

***In vivo* cleavage map illuminates the central role
of RNase E in coding and noncoding RNA pathways**

Yanjie Chao¹, Lei Li^{1,2}, Dylan Girodat³, Konrad U. Förstner^{1,2}, Nelly Said⁴,

Colin Corcoran¹, Michał Śmiga¹, Kai Papenfort^{1,5}, Richard Reinhardt⁶,

Hans-Joachim Wieden³, Ben F. Luisi⁷, Jörg Vogel^{1,8,9*}

1. Institute of Molecular Infection Biology, University of Würzburg, 97080 Würzburg, Germany
2. Core Unit Systems Medicine, University of Würzburg, 97080 Würzburg, Germany
3. Alberta RNA Research and Training Institute, Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada
4. Laboratory of Structural Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany
5. Department of Biology I, Microbiology, Ludwig-Maximilians-University Munich, 82152 Martinsried, Germany
6. Max Planck Genome Centre Cologne, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany
7. Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, United Kingdom
8. Helmholtz Institute for RNA-based Infection Research (HIRI), 97080 Würzburg, Germany

*Corresponding author

E-mail: joerg.vogel@uni-wuerzburg.de

Phone: +49-931-3182575

Fax: +49-931-3182578

Running title: TIER-Seq mapping of RNase E sites

SUMMARY

Understanding RNA processing and turnover requires knowledge of cleavages by major endoribonucleases within a living cell. We have employed TIER-Seq (transiently inactivating an endoribonuclease followed by RNA-Seq) to profile cleavage products of the essential endoribonuclease RNase E in *Salmonella enterica*. A dominating cleavage signature is the location of a uridine two nucleotides downstream in a single-stranded segment, which we rationalize structurally as a key recognition determinant that may favor RNase E catalysis. Our results suggest a prominent biogenesis pathway for bacterial regulatory small RNAs, whereby RNase E acts together with the RNA chaperone Hfq to liberate stable 3'-fragments from various precursor RNAs. Recapitulating this process *in vitro*, Hfq guides RNase E cleavage of a representative small RNA precursor for interaction with a mRNA target. *In vivo*, the processing is required for target regulation. Our findings reveal a general maturation mechanism for a major class of post-transcriptional regulators.

Keywords: RNase E, RNA degradome, noncoding RNA, Hfq, 3'UTR, ArcZ, RprA, sRNA maturation, uridine ruler, TIER-seq

Highlights:

- TIER-seq precisely maps ~22,000 endogenous RNase E cleavage sites in *Salmonella*
- Consensus motif of RNase E reveals a 2-nucleotide uridine ruler-and-cut mechanism
- RNase E as a central component in both maturation and degradation of small RNAs
- A general small RNA biogenesis pathway requiring RNase E and Hfq

INTRODUCTION

Small noncoding RNAs (sRNAs) that associate with the RNA chaperon Hfq constitute the largest class of post-transcriptional regulators in Gram-negative bacteria (De Lay et al., 2013; Storz et al., 2011; Vogel and Luisi, 2011; Wagner and Romby, 2015). Originally defined as a class in nonpathogenic *Escherichia coli* (Zhang et al., 2003), Hfq-dependent sRNAs have been globally mapped in numerous important human pathogens (Barquist and Vogel, 2015; Holmqvist et al., 2016; Koo et al., 2011; Melamed et al., 2016; Tree et al., 2014). These sRNAs generally act as multi-target repressors and activators through seed pairing interactions with the 5' untranslated region (UTR) of mRNAs (Desnoyers et al., 2013; Feng et al., 2015; Pappenfort and Vanderpool, 2015). A full understanding of these sRNA-mediated networks requires knowledge of how their RNA constituents are synthesized and turned over.

Many of the bacterial sRNAs characterized to date are transcribed from noncoding intergenic regions and operate as full-length, primary transcripts capped with a 5' triphosphate (5'PPP). However, some primary sRNAs such as ArcZ and RprA are converted into shorter stable species that retain the seed region for target mRNA recognition (Mandin and Gottesman, 2010; Pappenfort et al., 2015; Pappenfort et al., 2009), but it is currently unclear whether such processing generates the active sRNAs, as with eukaryotic microRNAs (Kim, 2005). Moreover, several recent studies reported sRNAs that are produced from the 3' region of mRNA genes (Miyakoshi et al., 2015b), only a subset of which are the result of gene-internal promoters (Chao et al., 2012; Guo et al., 2014), while many others appear to originate from mRNA processing. These 3' derived sRNAs are likely to be functional since they abundantly associate with Hfq (Chao et al., 2012) whose cellular concentration is limited (Wagner, 2013). Their physiological importance is further supported by established roles of the 3' mRNA-derived sRNAs CpxQ and SroC in the envelope stress response or amino acid pathways, respectively (Chao and Vogel, 2016; Miyakoshi et al., 2015a). Furthermore, 3' fragments of *E. coli* tRNA precursors function as molecular sponges of conserved sRNAs (Lalaouna et al., 2015). Collectively, the growing findings suggest that sRNA processing is a prevalent event, but both its functional relevance and the major responsible nuclease(s) remain to be established.

Of several candidate nucleases involved in sRNA processing and turnover, the conserved and essential endoribonuclease E (RNase E) is the likely central player (Mackie, 2013; Masse et al., 2003; Saramago et al., 2014). Inferred from transcript accumulation upon its inactivation, RNase E

drives the decay of most mRNAs in *E. coli* (Bernstein et al., 2004; Clarke et al., 2014), and in *Salmonella* processes the mRNA 3' end-derived CpxQ and SroC sRNAs (Chao and Vogel, 2016; Miyakoshi et al., 2015a). RNase E also degrades several sRNAs in the absence of Hfq or upon base pairing with target mRNAs (Bandyra et al., 2012; Masse et al., 2003; Moll et al., 2003). Conversely, some sRNAs activate gene expression by blocking RNase E cleavage sites in target mRNAs (Fröhlich et al., 2013; Papenfort et al., 2013). In addition, RNase E is known to engage in rRNA and tRNA precursor processing (Apirion and Lassar, 1978; Bessarab et al., 1998; Kime et al., 2014; Li and Deutscher, 2002; Ow and Kushner, 2002).

Despite the importance of RNase E in post-transcriptional control, its activity towards most noncoding RNAs is not known. Previous studies have characterized major RNase E cleavage sites in a few abundant model transcripts (e.g. (Apirion and Lassar, 1978; Delvillani et al., 2011; Ehretsmann et al., 1992; Mackie, 1991; Ow and Kushner, 2002; Patel and Dunn, 1992; Regnier and Hajnsdorf, 1991; Roy and Apirion, 1983), and concluded that the enzyme preferentially cleaves AU-rich regions in single stranded RNA (Arraiano et al., 2010; Huang et al., 1998; McDowall et al., 1995; McDowall et al., 1994). Here, to achieve a systems-level understanding of RNase E activity, we have analyzed in depth the *in vivo* RNase E cleavage events in *Salmonella* Typhimurium, a close relative of *E. coli* and a pathogenic model organism to study post-transcriptional regulation (Westermann et al., 2016). Our genome-wide capture of tens of thousands of endogenous cleavage sites reveals a minimal consensus sequence and a 2 nt uridine ruler-and-cut structural mechanism for this major endoribonuclease. Intriguingly, RNase E employs this mechanism to cleave many coding and noncoding transcripts at the 3' end and releases stable Hfq-bound RNA fragments, indicating that sRNA biogenesis through endonucleolytic processing is widespread. Searches for these predicted critical uridines in sRNAs enabled us to show that maturation by RNase E is essential for target regulation by the ArcZ sRNA. Moreover, our data reveal a high frequency of RNase E-mediated cleavages in Hfq-dependent sRNAs, supporting the functional link between RNase E and Hfq for the first time on a global level.

