcZ (121 nt) RNAs were synthesized using a MEGAshortscript T7 Transcription kit (Ambion) using manufacturer's guide-

lines. Purified RNA was quantified by OD₂₆₀ measurements. For experiments that called for 9S to be dephosphorylated, 20 pmol T7-transcribed RNA was treated with Calf Intestinal Phosphatase (CIP) (New England Biolabs) using manufacturer's guidelines and then extracted with phenol:chloroform:IAA and ethanol precipitated.

In vitro assay of RNase E activity

The *in vitro* cleavage assays with the purified RNase E derivatives and *in vitro* transcribed RNAs were carried out with slight modifications of a previously reported protocol (1). Briefly, the RNA was diluted in Buffer A (25 mM Tris pH 7.5, 50 mM NaCl, 50 mM KCl, 10 mM MgCl₂, 1 mM DTT and 0.5 U μl⁻¹ RNase Out (Thermo Fisher Scientific)), and heat denatured at 95°C for 1 min and cooled on ice for 5 min. In a 10 μl reaction, 300 nM RNA was incubated in buffer or with 300 nM Hfq₆ for 10 min at 30°C. Buffer or purified RNase E (1-529) or D26N,D28N,D338N mutant RNase E (1-529) or 8× mutant RNase E (1-529) (300 nM) was added, and reactions incubated for an additional 30 min at 30°C. Ammonium acetate stop solution (Ambion) (7.5 μl) was then added along with 82.5 μl dH₂O and 100 μl of phenol:chloroform:IAA. The aqueous phase was separated from the organic phase by PLG-tubes (5PRIME) and the RNA was precipitated from the aqueous phase by the addition of 10 μl of 3M sodium acetate, 250 μl of 100% ethanol and 3 μl of glycogen and incubation at -80°C for 30 min. The RNA was then centrifuged at 15 000 rpm for 30 min at 4°C, and the pellet washed with 200 μl of 70% ethanol and resuspended in 20 μl of Gel Loading Buffer II (Ambion). An aliquot (10 μl) was analyzed via northern analysis as described above.

RESULTS

ArcZ, CpxQ and MicL are differentially sensitive to the *rne-3071* allele

RNase E is capable of cleaving hundreds of RNAs and has been shown to play the most prominent role in sRNA cleavage in *E. coli* (reviewed in (9)). The enzyme is essential though overexpression of RNase G, which can cleave many of the same targets, can rescue an *rne*- strain (19). To determine the impact of RNase E and RNase G on cleavage of MicL, total RNA was isolated from strains carrying combinations of wild-type and temperature-sensitive *rne-3071* (21) and *Δrng* alleles. Strains initially grown at 30°C to OD₆₀₀ ~1.0 were split and cultured for an additional h at either 30°C or 43.5°C, after which total RNA was extracted.

We first assayed the levels of ArcZ and CpxQ, previously reported to be cleaved by RNase E (1,6). As expected, while the levels of the cleavage products were similar for all of the strains grown at 30°C (Figure 1A, lanes 1-4), the levels of the processed transcripts were significantly reduced in the *rne-3071* single mutant while the levels of the longer, un-cleaved transcripts increased at the non-permissive temperature of 43.5°C (Figure 1A, lane 6). The products detected for the *Δrng* single mutant were similar to the products seen for the wild type strain at 43.5°C (Figure 1A, lanes 5 and 7). In contrast, the cleavage products were almost entirely ab-

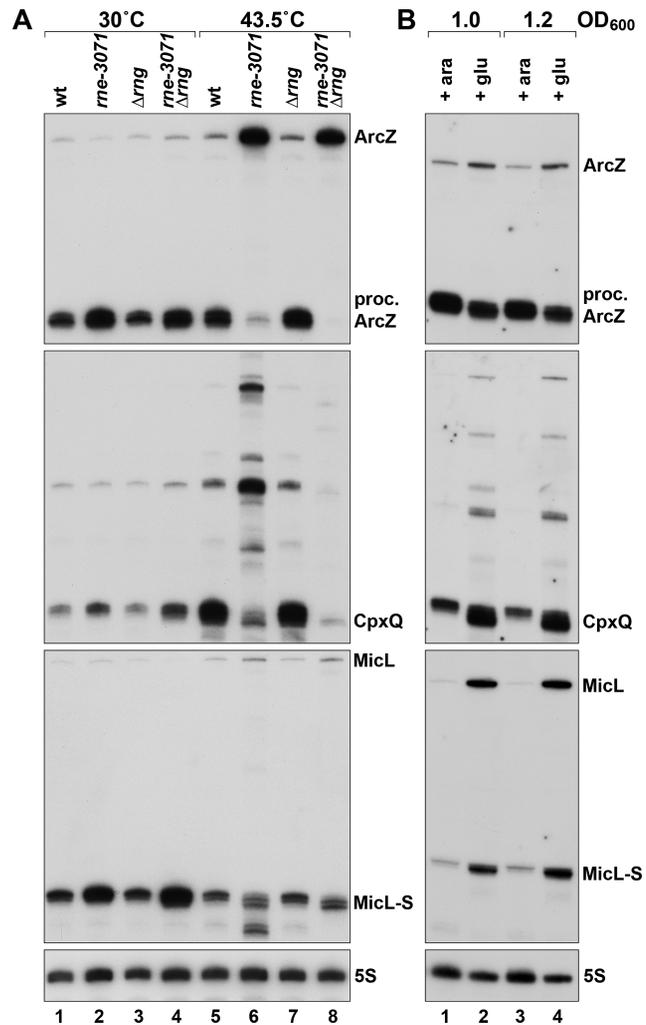


Figure 1. Varied effects of *rne-3071*, *Δrng* and RNase E depletion on ArcZ, CpxQ and MicL cleavage. (A) Four *E. coli* strains, wild type (EM1279), *rne-3071* (EM1277), *Δrng* (EM1279 *Δrng*) and *rne-3071 Δrng* double mutant (EM1277 *Δrng*) were cultured in LB at 30°C to an OD₆₀₀ ~1.0. Cultures were then split in half and incubated at either 30°C or 43.5°C for 1 h. (B) *E. coli* KSL2000, with the pBAD-RNE plasmid was cultured in LB arabinose at 37°C to an OD₆₀₀ ~0.5. Cultures were washed in LB, split and half grown in LB arabinose and the other half grown in LB glucose to OD₆₀₀ ~1.0 or 1.2. For both panels, total RNA extracted from the cultures was subjected to northern analysis using probes against the ArcZ, CpxQ, MicL and 5S RNAs. For panel B, the MicL-S: MicL ratio is 8.1, 0.9, 13.7 and 1.1 for lanes 1, 2, 3 and 4, respectively.

