Domain swapping between homologous bacterial small RNAs dissects processing and Hfq binding determinants and uncovers an aptamer for conditional RNase E cleavage

Yvonne Göpel, Muna Ayesha Khan and Boris Görke[*]

Department of Microbiology, Immunobiology and Genetics, Max F. Perutz Laboratories (MFPL), University of Vienna, Vienna Biocenter (VBC), Dr Bohr-Gasse 9, 1030 Vienna, Austria

ABSTRACT

In *E. coli*, small RNA GlmZ activates the *glmS* mRNA by base-pairing in an Hfq dependent manner. When not required, GlmZ is bound by adaptor protein RapZ and recruited to RNase E, which cleaves GlmZ in its base-pairing sequence. Small RNA GlmY counteracts cleavage of GlmZ by sequestration of RapZ. Although both sRNAs are highly homologous, only GlmZ specifically binds Hfq and undergoes cleavage by RNase E. We used domain swapping to identify the responsible modules. Two elements, the 3′ terminal oligo(U) stretch and the base-pairing region enable GlmZ to interact with Hfq. Accordingly, Hfq inhibits cleavage of GlmZ, directing it to base-pairing. Intriguingly, the central stem loop of GlmZ is decisive for cleavage, whereas the sequence comprising the actual cleavage site is dispensable. Assisted by RapZ, RNase E cleaves any RNA fused to the 3′ end of this module. These results suggest a novel mode for RNase E recognition, in which one of the required handholds in the substrate is replaced by an RNA binding protein. This device can generate RNAs of interest in their 5′ monophosphorylated form on demand. As these species are rapidly degraded, this tool allows to regulate gene expression post-transcriptionally by modulation of RapZ levels.

INTRODUCTION

Small RNAs (sRNAs) regulate protein synthesis at the post-transcriptional level. A major class of ribo-regulators in bacteria are *trans*-encoded sRNAs, which regulate functions encoded at distinct locations. Many *trans*-encoded sRNAs base-pair with mRNAs thereby modulating translation, transcription elongation or RNA decay. A minor group of sRNAs acts through sequestration of RNA or DNA binding proteins by molecular mimicry mechanisms. Although quite heterogeneous in sequence and structure, base-pairing sRNAs share discernible modules ([1]). They carry an intrinsic terminator composed of a stem loop followed by an oligo(U) sequence at the 3′ end. Furthermore, sRNAs contain a single-stranded seed region, which is part of the base-pairing site and initially contacts target RNAs. Binding sites for Hfq present a third essential feature, at least for enterobacterial sRNAs. Additional secondary structures may help to align modules or provide binding sites for further factors.

RNA chaperone Hfq and endoribonuclease RNase E, emerged as global players intimately connected with the function of base-pairing sRNAs in Gram-negative bacteria ([2,3]). Hfq of *E. coli* assembles into a homo-hexameric ring providing at least three distinct RNA binding sites. The distal site binds AAN sequence repeats present in mRNAs that are targeted by sRNAs ([4–6]). The proximal face preferentially binds single-stranded oligo(U) sequences and for some sRNAs an interaction with the 3′ terminal oligo(U) stretch has been shown ([7,8]). Additionally, single-stranded U-rich regions from the sRNA body may wrap around the rim of the Hfq hexamer and the outward protruding C-terminal tail of Hfq might support sRNA association ([7–10]). Thereby, Hfq protects sRNAs from degradation and facilitates annealing with RNA targets. sRNAs compete for Hfq suggesting that they share binding sites ([11–13]). For a couple of sRNAs binding to the distal site of Hfq by internal adenosine-rich sequences has been proposed ([12,14]). Thus, sRNAs may contact Hfq in different manners ([12,15]), suggesting that the features conferring binding must be determined for each sRNA individually.

RNase E initiates bulk RNA decay via two different pathways ([16–18]). For efficient cleavage each protomer within the principal RNase E dimer must contact the substrate. In the 5′ tethering pathway, the 5′ monophosphate of the RNA interacts with the 5′ sensor domain of one protomer,

---

[*]To whom correspondence should be addressed. Tel: +43 1 4277 54603; Fax: +43 1 4277 9546; Email: boris.goerke@univie.ac.at

which stimulates cleavage of the substrate by the second protomer. Therefore, RNA is cleaved more efficiently when 5′ monophosphorylated. These species are generated from primary transcripts by pyrophosphohydrolase RppH ([19]). The second pathway, referred to as 'direct entry', operates on 5′ triphosphorylated transcripts. This mechanism apparently requires that both RNA binding channels of the principal dimer interact with distinct single-stranded regions in the substrate ([20]). RNase E preferentially cleaves single-stranded AU-rich regions and generates 3′ hydroxylated and 5′ monophosphorylated cleavage products, the latter further accelerate subsequent RNase E attacks by 5′ tethering. Cleavages are specific, but the reason is poorly understood. RNase E is composed of an N-terminal catalytic domain and a C-terminal scaffold for binding of additional proteins involved in RNA decay, collectively forming the degradosome ([21,22]). Notably, Hfq recruits RNase E by interaction with its C-terminus contributing to sRNA-mediated silencing by coupled degradation of the sRNA with its target ([3,23]). Interestingly, by presenting a 5′ terminal monophosphate group an sRNA was shown to allosterically activate RNase E, allowing cleavage of base-paired substrates lacking RNase E recognition sites ([24]). On the other hand, sRNAs can stabilize transcripts by sequestration of RNase E target sites upon base-pairing ([25,26]). Finally, some sRNAs themselves are cleaved by RNase E when not paired. This may inactivate the sRNA or lead to species with even higher regulatory potential ([27,28]). Evidence is accumulating that cleavage of sRNAs by RNase E could also be regulated, providing a basis for post-transcriptional control of sRNAs ([29,30]).