RESULTS

A transcriptome-wide map of RNase E cleavage sites in vivo

To globally map RNase E cleavage events *in vivo*, we profiled 5' ends of cellular transcripts by comparative RNA-seq before and 30 minutes after programmed inactivation of the enzyme using a

temperature-sensitive *rne*^{TS} mutant (*rne*-3071, (Apirion and Lassar, 1978; Figueroa-Bossi et al., 2009)). We refer to this approach, which builds upon work by Clarke and colleagues (Clarke et al., 2014) as Transient Inactivation of Endoribonuclease followed by RNA-seq (TIER-Seq; see Fig. 1A). At the permissive temperature (28°C), *Salmonella* wild-type (WT) *rne* or mutant *rne*^{TS} strains both exhibit full RNase E activity; whereas upon shift to 44°C, only WT RNase E retains its activity to process RNA. To achieve a comprehensive RNase E-specific ‘degradome’ analysis at single-nucleotide resolution (Fig. 1A), we analyzed biological duplicates of all four of the above strains and conditions in the early stationary growth phase (OD₆₀₀ of 2) by RNA-seq, obtaining ~130 million reads (Fig. S1A). In agreement with previous work showing that RNase E cleaves AU-rich sequences (McDowall et al., 1995; McDowall et al., 1994), the inactivation of RNase E leads to a ~5% reduction of cDNA reads with 5'-A/T bases (Fig. S1B).

To pinpoint cleavage sites, we aligned all reads to the *Salmonella* genome, mapping a total of ~500,000 unique 5' ends (Fig. 1B-C). WT and *rne*^{TS} samples from growth at 28°C gave nearly identical 5' end profiles ($R^2=0.98$; Fig. 1B & S1C), confirming that the mutant RNase E is fully functional at the permissive temperature; whereas at the non-permissive temperature (44°C), many positions were selectively depleted in the *rne*^{TS} cDNA libraries (Fig. 1C). Since *Salmonella* has no 5'→3' exoribonuclease (Hui et al., 2014), we interpret these depleted positions as RNase E cleavage sites (Fig. 1A). This classification is supported by the capture of many previously known *E. coli* RNase E cleavage sites (Fig. 1D), for example, in the *rpsO*, *cspE*, *uncC/atpC*, or *glmUS* mRNAs (Delvillani et al., 2011; Joanny et al., 2007; Patel and Dunn, 1992; Regnier and Hajnsdorf, 1991), in the 9S precursor of 5S rRNA (Roy and Apirion, 1983), or near the 3' end of tRNAs (Ow and Kushner, 2002). Applying a threshold of >3-fold as significant depletion ($P<0.05$, FDR<0.05) in the *rne*^{TS} samples at 44°C, we assigned 22,033 RNase E-mediated cleavages in the *Salmonella* transcriptome, expanding by several orders of magnitude the database of *in vivo* target sites for this ribonuclease. The full list of cleavage sites is available in Table S1.

A systems-level view on RNase E activity in RNA metabolism

Systematic analysis of the 22,033 RNase E cleavage sites revealed their distribution in coding and noncoding transcripts from the *Salmonella* chromosome and virulence plasmids (Fig. 2A): ~80% occurred in mRNAs, primarily in the coding sequence (CDS), indicating that a major activity of RNase E is to degrade mRNAs, in addition to processing housekeeping RNAs. Altogether, we

detected a total of 2,557 mRNAs cleaved by RNase E, with a different number of cleavage sites per transcript (Fig. 2B); these represent 78% of 3,286 of *Salmonella* mRNAs that are well-expressed (RPKM>10, Table S2) in the early stationary phase. Notably, the assay captured many essential genes and virulence genes required for intracellular growth (Table S3), which provide insights into the processing of transcripts from indispensable genes and the roles of RNase E in *Salmonella* pathogenesis (Viegas et al., 2013), respectively. Longer transcripts generally tend to contain a higher number of cleavage sites (Fig. S1 D-E). After normalizing the number of cleavage sites to gene length, RNase E cleavage frequency in these genes (RPKM>10) ranges from 0 to ≥ 30 sites per kilo base with a median value at ~ 5.7 cleavages per kilo bases; or one site every ~ 175 nt of mRNA (Fig. 2C). This non-saturating cleavage pattern may suggest that most sites in mRNAs are inaccessible perhaps due to structural constraints or protein binding.

The position of an RNase E site within a transcript may provide information about the function of the cleavage. For example, RNase E autoregulates its synthesis by cutting in the 5' UTR of its own mRNA (Jain and Belasco, 1995); our analysis readily captured this critical site (Fig. S2A). As another example, we detect the RNase E site in the 5'UTR of *cfa* mRNA (Fig. S2A) that becomes protected by the *trans*-acting RydC sRNA with the consequence that the transcript is stabilized (Dimastrogiovanni et al., 2014; Fröhlich et al., 2013)(Fröhlich et al., 2014). Thus, our candidate list of $\sim 1,300$ RNase E cleavage sites identified in the 5' UTRs of 548 genes (Table S4) provides a resource to predict sites for post-transcriptional control by sRNAs and/or RNA-binding proteins.

A specific sequence motif recognized by RNase E

Even seemingly non-specific nucleases often exhibit a certain degree of sequence or structural preference. To understand the substrate determinants of RNase E activity, we analyzed the primary sequences and putative secondary structures around all the 22,033 cleavage sites. At the cleavage site, we observed an overall increase in the calculated folding energy (ΔG) indicating little secondary structure (Fig. 2D) and a spike of AU-rich sequences (Fig. 2E), both of which agree with previously studied individual RNase E sites (McDowall et al., 1995; McDowall et al., 1994). Importantly, sequence alignment of all the 22,033 sites predicts a minimal RNase E consensus sequence (Fig. 2F), with a marked preference for uridine at the +2 position in the 5-nt 'RN \downarrow WUU' core motif (with R as G/A, W as A/U, and N as any nucleotide). This RNase E motif, based entirely on global *in vivo* data, fully recapitulates preferences previously documented with model substrates *in*

vitro (Ehretsmann et al., 1992; Kaberdin, 2003; Mackie, 1991) and cell-derived RNA (Del Campo et al., 2015), while it clearly differs from recognition motifs of other major bacterial endoribonucleases such as tRNA-processing RNase P (McClain et al., 1987) or RNase III which cleaves double-stranded RNA (Gan et al., 2005).