sent in the *rne-3071 Δrng* double mutant (Figure 1A, lane 8).

The effects of the *rne-3071* and *rne-3071 Δrng* mutations were somewhat different for MicL when the same northern membrane was probed for this sRNA (Figure 1A). Similarly, low levels of MicL and high levels of cleaved MicL (MicL-S) transcripts were found across each of the 30°C samples (Figure 1A, lanes 1-4). For the cells shifted to 43.5°C, the overall levels of MicL were higher and MicL-S were lower. Again, the patterns for wild type and *Δrng* strains were similar (Figure 1A, lanes 5 and 7). However, unlike for ArcZ and CpxQ, the levels of cleaved MicL-S were not greatly reduced in the *rne-3071* single and *rne-3071*

Δrng double mutant strains. The levels of MicL increased, and the pattern of MicL-S was changed somewhat with the detection of one slightly shorter transcript in both mutants as well as somewhat shorter RNase G-dependent products in the *rne-3071* mutant (Figure 1A, lanes 6 and 8).

ArcZ, CpxQ and MicL are differentially sensitive to RNase E depletion

We also examined the consequences of depleting wild type RNase E in an *rne*- strain with RNase E expressed from a P_{BAD} promoter on a plasmid (19). A culture of these cells was grown to mid-exponential phase in LB with arabinose, washed, split and the two halves grown in either LB with arabinose, which induces expression from the P_{BAD} promoter, or LB with glucose, which represses expression. Total RNA isolated from the two cultures further grown to either $OD_{600} \sim 1.0$ or ~ 1.2 was again probed for the ArcZ, CpxQ and MicL sRNAs (Figure 1B). Full-length and processed ArcZ as well as CpxQ were present at approximately wild type levels for the cells grown with arabinose to induce RNase E (Figure 1B, lanes 1 and 3). In contrast, the levels of MicL-S were reduced. The shorter form of the sRNA may be further digested by elevated RNase E likely expressed from the plasmid. As expected, depletion of RNase E by growth in glucose led to increased levels of full-length and decreased levels of the processed ArcZ though cleavage was not completely abolished (Figure 1B, lanes 2 and 4). The levels of CpxQ, as well as intermediate cleavage products observed with the *rne-3071* strain, all increased for the glucose-grown cultures (Figure 1B, lanes 2 and 4). MicL and MicL-S levels also were both increased with glucose with a greater increase for full-length MicL (Figure 1B, lanes 2 and 4). Thus, as observed for the *rne-3071* strain, the three sRNAs showed differential sensitivity to altered RNase E levels.

MicL is cleaved by RNase E *in vitro*

To further test whether RNase E is capable of cleaving MicL at the position observed *in vivo*, we carried out *in vitro* cleavage assays with the purified catalytic N-terminal domain (NTD; residues 1–529) of RNase E in the presence and absence of the RNA chaperone Hfq. As shown in Figure 2, when we incubated full length MicL with RNase E (1–529) and Hfq, we observed faint cleavage of MicL at the position where cleavage is seen *in vivo*, possibly due to secondary structure occluding the cleavage site (see below). However, the levels of this product increased when the RNA was incubated with Hfq and a D26N,D28N,D338N mutant RNase E (1–529) that is more catalytically active than wild-type RNase E (1–529) (22).

A truncated version of MicL lacking the first 222 nt, which is processed as efficiently as wild type MicL *in vivo* (see below), is cleaved robustly by both wild type RNase E (1–529) and the hyperactive D26N,D28N,D338N mutant RNase E (1–529) in the presence of Hfq (Figure 2). Together the results of the *in vivo* and *in vitro* assays indicate that MicL is cleaved by RNase E. We suggest that MicL may be a particularly sensitive substrate since it is still processed by low levels of the endonuclease *in vivo*.

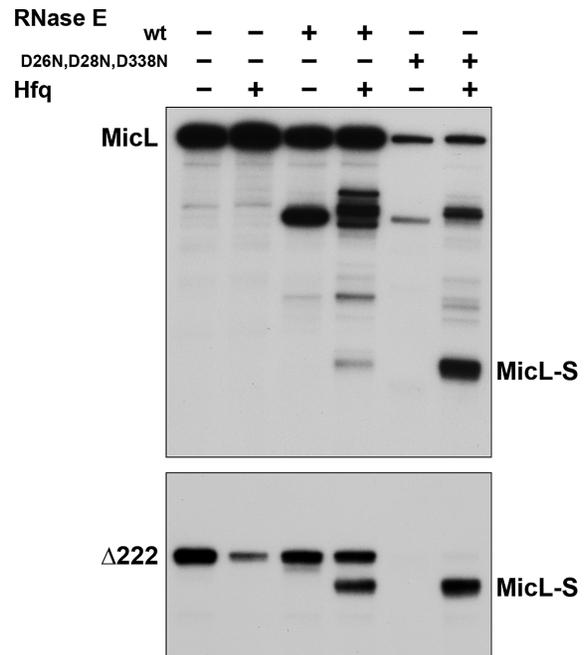


Figure 2. Purified RNase E (1–529) cleaves MicL *in vitro*. *In vitro* transcribed full-length MicL (308 nt) and $\Delta 222$ (86 nt) with 5'PPP was incubated with purified Hfq and purified wild type RNase E (1–529) or D26N,D28N,D338N mutant RNase E (1–529) at 30°C for 30 min. The RNA was then subject to northern analysis using an oligonucleotide probe complementary to the 3'-end of MicL. We suggest the *in vitro* transcribed full-length RNA may fold into some non-native configurations that could contribute to the low levels of cleavage observed for the full-length RNA with wild type RNase E (1–529).

Cleavage of MicL is not affected by sequences at the 5' end

We next wanted to examine what sequences directed the very specific cleavage of MicL to give MicL-S. To determine if sequences 5' of the MicL cleavage site are important, we examined the consequences of one insertion as well as sequential deletions of this region (Figure 3A and Supplementary Figure S1A). The insertion and series of 5' truncations were cloned into the pBR* expression plasmid and introduced into a $\Delta cutC$ strain, which lacks the native *micL* promoter and the region of *micL* overlapping the *cutC* coding sequence (4). Total RNA isolated from these strains was examined by northern analysis using an oligonucleotide probe complementary to the 3' end of MicL (Figure 3B and Supplementary Figure S1B).

All of the constructs still gave rise to the MicL-S cleavage product, though the levels were reduced for +20, $\Delta 70$, $\Delta 72$, $\Delta 82$, $\Delta 130$, $\Delta 132$ and particularly $\Delta 162$. Structure predictions suggested that for the constructs with reduced MicL-S levels, the cleavage site might be occluded by the formation of an alternative stem-loop (Supplementary Figure S1A and C). To test this possibility, we mutated residues 177–180 and 204–208 predicted to be involved in this pairing and observed that cleavage is increased (Supplementary Figure S1D). Overall, given that the same MicL-S cleavage product was observed for all 20 constructs assayed, we concluded that the sequences 5' of the cleavage site are not important for specific cleavage, but that, as has also been

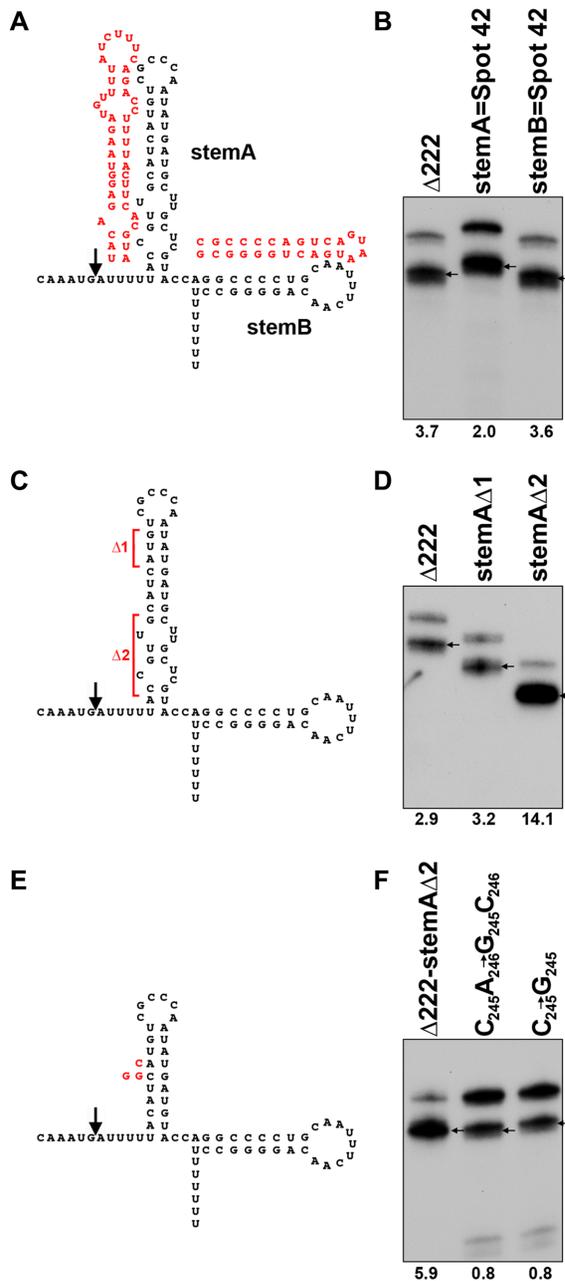


Figure 5. Two stable 3' stem-loops are required for efficient MicL cleavage. (A) Predicted secondary structure of $\Delta 222$ constructs with stem-loop A and stem-loop B replaced by similar stem-loops from Spot 42. The arrow indicates site of RNase E cleavage. (B) The $\Delta 222$ construct and the two $\Delta 222$ derivatives depicted in Figure 5A were expressed from pBR* in $\Delta cutC$ background. Arrows denote cleavage products corresponding to the site indicated in Figure 5A. (C) Predicted secondary structure of $\Delta 222$ constructs with regions of stem A deleted indicated by red brackets. The arrow indicates site of RNase E cleavage. (D) The $\Delta 222$ construct and the two $\Delta 222$ derivatives depicted in 5C were expressed from the pBR* vector in $\Delta cutC$ background. Arrows denote cleavage products corresponding to the site indicated in Figure 5C. (E) Predicted secondary structure of $\Delta 222$ -stemA $\Delta 2$ construct with red nucleotides indicating residues changed in the $C_{245}A_{246} \rightarrow G_{245}C_{246}$ and $C_{245} \rightarrow G_{245}$ constructs. The arrow indicates cleavage site. (F) The $\Delta 222$ -stemA $\Delta 2$ construct and the two $\Delta 222$ -stemA $\Delta 2$ derivatives depicted in Figure 5E were expressed from pBR* in $\Delta cutC$ background. Arrows denote cleavage products corresponding to the site indicated in E. For samples in panels B, D and F, total RNA was extracted and probed for MicL as in Figure 3B. The bottom band:top band ratio is given below each lane.

sRNA, which have a similar size and stability despite significantly different sequences (Figure 5A). Analysis of total RNA isolated from cells expressing these constructs revealed that the heterologous stem-loops did not affect the location or extent of cleavage; the pattern was very similar for $\Delta 222$, stemA = Spot 42 ($\Delta G = -10.8$ kcal/mol) and stemB = Spot 42 ($\Delta G = -21.2$ kcal/mol) (Figure 5B) indicating the sequences of the 3' stem-loops do not strongly impact processing.