RNase E cleavages underlie sRNA biogenesis from 3'UTRs

In analyzing cleavage site distributions relative to mRNA start or stop codons (Fig. 3A-B), we observed that, on average, 5' UTRs and the coding regions showed similar cleavage frequencies. Translation initiation regions were slightly counter-selected, perhaps because the prominent Shine-Dalgarno sequence (GGAGGA) is devoid of RNase E cleavage motifs. In contrast, RNase E sites were enriched around mRNA stop codons (Fig. 3B); the high-AU rich content and/or translation termination may favor this enrichment. Since bacterial 3' UTRs are generally short (Belasco, 2010), many of these stop codon sites may represent the most downstream site, leaving 3' fragments for degradation by 3'→5' exoribonucleases. Interestingly, approximately one third of these mRNAs carry protective ρ -independent terminators (Arraiano et al., 2010) that can, in principle, interact with the sRNA chaperone Hfq (Otaka et al., 2011; Sauer and Weichenrieder, 2011). These data point to the possibility that stable 3' UTR fragments may accumulate with functional consequence in the guise of regulatory sRNAs (Table S5) (Chao et al., 2012; Miyakoshi et al., 2015b). Indeed, we have detected the mRNA 3'UTR processing sites that produce the CpxQ and SroC sRNAs (Fig. S2B). Northern blot probing of several selected candidates revealed distinct RNA species from mRNA 3' ends, the generation of which required both active RNase E and presence of Hfq (Figs. 3C, S2C). Most of these 3' derived sRNAs co-accumulate with their parental mRNA transcripts and possess potential seed regions (Fig. S3), suggesting that they are bona-fide regulatory sRNAs with conserved targets and functions. In addition, the cleavage sites in these sRNAs resemble the 'RNWUU' sequence (Fig. S2D), supporting the recognition of this consensus by RNase E (Fig. 2F).

Cleavage by RNase E produces sRNAs from noncoding RNA precursors

The majority of well-characterized Hfq-dependent sRNAs in *E. coli* and *Salmonella* are primary transcripts of 50-250 nt in length. Although previous work on a few model sRNAs has implicated RNase E in their decay (Göpel et al., 2013; Madhugiri et al., 2010; Miyakoshi et al., 2015a; Viegas et

al., 2007), it is unknown whether this sRNA class is generally processed by RNase E. Here, we have mapped ~600 RNase E cleavage sites in 107 experimentally validated sRNAs (Table S6), corroborating previously proposed sites in model sRNAs such as DsrA and MicA (Fig. S4 A-B) (Moll et al., 2003). RNase E seems to target preferentially sRNAs that are bound by Hfq, as there are more cleavage sites in Hfq-dependent sRNAs compared to those that are Hfq-independent (Chao et al., 2012) (Fig. S4C). Additionally, many cleavage sites in these sRNAs mapped to the vicinity of the seed region (Fig. S4D), as exemplified by their clustering in the well-characterized seed of SgrS and RyhB (Fig. S4 E-F). These data suggest that RNase E may inactivate sRNAs by removing the seed region, which is in agreement with previous results for MicC (Bandyra et al., 2012) and RyhB (Masse et al., 2003; Moll et al., 2003). Both MicC and RyhB are turned over by RNase E through seed cleavage if the target is absent, and this could provide a surveillance mechanism for accurate seed matching (Bandyra et al., 2012).

Another group of sRNAs is spared from immediate degradation following RNase E cleavage; instead, these RNAs appear to be processed by the enzyme. The highly conserved ArcZ and RprA sRNAs, which each regulate a number of targets including *rpoS* (Majdalani et al., 2001; Mandin and Gottesman, 2010; Papenfort et al., 2015; Papenfort et al., 2009), provide cogent examples where RNase E converts a precursor into a stable shorter sRNA form (Fig. 4). For both ArcZ and RprA, the detected cleavages sites precisely match the RNase E consensus motif (Fig. 4 A-C) and are fully consistent with the size of the previously documented ~50 nt 3' species of these sRNAs (Argaman et al., 2001; Mandin and Gottesman, 2010; Papenfort et al., 2015; Papenfort et al., 2009). These 3' species accumulated to significantly higher levels than the primary sRNAs, in an Hfq-dependent manner (Fig. 4 D-E). When RNase E was inactivated for 30 min, these shorter ArcZ and RprA species became undetectable on northern blots (Fig. 4D-E; lane *rne*^{TS}, 44°C), suggesting a primary role for the enzyme in the processing event. To independently evaluate the function of RNase E in processing these sRNAs *in vivo*, each RNA was expressed from a plasmid-borne promoter subsequent to heat inactivation of the enzyme. While the full-length sRNAs accumulated under this condition, they were not converted into the short 3' species (Figs. 4G & S5B). These findings establish RNase E as a primary nuclease for generating functional short ArcZ and RprA, both of which regulate numerous *trans*-encoded target mRNAs (Papenfort et al., 2015; Papenfort et al., 2009).

Hfq-dependent regulatory RNA can also originate from other types of precursors, such as polycistronic tRNA transcripts. One such precursor is the sRNA sponge *leuZ*-3'ETS (Lalaouna et al., 2015) which was suggested to be processed by RNase E during *leuZ*-tRNA maturation (Li and Deutscher, 2002). Our TIER-seq data confirm that the 5' end of *leuZ*-3'ETS is generated by RNase E and pinpoints the cleavage site to an adenine 15 nt downstream of the mature *leuZ*-tRNA (Fig. 4C). Using the *rne*^{TS} strain, we observe RNase E to be essential for the production of this sRNA sponge (Fig. 4F). Together, these results argue for a major role of RNase E in maturing noncoding regulatory RNAs from different types of cellular transcripts.

Determinants of RNase E in sRNA processing

To understand how RNase E matures Hfq-associated sRNAs, we chose ArcZ for further characterization (Fig. 5A). Using the purified catalytic domain (NTD) of RNase E in combination with Hfq, we could readily reconstitute *in vitro* the release of 3' ArcZ (56 nt) from its 118 nt-long precursor (pre-ArcZ) prepared with T7 RNA polymerase (Fig. 5B). Within 3 min, the reaction produced the mature ArcZ fragment, which accumulated over time; and the cleavage occurred precisely at the expected sites identified by TIER-seq *in vivo* (Fig. 5D). However, in the absence of Hfq, RNase E rapidly hydrolyzed pre-ArcZ into fragments without producing 3' ArcZ (Fig. 5B). This suggests that Hfq plays a role in directing the correct processing of ArcZ by RNase E.

The maturation site in ArcZ in *Salmonella* and *E. coli* matches well with our TIER-seq derived RNase E consensus (GA↓U₊₁U₊₂U₊₃, Figs. 5A vs. 2F), featuring uridines (U₊₂ U₊₃) at the second and third position downstream of the cleavage site that are highly conserved in numerous enterobacterial species (Papenfort et al., 2009). Strikingly, changing U₊₂ to a disfavored G in the RNase E motif strongly diminished ArcZ processing by RNase E *in vitro* (Fig. 5 C-D), and processing was fully inhibited by further mutating U₊₃. By contrast, the same change at U₊₁ alone had little if any effect (Fig. 5 C-D). To explore if these findings have bearing on the maturation process *in vivo*, we expressed mutant ArcZ variants from inducible pBAD plasmids and analyzed the status of the ArcZ sRNA. Consistent with the *in vitro* results, the U₊₂→G₊₂ mutation strongly reduced the levels of 3' ArcZ in *Salmonella* (Fig. 5E), with further reductions upon additional mutation of the upstream (U₊₁U₊₂ → G₊₁G₊₂) and downstream (U₊₂U₊₃ → G₊₂G₊₃) uridines. Of note, the processing of ArcZ seems to be required for the regulation of its target *tpx* mRNA (see below).