To further examine whether stem-loop A was important, we shortened this stem in two mutants. In stemA $\Delta 1$ ($\Delta G = -8.7$ kcal/mol), the stem was shortened by the removal of three base pairs towards the top of the stem-loop and in stemA $\Delta 2$ ($\Delta G = -13.0$ kcal/mol), the bulges in the lower portion of the stem were removed thereby increasing stability (Figure 5C). Again, synthesis from these constructs was examined by northern analysis. While the stemA $\Delta 1$ construct showed cleavage similar to the control $\Delta 222$ construct, the stemA $\Delta 2$ derivative, which is predicted to be more stable than the wild type stemA, gave higher levels of MicL-S (Figure 5D). These data suggest that stem-loop A influences cleavage efficiency, possibly along with contributing to transcript stability.

To further explore the effect of stem A stability on processing, we additionally introduced mutations to destabilize the stem in the context of $\Delta 222$ with stemA $\Delta 2$; $C_{245}A_{246} \rightarrow G_{245}C_{246}$ ($\Delta G = -4.1$ kcal/mol) and $C_{245} \rightarrow G_{245}$ ($\Delta G = -5.8$ kcal/mol) (Figure 5E). The transcripts made from these constructs were then compared to the $\Delta 222$ -stemA $\Delta 2$ ($\Delta G = -13.0$ kcal/mol) RNA. Both destabilized constructs showed a decrease in the cleavage product when compared to $\Delta 222$ -stemA $\Delta 2$, further suggesting that the stability of stem-loop A affects cleavage efficiency (Figure 5F).

Stem-loop A determines position of cleavage

To investigate whether stem-loop A also governs the position of cleavage, we inserted heterologous sequences of ACACAC, UCUCUC or UGUGUG in the single-stranded region between the cleavage site and stem-loop A and examined the influence of the insertions on the cleavage of the $\Delta 222$ construct (Figure 6A). Northern (Figure 6B) and primer extension (Supplementary Figure S4) analysis of all three mutants expressed from a plasmid clearly shows cleavage to generate a product similar to the size of MicL-S, again suggesting cleavage can occur at a very precise distance from stem-loop A regardless of sequence. We also detect an extra band, particularly for the UCUCUC and UGUGUG mutants that is slightly larger than MicL-S. This extra band could be resulting from cleavage at the AAAUGIAUUU sequence. However, an extended stem-loop A that can be predicted for the UCUCUC and UGUGUG constructs also might be directing cleavage at the second position.

Distances between two 3' stem-loops of MicL affect cleavage

We noted that the two stem-loops are directly adjacent to each other and next examined the consequences of altering the distance between the two structures by inserting repeating U and C 'spacers' of 10 or 15 nucleotides between

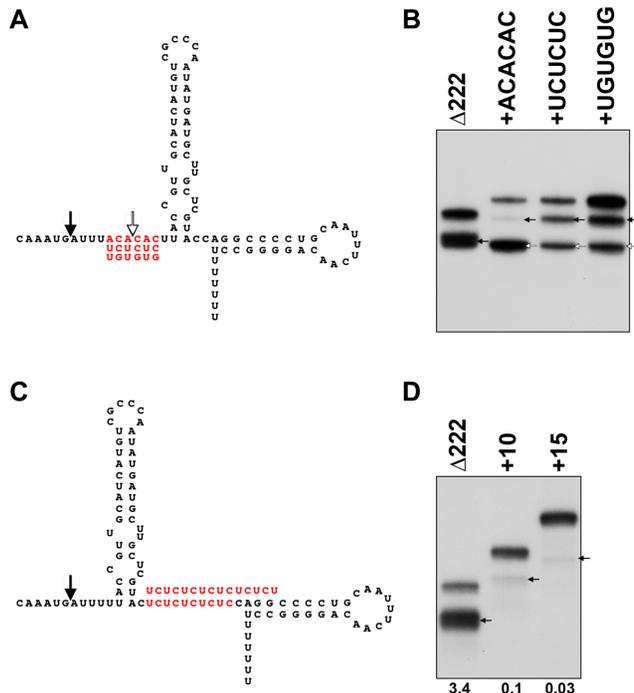


Figure 6. Stem-loop A and distance between two stem-loops of MicL affect cleavage. (A) Predicted secondary structure of $\Delta 222$ constructs with insertions made in separate constructs given in red. The black arrow indicates cleavage site in $\Delta 222$, and the white arrow indicates the new cleavage site observed. (B) The $\Delta 222$ construct and the three $\Delta 222$ derivatives depicted in Figure 6A were expressed from pBR* in $\Delta cutC$ background. Arrows denote cleavage products corresponding to the sites indicated in Figure 6A. (C) Predicted secondary structure of $\Delta 222$ constructs with spacer sequences introduced between stem-loop A and stem-loop B given in red. The arrow indicates cleavage site. (D) The $\Delta 222$ construct and the two $\Delta 222$ derivatives depicted in Figure 6C were expressed from pBR* in $\Delta cutC$ background. Arrows denote cleavage products corresponding to the site indicated in C. For samples in panels B and D, total RNA was extracted and probed for MicL as in Figure 3B. For panel D, the bottom band:top band ratio is given below each lane.

stem-loop A and B in the $\Delta 222$ context (Figure 6C). Secondary structure predictions indicate the spacer constructs create a single stranded region between the stem-loops and do not alter the stem-loop structures and their stabilities. Again, total RNA isolated from strains carrying the resulting constructs was subjected to northern analysis (Figure 6D). Compared to the $\Delta 222$ control, both spacer constructs showed decreased levels of cleavage with less product for the +15 construct than for the +10 construct. These data indicate that two stable stem-loops in close proximity to the RNase E cleavage site are required for efficient MicL cleavage, superseding recognition of a specific sequence.

Strongest cleavage is observed with two adjacent 3' stem-loops

To further test whether two stem-loops are needed for efficient cleavage and to examine the influence of these two stem-loops with respect to the distance from and position 5' or 3' of the cleavage site, we synthesized a number of synthetic constructs (Figure 7A) starting with the $\Delta 222$ -stemA $\Delta 2$ MicL derivative. For all of the synthetic

constructs, the same linear sequence of four repeats of the MicL cleavage site (UGAUU) separated by a CC or UC spacer was flanked by different stem-loops. The repeat sequence was predicted to be single-stranded and not affect the structures of added or altered stem-loops in each of the constructs. Total RNA extracted from the strains harboring the resulting vectors was subject to northern analysis (Figure 7B). For the construct with the repeats followed by stemA $\Delta 2$ and the wild type stem-loop B terminator (repeats+stemA $\Delta 2$ +term, Figure 7B, lane 2), the 95 nt transcript gave the same predominant MicL-S product as the $\Delta 222$ -stemA $\Delta 2$ construct (Figure 7B, lane 1), though one additional minor cleavage product was observed. These data suggest that although additional putative cleavage sites are present in the 5'-single stranded region, cleavage was directed towards the site nearest the two 3' stem-loops.

The requirement for having two 3' stem-loops was again tested by deleting stemA (repeats+term). Analysis of the total RNA from the strain carrying this construct showed there was very little cleavage of the 69 nt transcript with just stem-loop B (Figure 7B, lane 3). We suggest that the four minor products observed correspond to cleavage at the four consensus sequences in the linear repeat. The addition of a Spot 42 stem-loop in place of stem A (repeats+Spot 42 stem+term) restored strong cleavage at the position closest to the stem-loops (Figure 7B, lane 4), further indicating that efficient cleavage requires two stable stem-loops, independent of the sequences of the stem-loops.

To test whether stem-loops at the 5' end of the repeat sequence could also influence cleavage, we examined one additional construct carrying stemA $\Delta 2$ upstream and the stem-loop B terminator downstream of the repeat sequence (stemA $\Delta 2$ +repeats+term). Only very faint cleavage products were detected for this construct (Figure 7B, lane 5). Thus, two adjacent stem-loops are required for robust cleavage.

Mutations in MicL and RNase E decrease MicL cleavage *in vitro*

Finally, we tested whether the effects of the MicL mutations and insertions similarly impacted cleavage by purified RNase E. *In vitro* transcribed $\Delta 222$, stemA $\Delta 2$, +10 and +15 were all incubated with purified RNase E (1–529) and examined by northern analysis. The results of the *in vitro* experiments are consistent with the *in vivo* findings with stemA $\Delta 2$ showing somewhat increased cleavage and +10 and +15 showing decreased cleavage (Figure 8A).

Two recent crystal structures of a D303R,D346R catalytically-inactive RNase E (1–529) with sRNA fragments (22) have shown that helices of both RprA and SdsR are recognized by RNase E. The structures led to the identification of eight amino acid residues that when mutated (8 \times RNase E 1-529) impaired RNase E binding to and cleavage of RNA substrates with helical elements adjacent to the cleavage site. Among the substrates tested was the 9S rRNA, previously reported to contain structural elements important for RNase E recognition (23). We similarly did not observe cleavage of the 9S rRNA with the 8 \times RNase E mutant unless the *in vitro* transcript was treated with calf intestinal phosphatase (CIP) to convert the 5' triphosphate