The crucial roles of U₊₂ and Hfq in RNase E cleavage were also evident for the RprA sRNA (Fig. S5). Full-length RprA precursor (pre-RprA) was processed by RNase E *in vitro* at its internal seed sequence (GA↓A₊₁U₊₂U₊₃), producing mature RprA only in the presence of Hfq. Mutating U₊₂ alone significantly reduced the maturation of RprA by RNase E, which was fully abolished by changing both U₊₂U₊₃ to non-preferred guanines. The essentiality of U₊₂ in the RprA processing could also be demonstrated *in vivo* (Fig. S5C) as well as for cleavage of the *cfa* mRNA (Fig. S2E). Together, these mutational studies further validate our TIER-seq based prediction of U₊₂ as a key nucleotide for specific RNase E cleavage of cellular transcripts.

RNase E-dependent sRNA maturation is essential for target regulation

To consider RNase E as an sRNA maturation factor with functional consequences requires that its processing activity is essential for sRNA function. Demonstrating such a property requires first the development of a system in which processing of an sRNA precursor can be impeded without changing or losing the seed region. The ArcZ sRNA offers such a system: Exploiting our finding that mutation of the crucial U₊₂ in the RNase E motif of ArcZ abolished cleavage enabled us to produce pre-ArcZ with diminished amounts of 3' ArcZ *in vivo* (Fig. 5). We examined the ability of the pre-ArcZ to repress the synthesis of Tpx (Fig. 5E), whose mRNA is targeted by the conserved seed region of ArcZ (Papenfort et al., 2009) (Figs. 6A, S6). While a 10 min expression of WT ArcZ down-regulated the *tpx* mRNA by 7-fold, the U₊₂→G₊₂ mutant (variant GAU[G]U) achieved only 3-fold repression, despite the higher levels of precursor (Fig. 5E). Additional mutation of an adjacent uridine (variants GA[G]U or GAU[G]) fully inhibited 3' ArcZ production and abrogated *tpx* regulation despite higher levels of the precursor, strongly suggesting that only the mature 3' ArcZ is the functional regulator.

According to previous work (Papenfort et al., 2009), the U₊₂U₊₃ residues in the RNase E site of ArcZ may not engage in base pairing with *tpx* (Fig. 6A). If they do at all, they might extend the duplex by two additional A:U pairs, which could be disrupted by the non-functional, locked GAU[G] variant of Pre-ArcZ. To rule out that the failure of the GAU[G] variant (ArcZ-GG) to repress *tpx* was simply due to insufficient base pairing, we introduced a compensatory AU→CC mutation in the *tpx*-GFP fusion (Tpx-CC), but again no regulation by the ArcZ-GG variant was observed (Fig. 6B). Likewise, the processing-deficient ArcZ-GG variant also failed to regulate the *sdaC* mRNA target, in

either its wild-type form or with a duplex-extending CC mutation (Fig. S6B). Thus, RNase E cleavage is essential for the production of functional ArcZ.

A likely explanation for ArcZ maturation to be essential for regulation is that the ArcZ seed may only become available for target pairing upon RNase E cleavage. To test this, we examined sRNA duplex formation with *tpx* mRNA *in vitro*. Electrophoretic mobility shift assays with radiolabeled sRNA showed that the mature 3' ArcZ binds to the target region of *tpx* mRNA (a 216-nt fragment containing 5'UTR and early coding sequence) with very high affinity ($K_D \approx 15$ nM; Fig. 6C); by contrast, a ~500-fold excess of pre-ArcZ over target was insufficient for full duplex formation (Fig. 6D), similar to the low affinity observed for pre0ArcZ binding to the *rpoS* mRNA (Soper et al., 2010). In addition, Hfq promotes formation of the sRNA-target duplex in the case of mature ArcZ, but less so for pre-ArcZ (Fig. S6 C-D). These results were further confirmed by reciprocal experiments with labeled *tpx* mRNA. Again, mature ArcZ readily bound to the target and formed a stable ArcZ-*tpx*-Hfq ternary complex (Fig. 6E), whereas excess of the pre-ArcZ RNA only competed with the *tpx*-Hfq complex formation and released free *tpx* mRNA. These combined *in vivo* and *in vitro* results show that the pre-ArcZ undergoes an RNase E-dependent maturation to activate ArcZ for repression of *tpx* and perhaps other targets. This demonstrates for the first time that RNase E cleavage is required to activate an Hfq-dependent sRNA.

DISCUSSION

Bacterial transcripts are generally short-lived (Bernstein et al., 2002; Chen et al., 2015) and subject to rapid turnover by cellular ribonucleases (Hui et al., 2014; Mackie, 2013). Disregarding tRNA and rRNA, gene expression and regulation typically take place at the level of primary transcripts bearing the native 5'PPP end. This is fundamentally different from higher eukaryotes where nearly all types of regulatory transcripts undergo processing and maturation as a prerequisite for function. Our identification of numerous conserved regulatory sRNAs that result from RNase E cleavage (Figs. 3 & S3) illustrates the complexity of the bacterial 'RNA degradome'. These increasing numbers of processing-derived RNA species (Chao and Vogel, 2016; Davis and Waldor, 2007; Deltcheva et al., 2011; Guo et al., 2014; Miyakoshi et al., 2015a) contrast with the general perception that cleaved bacterial transcripts are usually labile species of little biological relevance.

TIER-seq offers a generic approach for global analysis of processed transcripts and cleavage sites in living cells with single-nucleotide resolution, and for mechanistic understanding of ribonuclease activities at a systems-level. We have here employed a temperature-sensitive strain to transiently inactivate the endogenous RNase E, which minimizes the potentially confounding effects of 'non-native' conditions used in previous degradome studies where the nucleases were either genetically deleted (Linder et al., 2014), ectopically over-expressed (Schifano et al., 2014), or supplemented *in vitro* (Clarke et al., 2014). To circumvent the need of a thermosensitive mutant and temperature-induced transcriptomic changes (Table S7), future TIER-seq studies may benefit by using alternative means of transient nuclease inactivation such as small molecules (Kime et al., 2015), small inhibitory proteins (Kim et al., 2008; Lee et al., 2003), target-specific proteases (Cameron and Collins, 2014), or conditionally spliced inteins (Zeidler et al., 2004).

RNase E-dependent sRNA biogenesis and maturation in bacteria

We identify RNase E as a key factor for both the biogenesis of many 3'UTR-derived sRNAs and the maturation of active sRNAs from their noncoding precursors. This establishes RNase E cleavage as a second major pathway for the biogenesis of Hfq-dependent sRNAs (Fig. 7A). As compared to the canonical pathway of *de novo* transcription, this cleavage-based biogenesis may confer several advantages. RNase E can generate sRNAs from diverse origins including essentially all existing transcripts (Fig. 7A), greatly expanding the sRNA repertoire in the cell. This pathway could reduce regulatory overhead during evolution of new genes (Mattick, 2004), using the existing regulatory elements of the parental transcripts to control the expression of 3' derived sRNAs. Activating an internal seed sequence by sRNA precursor cleavage as shown here for ArcZ offers an additional layer of control in post-transcriptional regulation, for example, via an adaptor protein such as RapZ which facilitates specific RNase E cleavage in certain sRNAs (Göpel et al., 2013). Lastly, RNase E cleavage derived sRNAs carry a 5'P end which promotes mRNA target degradation (Bandyra et al., 2012; Chao and Vogel, 2016; Pfeiffer et al., 2009) and as a consequence, different regulation kinetics than translational control alone (Levine and Hwa, 2008).

The key role of RNase E in sRNA biogenesis mirrors the central role of this enzyme in mRNA target regulation by many Hfq-dependent sRNAs (Masse et al., 2003; Saramago et al., 2014; Vogel and Luisi, 2011). Importantly, target degradation was proposed to involve tripartite RNase E-based ribonucleoprotein complexes with sRNA and Hfq (Ikeda et al., 2011; Morita et al., 2005; Worrall et

al., 2008). Our results indicate that this complex may form to mediate the alternative biogenesis of sRNAs prior to their target decay. For example, an ArcZ-Hfq-RNase E complex must form in the course of ArcZ maturation from the Hfq-bound pre-ArcZ sRNA (Fig. 5A). In this respect, the Hfq-RNase E complex in bacteria could have a dual function: it processes precursor transcripts to stable mature sRNA, and guides the mature sRNA for target regulation.

A U₊₂ ruler-and-cut mechanism mediates specific RNase E cleavage

The hallmark of the RNase E consensus motif inferred from our *in vivo* map (Fig. 2) is a predominant uridine at 2 nt downstream of the cleavage sites (U₊₂) and we provide *in vivo* and *in vitro* evidence that the U₊₂ is crucial for RNase E cleavage. Analysis of the available crystal structure of RNase E-RNA complex shows that the enzyme interacts with RNA at +2 nt via a stable stacking interaction of the nucleobase with Phe₆₇ and Lys₁₁₂ (Callaghan et al., 2005; Mackie, 2013). However, this structure contains a non-cognate substrate with G₊₂, representing a stable RNA-binding conformation trapped at the pre-cleavage state. Why is a uridine at this position preferred for cleavage? By substituting the G₊₂ to U *in silico*, a molecular dynamics simulation analysis suggests that RNase E-RNA complex undergoes a conformational change favored by the presence of U₊₂, allowing us to propose a new model (Fig. 7B) whereby RNase E mediates specific cleavage using a U₊₂ ruler-and-cut mechanism. Simulations of the pre-cleavage state show that U₊₂ was tightly bound in a crevice of the protein formed by the backbone of the Lys₁₁₂Gly₁₁₃Ala₁₁₄Ala₁₁₅ loop, and the Lys₁₁₂ side chain resulting in a binding pocket that favors uracil (uracil pocket, Fig. S7 A-C and supplemental discussion). Importantly, the presence of the cognate U₊₂ promotes a distortion of the phosphodiester backbone angles at the cleavage site 2 nt upstream. The new conformation of the scissile phosphate may closely resemble, with slight deviation, the pseudo trigonal bipyramidal geometry that facilitates in line nucleophilic attack of scissile phosphate (Figs. 7B, S7 D-F). While we have shown here that mutating U₊₂ in RNA abolishes cleavage, mutation of the critical Lys₁₁₂ also abrogates RNase E cleavage of cognate substrates (Callaghan et al., 2005). The high conservation of residues forming the uracil pocket (e.g. Phe₆₇ and Lys₁₁₂) indicates this may be a conserved mechanism for the RNase E protein family.

The uridine ruler-and-cut mechanism is also employed by other endoribonucleases including an unrelated human nuclease RNase L. RNase L recognizes uridine in single-stranded RNAs and cleaves 2 nt downstream (Han et al., 2014), whereas RNase E cuts 2 nt upstream (RNase E) due to

different dimeric structure arrangements. Interestingly, a fraction of RNase E sites contains C₊₂ (Fig. 2F), indicating that RNase E displays certain degree of flexibility by accepting a cytosine in the absence of other specificity signals. Indeed, *in vitro* experiments with poly(A) RNA demonstrate that C₊₂ can serve as a cleavage signal (Kaberdin, 2003), suggesting that RNase E may distinguish the smaller pyrimidine from purine bases by steric hindrance (Figs. 7C & S7 A-B). Nevertheless, U₊₂ is the preferred signal (Fig. 2F), likely because its C₄ oxygen possesses hydrogen bonding potential with RNase E (Fig. S7B). In addition, some flexibility of RNase E is reflected near the cleavage sites, as RNase E frequently cuts 1 nt upstream or downstream of the determined cleavage site. To compensate for this, short stretches of uridines (1-4 Us) are often found at the +2 positions, which may serve to reinforce RNase E recognition and cleavage (e.g. ArcZ, Fig. 5).

Our identification of crucial U₊₂ residues for RNase E specific cleavage enables straightforward mutations of individual cleavage sites of interest, instead of global inactivation of the enzyme. This will aid the molecular investigation of 3'UTR-derived sRNAs and of RNase E-mediated post-transcriptional regulations, not only in the Hfq regulon but also for the recently discovered class of ProQ-associated sRNAs (Smirnov et al., 2016) many of which might be RNase E targets, too. This information may also help design novel CRISPR-Cas or antisense RNA-based synthetic tools to activate gene expression by specifically blocking a cleavage site, and to engineer stable mRNAs for better gene expression.

ACKNOWLEDGEMENTS

We thank L. Bossi for *Salmonella* strain *rne*-3071; K. Bandyra for purified RNase E protein; S. Gorski for editing; K. McDowall for sharing data; T. Yano and B. Plaschke for technical assistance. This study was funded by DFG (Vo875/14-1) and BioSysNet grants. BFL is supported by the Wellcome Trust. KP was supported by the Human Frontiers Science Program (CDA00024/2016-C).

AUTHOR CONTRIBUTION

YC and JV conceived the research; YC, NS, CC, MS, KP conducted experiments; YC, LL, KUF and BFL analyzed data; DG and HJW performed MD simulations; RR performed RNA-seq; YC, BFL and JV wrote the manuscript.

EXPERIMENTAL PROCEDURES

Full methods are described in the Supplemental Experimental Procedures; so are details of bacterial strains, plasmids, and oligonucleotides.

Transient inactivation of RNase E

The *Salmonella rne*^{TS} strains refer to *rne*-3071 and its isogenic WT control previously established in (Figuroa-Bossi et al., 2009). Bacteria were grown in LB-Lennox medium at 28°C to an OD₆₀₀ of 2, then shifted to 44°C for 30 min to inactivate RNase E.