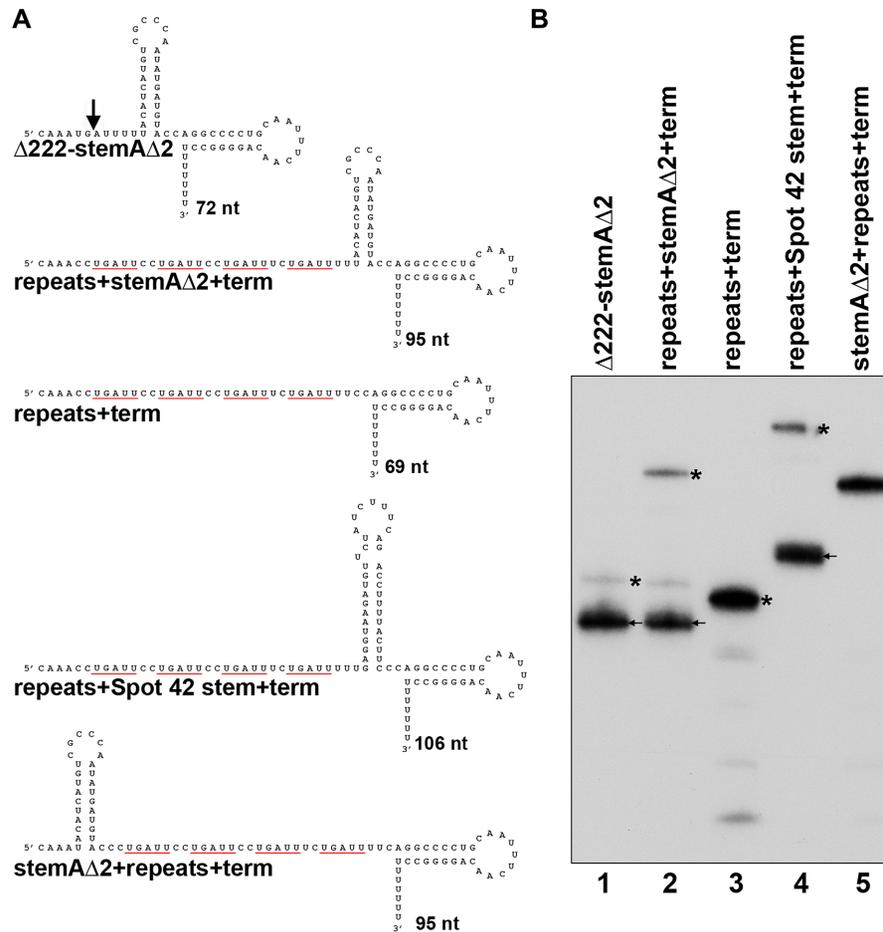


Figure 7. Strongest cleavage is observed with two 3' stem-loops. (A) Predicted secondary structures of constructs with repeat sequences between different stem-loop structures. Secondary structures were predicted using the Mfold software package. The arrow indicates the site of RNase E cleavage. UGAUU repeats are underlined in red. (B) Cleavage is directed to the site nearest to two stable stem-loops. The constructs from Figure 7A were expressed from pBR* in $\Delta cutC$ background. Total RNA was extracted and probed for MicL as in Figure 3B. Full length products are denoted by asterisks, and prominent cleavage products are denoted by arrows.

to monophosphate to allow 5' end recognition (Supplementary Figure S5). Consistent with RNase E recognition of stem-loop structures in MicL, the $\Delta 222$ derivative was not cleaved by the 8 \times RNase E mutant (Figure 8B). In contrast, *in vitro* synthesized ArcZ was partially cleaved by this mutant (Supplementary Figure S6).

As also observed in Figure 2, in the absence of Hfq, we no longer detected $\Delta 222$ cleavage with wild type RNase E (1–529), but the RNA was completely degraded by the hyperactive D26N,D28N,D338N mutant (Figure 8B). The ArcZ RNA was almost completely degraded by both the wild type and the D26N,D28N,D338N mutant proteins and partially degraded by the 8 \times mutant in the absence of Hfq (Supplementary Figure S6). Together these results indicate that Hfq may have two roles in promoting MicL cleavage, positioning RNase E to help direct cleavage at a specific position as well as protecting the cleaved product from excess cleavage.

DISCUSSION

The 308 nt σ^E -dependent MicL sRNA, transcribed from within the *cutC* gene, was previously reported to be cleaved

to give a 80 nt derivative capable of binding Hfq and repressing the synthesis of the abundant *lpp* mRNA (4). The results presented here show that MicL is cleaved by RNase E, and that precise cleavage is dependent on the presence of two adjacent hairpin structures 3' to the cleavage site. Cleavage was found to be affected by both the stability of the two hairpins and the distance between them, with less impact of the sequence of the two hairpins or the sequence in and around the cleavage site. We also observed that the first stem-loop dictated the position of cleavage, and that Hfq promoted the specific cleavage. Our results suggest that RNase E in conjunction with Hfq is recognizing helical structures leading to cleavage at a specified distance from this recognition element.

sRNAs are differentially affected by *rne-3071*

The sequence of the MicL cleavage site is similar to the RNase E core motif identified in a recent study mapping RNase E cleavage sites transcriptome-wide by transient inactivation of RNase E followed by high-throughput RNA sequencing in *S. enterica* (6). However, even in this high-

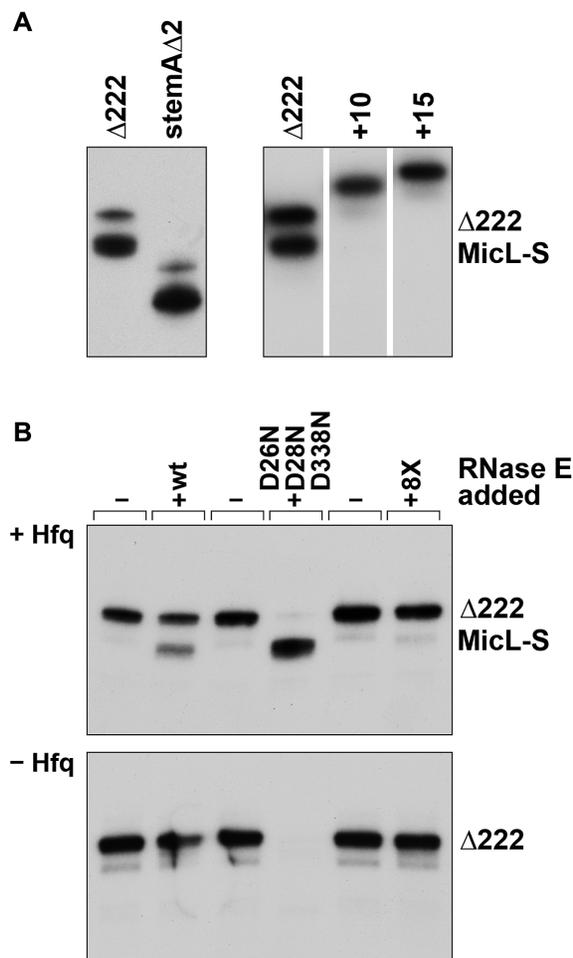


Figure 8. Effects of mutations on MicL cleavage *in vitro*. (A) *In vitro* transcribed $\Delta 222$, $\Delta 222$ -stemA $\Delta 2$, $\Delta 222+10$ and $\Delta 222+15$ with 5'PPP were incubated with purified Hfq and purified wild type RNase E (1–529) at 30°C for 30 min. The blot shown is from a single image with the indicated lanes spliced together. (B) *In vitro* transcribed $\Delta 222$ was incubated with and without purified Hfq and purified wild type RNase E (1–529), D26N, D28N, D338N mutant RNase E (1–529) or 8X mutant RNase E (1–529) at 30°C for 30 min. For both A and B, the RNA was subject to northern analysis using an oligonucleotide probe complementary to the 3'-end of MicL.

throughput study, the RNA levels of MicL and MicL-S were not severely affected by the transient inactivation of the RNase E, at least compared to other known RNase E sRNA substrates, analogous to what we observed (Figure 1).