RNA-seq and data analysis

cDNA libraries were constructed following a standard protocol (Chao et al., 2012; Westermann et al., 2016). Briefly, RNA was polyadenylated at 3' end, and ligated to an adaptor at 5' end after treatment with tobacco acid pyrophosphatase. First-strand cDNA was synthesized using oligo(dT)-adaptor and M-MLV reverse transcriptase. The linear amplified cDNAs were multiplexed and sequenced using Illumina HiSeq. Reads were mapped to *Salmonella* genome using READemption; 5' ends coverage was visualized in IGB. The RNase E sites, which are depleted 5' ends in the *rne*^{TS} samples relative to WT at 44°C, were identified using DESeq2. The sequencing data have been deposited in the GEO database (GSE81869).

REFERENCES

- Apirion, D., and Lassar, A.B. (1978). A conditional lethal mutant of *Escherichia coli* which affects the processing of ribosomal RNA. *J Biol Chem* *253*, 1738-1742.
- Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H., and Altuvia, S. (2001). Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr Biol* *11*, 941-950.
- Arraiano, C.M., Andrade, J.M., Domingues, S., Guinote, I.B., Malecki, M., Matos, R.G., Moreira, R.N., Pobre, V., Reis, F.P., Saramago, M., *et al.* (2010). The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev* *34*, 883-923.
- Bandyra, K.J., Said, N., Pfeiffer, V., Gorna, M.W., Vogel, J., and Luisi, B.F. (2012). The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol Cell* *47*, 943-953.
- Barquist, L., and Vogel, J. (2015). Accelerating Discovery and Functional Analysis of Small RNAs with New Technologies. *Annu Rev Genet* *49*, 367-394.
- Belasco, J.G. (2010). All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat Rev Mol Cell Biol* *11*, 467-478.
- Bernstein, J.A., Khodursky, A.B., Lin, P.H., Lin-Chao, S., and Cohen, S.N. (2002). Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci U S A* *99*, 9697-9702.
- Bernstein, J.A., Lin, P.H., Cohen, S.N., and Lin-Chao, S. (2004). Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc Natl Acad Sci U S A* *101*, 2758-2763.
- Bessarab, D.A., Kaberdin, V.R., Wei, C.L., Liou, G.G., and Lin-Chao, S. (1998). RNA components of *Escherichia coli* degradosome: evidence for rRNA decay. *Proc Natl Acad Sci U S A* *95*, 3157-3161.
- Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G., and Luisi, B.F. (2005). Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* *437*, 1187-1191.
- Cameron, D.E., and Collins, J.J. (2014). Tunable protein degradation in bacteria. *Nature biotechnology* *32*, 1276-1281.
- 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.
- 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.
- Chen, H., Shiroguchi, K., Ge, H., and Xie, X.S. (2015). Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*. *Mol Syst Biol* *11*, 781.
- 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.
- Davis, B.M., and Waldor, M.K. (2007). RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol Microbiol* *65*, 373-385.
- De Lay, N., Schu, D.J., and Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J Biol Chem* *288*, 7996-8003.
- Del Campo, C., Bartholomaeus, 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.
- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* *471*, 602-607.
- Delvillani, F., Papiani, G., Deho, G., and Briani, F. (2011). S1 ribosomal protein and the interplay between translation and mRNA decay. *Nucleic Acids Res* *39*, 7702-7715.
- Desnoyers, G., Bouchard, M.P., and Masse, E. (2013). New insights into small RNA-dependent translational regulation in prokaryotes. *Trends Genet* *29*, 92-98.
- Dimastrogiovanni, D., Frohlich, K.S., Bandyra, K.J., Bruce, H.A., Hohensee, S., Vogel, J., and Luisi, B.F. (2014). Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *Elife* *3*.
- 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.
- Feng, L., Rutherford, S.T., Papenfort, K., Bagert, J.D., van Kessel, J.C., Tirrell, D.A., Wingreen, N.S., and Bassler, B.L. (2015). A *qrr* noncoding RNA deploys four different regulatory mechanisms to optimize quorum-sensing dynamics. *Cell* *160*, 228-240.
- Figueroa-Bossi, N., Valentini, M., Malleret, L., Fiorini, F., and Bossi, L. (2009). Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev* *23*, 2004-2015.
- Frohlich, K.S., Papenfort, K., Fekete, A., and Vogel, J. (2013). A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *EMBO J* *32*, 2963-2979.

- Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2005). Intermediate states of ribonuclease III in complex with double-stranded RNA. *Structure* *13*, 1435-1442.
- Göpel, Y., Papenfort, K., Reichenbach, B., Vogel, J., and Gorke, B. (2013). Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev* *27*, 552-564.
- Guo, M.S., Updegrove, T.B., Gogol, E.B., Shabalina, S.A., Gross, C.A., and Storz, G. (2014). MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev* *28*, 1620-1634.
- Han, Y., Donovan, J., Rath, S., Whitney, G., Chitrakar, A., and Korennykh, A. (2014). Structure of human RNase L reveals the basis for regulated RNA decay in the IFN response. *Science* *343*, 1244-1248.
- Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., Backofen, R., and Vogel, J. (2016). Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *EMBO J* *35*, 991-1011.
- Huang, H., Liao, J., and Cohen, S.N. (1998). Poly(A)- and poly(U)-specific RNA 3' tail shortening by E. coli ribonuclease E. *Nature* *391*, 99-102.
- Hui, M.P., Foley, P.L., and Belasco, J.G. (2014). Messenger RNA degradation in bacterial cells. *Annu Rev Genet* *48*, 537-559.
- Ikeda, Y., Yagi, M., Morita, T., and Aiba, H. (2011). Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli*. *Mol Microbiol* *79*, 419-432.
- Jain, C., and Belasco, J.G. (1995). RNase E autoregulates its synthesis by controlling the degradation rate of its own mRNA in *Escherichia coli*: unusual sensitivity of the rne transcript to RNase E activity. *Genes Dev* *9*, 84-96.
- Joanny, G., Le Derout, J., Brechemier-Baey, D., Labas, V., Vinh, J., Regnier, P., and Hajnsdorf, E. (2007). Polyadenylation of a functional mRNA controls gene expression in *Escherichia coli*. *Nucleic Acids Res* *35*, 2494-2502.
- Kaberdin, V.R. (2003). Probing the substrate specificity of *Escherichia coli* RNase E using a novel oligonucleotide-based assay. *Nucleic Acids Res* *31*, 4710-4716.
- Kim, K.S., Manasherob, R., and Cohen, S.N. (2008). YmdB: a stress-responsive ribonuclease-binding regulator of E. coli RNase III activity. *Genes Dev* *22*, 3497-3508.
- Kim, V.N. (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol* *6*, 376-385.
- Kime, L., Clarke, J.E., Romero, A.D., Grasby, J.A., and McDowall, K.J. (2014). Adjacent single-stranded regions mediate processing of tRNA precursors by RNase E direct entry. *Nucleic Acids Res* *42*, 4577-4589.
- Kime, L., Vincent, H.A., Gendoo, D.M., Jourdan, S.S., Fishwick, C.W., Callaghan, A.J., and McDowall, K.J. (2015). The first small-molecule inhibitors of members of the ribonuclease E family. *Scientific reports* *5*, 8028.
- Koo, J.T., Alleyne, T.M., Schiano, C.A., Jafari, N., and Lathem, W.W. (2011). Global discovery of small RNAs in *Yersinia pseudotuberculosis* identifies *Yersinia*-specific small, noncoding RNAs required for virulence. *Proc Natl Acad Sci U S A* *108*, E709-717.
- Lalaouna, D., Carrier, M.C., Semsey, S., Brouard, J.S., Wang, J., Wade, J.T., and Masse, E. (2015). A 3' external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. *Mol Cell* *58*, 393-405.
- Lee, K., Zhan, X., Gao, J., Qiu, J., Feng, Y., Meganathan, R., Cohen, S.N., and Georgiou, G. (2003). RraA, a protein inhibitor of RNase E activity that globally modulates RNA abundance in E. coli. *Cell* *114*, 623-634.
- Levine, E., and Hwa, T. (2008). Small RNAs establish gene expression thresholds. *Curr Opin Microbiol* *11*, 574-579.
- Li, Z., and Deutscher, M.P. (2002). RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA* *8*, 97-109.
- Linder, P., Lemeille, S., and Redder, P. (2014). Transcriptome-wide analyses of 5'-ends in RNase J mutants of a gram-positive pathogen reveal a role in RNA maturation, regulation and degradation. *PLoS Genet* *10*, e1004207.
- Mackie, G.A. (1991). Specific endonucleolytic cleavage of the mRNA for ribosomal protein S20 of *Escherichia coli* requires the product of the ams gene in vivo and in vitro. *J Bacteriol* *173*, 2488-2497.
- Mackie, G.A. (2013). RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* *11*, 45-57.
- Madhugiri, R., Basineni, S.R., and Klug, G. (2010). Turn-over of the small non-coding RNA RprA in E. coli is influenced by osmolarity. *Mol Genet Genomics* *284*, 307-318.
- Majdalani, N., Chen, S., Murrow, J., St John, K., and Gottesman, S. (2001). Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol Microbiol* *39*, 1382-1394.
- Mandin, P., and Gottesman, S. (2010). Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* *29*, 3094-3107.
- Masse, E., Escorcia, F.E., and Gottesman, S. (2003). Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* *17*, 2374-2383.
- Mattick, J.S. (2004). RNA regulation: a new genetics? *Nat Rev Genet* *5*, 316-323.
- McClain, W.H., Guerrier-Takada, C., and Altman, S. (1987). Model substrates for an RNA enzyme. *Science* *238*, 527-530.
- McDowall, K.J., Kaberdin, V.R., Wu, S.W., Cohen, S.N., and Lin-Chao, S. (1995). Site-specific RNase E cleavage of oligonucleotides and inhibition by stem-loops. *Nature* *374*, 287-290.

- 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.
- Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y.E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., and Margalit, H. (2016). Global Mapping of Small RNA-Target Interactions in Bacteria. *Mol Cell* 63, 884-897.
- Miyakoshi, M., Chao, Y., and Vogel, J. (2015a). Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J* 34, 1478-1492.
- Miyakoshi, M., Chao, Y., and Vogel, J. (2015b). Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Curr Opin Microbiol* 24, 132-139.
- Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V.R., and Blasi, U. (2003). Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* 9, 1308-1314.
- Morita, T., Maki, K., and Aiba, H. (2005). RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* 19, 2176-2186.
- 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.
- Ow, M.C., and Kushner, S.R. (2002). Initiation of tRNA maturation by RNase E is essential for cell viability in *E. coli*. *Genes Dev* 16, 1102-1115.
- Papenfort, K., Espinosa, E., Casadesus, J., and Vogel, J. (2015). Small RNA-based feedforward loop with AND-gate logic regulates extrachromosomal DNA transfer in *Salmonella*. *Proc Natl Acad Sci U S A* 112, E4772-4781.
- Papenfort, K., Said, N., Welsink, T., Lucchini, S., Hinton, J.C., and Vogel, J. (2009). Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* 74, 139-158.
- Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C.K., and Vogel, J. (2013). Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* 153, 426-437.
- Papenfort, K., and Vanderpool, C.K. (2015). Target activation by regulatory RNAs in bacteria. *FEMS Microbiol Rev* 39, 362-378.
- Patel, A.M., and Dunn, S.D. (1992). RNase E-dependent cleavages in the 5' and 3' regions of the *Escherichia coli* unc mRNA. *J Bacteriol* 174, 3541-3548.
- Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J.C., and Vogel, J. (2009). Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16, 840-846.
- Regnier, P., and Hajsnsdorf, E. (1991). Decay of mRNA encoding ribosomal protein S15 of *Escherichia coli* is initiated by an RNase E-dependent endonucleolytic cleavage that removes the 3' stabilizing stem and loop structure. *J Mol Biol* 217, 283-292.
- Roy, M.K., and Apirion, D. (1983). Purification and properties of ribonuclease E, an RNA-processing enzyme from *Escherichia coli*. *Biochim Biophys Acta* 747, 200-208.
- Saramago, M., Barria, C., Dos Santos, R.F., Silva, I.J., Pobre, V., Domingues, S., Andrade, J.M., Viegas, S.C., and Arraiano, C.M. (2014). The role of RNases in the regulation of small RNAs. *Curr Opin Microbiol* 18, 105-115.
- 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.
- Schifano, J.M., Vvedenskaya, I.O., Knoblauch, J.G., Ouyang, M., Nickels, B.E., and Woychik, N.A. (2014). An RNA-seq method for defining endoribonuclease cleavage specificity identifies dual rRNA substrates for toxin MazF-mt3. *Nat Commun* 5, 3538.
- Smirnov, A., Forstner, K.U., Holmqvist, E., Otto, A., Gunster, R., Becher, D., Reinhardt, R., and Vogel, J. (2016). Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc Natl Acad Sci U S A* 113, 11591-11596.
- 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.
- Storz, G., Vogel, J., and Wassarman, K.M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 43, 880-891.
- Tree, J.J., Granneman, S., McAteer, S.P., Tollervey, D., and Gally, D.L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. *Mol Cell* 55, 199-213.
- Viegas, S.C., Mil-Homens, D., Fialho, A.M., and Arraiano, C.M. (2013). The virulence of *Salmonella enterica* Serovar Typhimurium in the insect model *Galleria mellonella* is impaired by mutations in RNase E and RNase III. *Appl Environ Microbiol* 79, 6124-6133.
- Viegas, S.C., Pfeiffer, V., Sittka, A., Silva, I.J., Vogel, J., and Arraiano, C.M. (2007). Characterization of the role of ribonucleases in *Salmonella* small RNA decay. *Nucleic Acids Res* 35, 7651-7664.
- Vogel, J., and Luisi, B.F. (2011). Hfq and its constellation of RNA. *Nat Rev Microbiol* 9, 578-589.
- Wagner, E.G. (2013). Cycling of RNAs on Hfq. *RNA Biol* 10, 619-626.
- Wagner, E.G., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv Genet* 90, 133-208.