The RNase E temperature-sensitive allele (*rne-3071*) in *E. coli* has a C742T transition resulting in the F68L substitution near the nucleotide binding motif (24,25). This temperature-sensitive mutation was found to increase the chemical half-life of total pulse-labeled RNA (26) and affect the steady state levels of individual RNAs cleaved by RNase E (27). Our results show different degrees of sRNA cleavage for the *rne-3071* strain at the non-permissible temperature; cleavage is almost completely abolished for ArcZ, significantly decreased for CpxQ and only slightly decreased for MicL (Figure 1A). Nevertheless, we think MicL is in fact a substrate for RNase E since the extent of MicL processing is

decreased upon RNase E depletion *in vivo* (Figure 1B) and cleavage is observed at the same position with the purified NTD of RNase E *in vitro* (1–529) (Figure 2).

Our observations are consistent with early studies of the known RNase E substrates RNAI and T4 mRNA, for which it was found that not all RNase E decay intermediates are reduced in the *rne-3071* strain at the non-permissible temperature (11,13). In fact, when comparing RNAI cleavage by wild-type and *rne-3071* at the non-permissible temperature, some decay products were decreased but several products were similar between the two strains and others were enhanced. Furthermore, the *rne-3071* strain seemed to generate new RNAI decay bands not seen in the *rne*⁺ strain at the non-permissible temperature, leading the authors to warn about assigning RNase E cleavage sites based entirely on changes in RNA abundance in the *rne-3071* strain (13).

Recognition of MicL stem-loops by RNase E

From our mutational analyses, we propose that the structural elements at the 3' end of the MicL RNA are being recognized by RNase E in what is a 'direct entry' pathway for cleavage. The 5' end of *in vivo* transcribed MicL RNA is triphosphorylated (28) and thus is less likely to be recognized by RNase E. It is interesting to note that several known RNase E substrates have two stem-loop structures either preceding or following the cleavage site. This is the case for 9S RNA where two stem-loops are predicted to precede the single-stranded RNase E cleavage site and disruptions of these structures perturb cleavage while sequences 3' of the cut site are dispensable (23). Similarly, three closely spaced stem-loop structures succeed the RNase E cleavage site of RNAI, while 5' sequences are dispensable (13). Although the predominant cleavage site occurs four nucleotides from the first 3' stem-loop structure of RNAI, a less intense cleavage site also occurs 4 nucleotides upstream of the second hairpin. Thus, it would seem, like with MicL, two closely spaced hairpins dictate cleavage of RNAI at a specified distance from an adjacent hairpin.

We suggest that the regions of RNase E, the RNase H like domain and small domain located distal to the catalytic active site of RNase E, observed to interact with helical structures of RprA and SdsR in the crystal structures (22) comprise the domain responsible for structural recognition of MicL. Consistent with this view, the RNase E (1–529) 8× mutant (with mutations to R3, Q22, H268, Q270 and Y269 in the RNase H domain and K433, R488 and R490 in the small domain) did not cleave $\Delta 222$ *in vitro* (Figure 8B). Due to the tetrameric organization of RNase E, it is conceivable that the two MicL 3' stem-loops can contact two adjacent subunits within the principle dimer and that these close contacts anchor the single stranded cleavage site near the active site of the protein, explaining the requirement for a specific distance between the two 3' stem-loops as well as the specific site of cleavage relative to the first stem-loop. Consistent with this possibility, a model of RNase E (1–529) with RprA shows that the distance between the helical recognition domain and the active site of the enzyme is approximately 5 nt (B. Luisi, personal communication). Given that MicL is a known Hfq binding sRNA (4) and Hfq likely interacts with the rho-independent terminator struc-

ture (29,30), another explanation for the distance requirement between the two stem-loops is the need for a ternary complex between MicL, RNase E and Hfq. In this model, RNase E could bind the first stem-loop, while Hfq binds the second stem-loop. The lack of $\Delta 222$ cleavage by RNase E (1–529) when Hfq is omitted *in vitro* could be explained by this hypothesis.

RNase E recognition of a combination of sRNA features could lead to a cleavage hierarchy

ArcZ, CpxQ as well as other 3' UTR-derived sRNAs that are cleaved by RNase E are predicted to have two closely spaced stem-loops at their 3' ends, the second one corresponding to the Rho-independent transcription terminator (1,2,31). Thus, we predict that structural features like what we have found for MicL help direct cleavage of other RNAs, especially 3' UTR-derived sRNAs. We note that features in addition to the stem-loops could contribute to the very specific cleavage such that an sRNA might be recognized by a combination of 5' sensing and direct entry. The individual recognition elements—5' end, single stranded AU-rich sequence, 3' stem loops and Hfq binding—each could contribute differently to the efficiency of cleavage. This possibility is underscored by the different sensitivities of MicL, ArcZ and CpxQ to the *rne-3017* allele and RNase E depletion, and the differences in MicL and ArcZ cleavage *in vitro*. Overall our results suggest that RNase E substrate recognition is far more nuanced than initially imagined, allowing for even more levels or regulation and a hierarchy of RNase E-mediated cleavage.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank B. Luisi for helpful discussions and the kind gifts of the purified 1–529 domains of wild type, D26N, D28N, D338N mutant and 8 \times mutant RNase E. We are grateful to K. Lee for sharing the RNase E depletion strain. We also thank S. Gottesman and B. Luisi for comments on this manuscript.

FUNDING

Intramural Research program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (to T.B.U., A.B.K. and G.S.); Wellcome Trust (to K.J.B.). Funding for open access charge: Intramural Research program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development [ZIA HD001608-27].

Conflict of interest statement. None declared.