- Westermann, A.J., Forstner, K.U., Amman, F., Barquist, L., Chao, Y., Schulte, L.N., Muller, L., Reinhardt, R., Stadler, P.F., and Vogel, J. (2016). Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* 529, 496-501.
- Worrall, J.A., Gorna, M., Crump, N.T., Phillips, L.G., Tuck, A.C., Price, A.J., Bavro, V.N., and Luisi, B.F. (2008). Reconstitution and analysis of the multienzyme *Escherichia coli* RNA degradosome. *J Mol Biol* 382, 870-883.
- Zeidler, M.P., Tan, C., Bellaiche, Y., Cherry, S., Hader, S., Gayko, U., and Perrimon, N. (2004). Temperature-sensitive control of protein activity by conditionally splicing inteins. *Nature biotechnology* 22, 871-876.
- Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G., and Gottesman, S. (2003). Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* 50, 1111-1124.

FIGURE LEGENDS

Figure 1. Global mapping of endogenous RNase E cleavage sites in *Salmonella* using TIER-seq

A) Scheme of the TIER-seq approach. Endogenous cleavage sites were identified by analyzing the 5' ends of RNase E cleavage products (purple) in the WT and *rne*^{TS} strains at the non-permissive temperature (44°C). Total RNA from WT and *rne*^{TS} was converted to cDNAs and sequenced; the 5' ends depleted in the *rne*^{TS} libraries at 44°C indicate the RNase E cleavage sites (e.g. purple U).

B,C) Global analysis of 5' end profile at the permissive temperature 28°C (B) and non-permissive temperature 44°C (C). The plots show the read counts for every 5' base in WT samples and the relative fold-change compared to *rne*^{TS} samples. Candidate RNase E cleavage sites that show >3-fold depletion in *rne*^{TS} samples ($P < 0.05$, FDR < 0.05) are colored in red.

D) TIER-seq captures known RNase E cleavage sites with single-nucleotide resolution. TS indicate the *rne*^{TS} samples. R1 and R2 are two biological replicates. The major RNase E sites are marked by red arrowheads and bold lettering; secondary cleavage sites are indicated by open arrowheads. The ORF or mature RNAs are shadowed by grey boxes. See also Figs. S1, S2.

Figure 2. Systems-wide analysis of cleavage sites reveals a consensus for RNase E

A) Classification of all RNase E cleavage sites. The proportion (%) of all sites mapped within a category is shown. See also Table S1.

B) The number of cleavage sites mapped per mRNA gene.

C) The distribution of RNase E cleavage frequencies in mRNAs (RPKM > 10). See also Fig. S1.

D) Sequences at the RNase E sites are less structured. Minimal folding energy (MFE) was calculated for each 25 nt using a sliding window, and compared to randomly shuffled sequences. Median Z-score is shown as a bold line; dotted lines indicate the upper and lower quartile.

E) Distribution of AU-content at the RNase E cleavage sites. Dashed line indicates the cleavage site (+1 nt).

F) The RNase E consensus motif based on alignment of all mapped cleavage sites. Error bars indicate 95% confidence intervals. See also Figs. S1 & S2.

Figure 3. RNase E cleaves mRNAs to produce 3'UTR-derived sRNAs

- A,B) Distribution of RNase E cleavage sites in mRNAs relative to their start codon (**A**) or stop codon (**B**). The grey lines in the lower panel indicate the distribution of consensus motif based on genomic sequence.
- C) RNase E and Hfq are required for the biogenesis of 3'UTR-derived sRNAs. WT and Δhfq strains were grown at 37°C to OD₆₀₀ of 2. The location of sRNAs (red arrows) and host genes are shown in the lower panel. Promoters (where available) and terminators are shown. The 5S rRNA served as loading control (Fig. S2C). See also Figs. S2, S3 & S4.

Figure 4. RNase E cleaves noncoding RNAs to release 3' mature sRNAs

- A-C) RNase E cleavage sites are identified in the ArcZ sRNA (**A**), RprA (**B**), and 3'ETS^{leuZ} (**C**). The major sites are marked by red arrowheads and bold lettering, whereas the minor sites are indicated by open arrowheads. See also Fig. S5-S6.
- D-F) RNase E is required for the processing of ArcZ (**D**), RprA (**E**), and 3'ETS^{leuZ} (**F**). Open arrowheads indicate precursor fragments and filled arrowheads indicate processed mature species. * indicates longer precursors of polycistronic LeuZ-tRNA fragments; 5S loading controls, see Fig. S2C.
- G) The maturation of ArcZ is dependent on RNase E activity. Expression of the full-length ArcZ precursor (pre-ArcZ) was induced by L-arabinose.

Figure 5. RNase E mediates the maturation of ArcZ sRNA *in vitro* and *in vivo*

- A) Alignment of ArcZ sequence. Conservation scores are plotted below the sequences, and the conserved seed is colored in green.
- B) Reconstitution of ArcZ maturation *in vitro*. Full-length pre-ArcZ RNA was incubated with RNase E in the presence or absence of Hfq. RNA was analyzed by northern blotting with an oligo antisense to the mature ArcZ. The lower set shows mature ArcZ signals with longer exposure.
- C) Mutation of RNase E cleavage site. Variants of ArcZ precursors were incubated with Hfq, and then subjected to RNase E cleavage. The lower set shows mature ArcZ signals with longer exposure.
- D) Primer extension to map the RNase E cleavage sites in ArcZ *in vitro*.

E) Validation of RNase E motif in ArcZ *in vivo*. See also Fig. S5, S6.

Figure 6. Maturation of ArcZ sRNA is essential for target regulation

- A) Established basepair interactions between ArcZ and *tpx* mRNA (Papenfort et al., 2009). The major cleavage site in ArcZ is indicated.
- B) Western blot detection of GFP levels. GFP was fused with *tpx* 5'UTR; and the introduced mutations are shown in A. "WT" refers to WT full-length ArcZ, "mat" refers to mature ArcZ, and "GG" refers to the GAU[GG] variant of ArcZ. GroEL served as loading control.
- C) Direct interaction of *tpx* with mature ArcZ by EMSA. Radiolabeled mature ArcZ was incubated with increasing concentration of *tpx* mRNA in the presence of Hfq (40 nM). The gel was resized, see Fig. S6.
- D) Direct interaction of *tpx* with pre-ArcZ by EMSA.
- E) Mature ArcZ was co-shifted with *tpx* mRNA. Radiolabeled *tpx* mRNA was incubated with increasing concentration of pre-ArcZ or mature ArcZ (0, 6, 25, 100, 400, 2000 nM) in the presence of 40nM Hfq. See also Fig. S6.

Figure 7. Mechanism of RNase E cleavage and an alternative sRNA biogenesis pathway

- A) RNase E cleavage constitutes a major sRNA biogenesis pathway in bacteria.
- B) Proposed model for the +2 uridine ruler-and-cut mechanism of specific RNase E cleavage. The scissile phosphate is attacked hydrolytically by a water molecule (not shown) that is coordinated by the magnesium ion bound by the carboxylates of D₃₄₆ and D₃₀₃. Stacking interactions (between F₆₇ and K₁₁₂) and hydrogen bonding (with the K₁₁₂GAA loop) with the base at position +2, favor uridine at this position. The interactions are predicted to help orientate the phosphate backbone into a geometry that would facilitate cleavage at the scissile phosphate. See also Fig. S7.