REFERENCES

- Chao, Y. and Vogel, J. (2016) A 3' UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. *Mol. Cell*, **61**, 352–363.
- Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M. and Vogel, J. (2012) An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J.*, **31**, 4005–4019.
- Kim, H.M., Shin, J.H., Cho, Y.B. and Roe, J.H. (2014) Inverse regulation of Fe- and Ni-containing SOD genes by a Fur family regulator Nur through small RNA processed from 3'UTR of the *sodF* mRNA. *Nucleic Acids Res.*, **42**, 2003–2014.
- Guo, M.S., Updegrove, T.B., Gogol, E.B., Shabalina, S.A., Gross, C.A. and Storz, G. (2014) MicL, a new σ E-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev.*, **28**, 1620–1634.
- Miyakoshi, M., Chao, Y. and Vogel, J. (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J.*, **34**, 1478–1492.
- Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., Šmiga, M., Papenfort, K., Reinhardt, R., Wieden, H.J. *et al.* (2017) *In vivo* cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell*, **65**, 39–51.
- Mackie, G.A. (2013) RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.*, **11**, 45–57.
- Bandyra, K.J., Bouvier, M., Carpousis, A.J. and Luisi, B.F. (2013) The social fabric of the RNA degradosome. *Biochim. Biophys. Acta*, **1829**, 514–522.
- Mohanty, B.K. and Kushner, S.R. (2016) Regulation of mRNA decay in bacteria. *Annu. Rev. Microbiol.*, **70**, 25–44.
- Del Campo, C., Bartholomäus, A., Fedyunin, I. and Ignatova, Z. (2015) Secondary structure across the bacterial transcriptome reveals versatile roles in mRNA regulation and function. *PLoS Genet.*, **11**, e1005613.
- Ehretsmann, C.P., Carpousis, A.J. and Krisch, H.M. (1992) Specificity of *Escherichia coli* endoribonuclease RNase E: *in vivo* and *in vitro* analysis of mutants in a bacteriophage T4 mRNA processing site. *Genes Dev.*, **6**, 149–159.
- Kaberlin, V.R. (2003) Probing the substrate specificity of *Escherichia coli* RNase E using a novel oligonucleotide-based assay. *Nucleic Acids Res.*, **31**, 4710–4716.
- McDowall, K.J., Lin-Chao, S. and Cohen, S.N. (1994) A+U content rather than a particular nucleotide order determines the specificity of RNase E cleavage. *J. Biol. Chem.*, **269**, 10790–10796.
- Deana, A., Celesnik, H. and Belasco, J.G. (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature*, **451**, 355–358.
- Clarke, J.E., Kime, L., Romero, A.D. and McDowall, K.J. (2014) Direct entry by RNase E is a major pathway for the degradation and processing of RNA in *Escherichia coli*. *Nucleic Acids Res.*, **42**, 11733–11751.
- Göpel, Y., Khan, M.A. and Görke, B. (2016) Domain swapping between homologous bacterial small RNAs dissects processing and Hfq binding determinants and uncovers an aptamer for conditional RNase E cleavage. *Nucleic Acids Res.*, **44**, 824–837.
- Guillier, M. and Gottesman, S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.*, **59**, 231–247.
- Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
- Lee, K., Bernstein, J.A. and Cohen, S.N. (2002) RNase G complementation of *rne* null mutation identifies functional interrelationships with RNase E in *Escherichia coli*. *Mol. Microbiol.*, **43**, 1445–1456.
- Hao, Y., Updegrove, T.B., Livingston, N.N. and Storz, G. (2016) Protection against deleterious nitrogen compounds: role of σ^S -dependent small RNAs encoded adjacent to *sdia*. *Nucleic Acids Res.*, **44**, 6935–6948.
- Ono, M. and Kuwano, M. (1979) A conditional lethal mutation in an *Escherichia coli* strain with a longer chemical lifetime of messenger RNA. *J. Mol. Biol.*, **129**, 343–357.
- Bandyra, K.J., Wandzik, J. and Luisi, B.F. (2018) Substrate recognition and autoinhibition in the central ribonuclease RNase E. *Mol. Cell*, **72**, 275–285.
- Cormack, R.S. and Mackie, G.A. (1992) Structural requirements for the processing of *Escherichia coli* 5 S ribosomal RNA by RNase E *in vitro*. *J. Mol. Biol.*, **228**, 1078–1090.

24. Apirion, D. (1978) Isolation, genetic mapping and some characterization of a mutation in *Escherichia coli* that affects the processing of ribonucleic acid. *Genetics*, **90**, 659–671.
25. McDowall, K.J., Hernandez, R.G., Lin-Chao, S. and Cohen, S.N. (1993) The *ams-1* and *rne-3071* temperature-sensitive mutations in the *ams* gene are in close proximity to each other and cause substitutions within a domain that resembles a product of the *Escherichia coli mre* locus. *J. Bacteriol.*, **175**, 4245–4249.
26. Kuwano, M., Ono, M., Endo, H., Hori, K., Nakamura, K., Hirota, Y. and Ohnishi, Y. (1977) Gene affecting longevity of messenger RNA: a mutant of *Escherichia coli* with altered mRNA stability. *Mol. Gen. Genet.*, **154**, 279–285.
27. Babitze, P. and Kushner, S.R. (1991) The Ams (altered mRNA stability) protein and ribonuclease E are encoded by the same structural gene of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 1–5.
28. Thomason, M.K., Bischler, T., Eisenbart, S.K., Förstner, K.U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C.M. and Storz, G. (2015) Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J. Bacteriol.*, **197**, 18–28.
29. Otaka, H., Ishikawa, H., Morita, T. and Aiba, H. (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 13059–13064.
30. Sauer, E. and Weichenrieder, O. (2011) Structural basis for RNA 3'-end recognition by Hfq. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 13065–13070.
31. Soper, T., Mandin, P., Majdalani, N., Gottesman, S. and Woodson, S.A. (2010) Positive regulation by small RNAs and the role of Hfq. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 9602–9607.