An intriguing mechanism employing regulated decay of an sRNA by RNase E has been unveiled in the GlmY/GlmZ sRNA circuit in *E. coli* ([31]). GlmZ activates the mRNA encoding glucosamine-6-phosphate (GlcN6P) synthase GlmS through sequestration of an anti-Shine–Dalgarno sequence upon base-pairing (Figure 1A right; ([32,33])). GlmS synthesizes GlcN6P, which is required for cell envelope biogenesis. When GlcN6P levels are high and GlmS is dispensable, RNase E inactivates GlmZ by cleaving its base-pairing region. However, this cleavage requires adaptor protein RapZ, which binds GlmZ and recruits RNase E by interaction with its catalytic domain (Figure 1A left; ([30])). GlmZ cleavage is antagonized by the homologous sRNA GlmY, which accumulates upon GlcN6P shortage ([33,34]). Owing to its similarity to GlmZ, GlmY sequesters RapZ. As a result, GlmZ stays unprocessed and up-regulates GlmS, thereby restoring GlcN6P homeostasis (Figure 1A, right; ([30])).

Although RapZ specifically binds both sRNAs, only GlmZ is processed by RNase E ([30]). As expected, Hfq binds GlmZ with high affinity. In contrast, GlmY is only weakly bound ([30]). These differences were surprising given that both sRNAs share 66% sequence identity and exhibit strikingly similar secondary structures (Supplementary Figure S1; Figure 1B). To provide insight, we dissected the molecular requirements for Hfq binding and RNase E cleavage by swapping sequences between GlmY and GlmZ. This analysis uncovered modules crucial for Hfq binding, residing in the 3′ end of GlmZ. As one of these modules overlaps the RNase E target site, binding of Hfq confers protection from

cleavage. Most importantly, the current work reveals an aptamer suited to induce RNase E-dependent cleavage in any RNA fused to its 3′ end. This tool generates 5′ monophosphorylated RNAs on demand, by modulation of RapZ availability in the cell. Using a model mRNA, we demonstrate that this aptamer allows regulating mRNA and protein abundance at the post-transcriptional level, providing a novel tool for synthetic biology.

MATERIALS AND METHODS

Growth conditions, plasmids and strains

Escherichia coli strains and plasmids used in this study are described in Supplementary Tables S1 and S2, respectively. Oligonucleotides are listed in Supplementary Table S3. Bacteria were grown in LB medium under shaking at 165 rpm and antibiotics were added when necessary: ampicillin (100 μg/ml), chloramphenicol (15 μg/ml), kanamycin (30 μg/ml), spectinomycin (75 μg/ml), streptomycin (100 μg/ml). General transduction by phage T4GT7 was used to transfer established alleles tagged with an antibiotic resistance marker between strains ([35]). Resistance markers flanked by FLP recombinase recognition sites were removed using the temperature-sensitive FLP recombinase delivery plasmid pCP20 ([36]). The F' plasmid of strain BMH71-18 was moved into strain Z897 by conjugation using a double selection for streptomycin resistance and proline prototrophy. Recombinant plasmids were constructed using strains DH5α or XL1-blue as described under 'Supplementary Material and Methods'.

In vitro transcription and labelling of small RNAs

Generation of radioactively labelled RNAs by *in vitro* transcription is described under Supplementary Materials and Methods.

Purification of Strep-RapZ and His$_6$-RNase E-N (1–529)

Strep-RapZ and the His$_6$-tagged catalytic domain of RNase E were purified as described previously (([30,37]); for details see Supplementary Materials and Methods).

Electrophoretic mobility shift assays (EMSAs)

Protein/RNA binding assays were carried out in 1x structure buffer (10 mM Tris-HCl pH 7.0, 100 mM KCl, 10 mM MgCl$_2$) in a volume of 10 μl ([30]). Four nM of the α-^{32}P-UTP-labelled heat-denatured RNAs were mixed with 1 μg yeast tRNA (Ambion) and various amounts of the respective proteins (Hfq or Strep-RapZ) and incubated for 30 min at 30°C. The protein dilutions were prepared in 1x structure buffer. Following incubation, 2 μl loading buffer (50% glycerol, 0.5x TBE, 0.2% bromophenol blue) were added and samples were subsequently separated by non-denaturing gel electrophoresis (8% polyacrylamide, 1x TBE) at 300 V for 3 h at 4°C using 0.5x TBE as running buffer. Gels were dried and analysed by phosphorimaging (TyphoonTM FLA 9500, GE Healthcare).
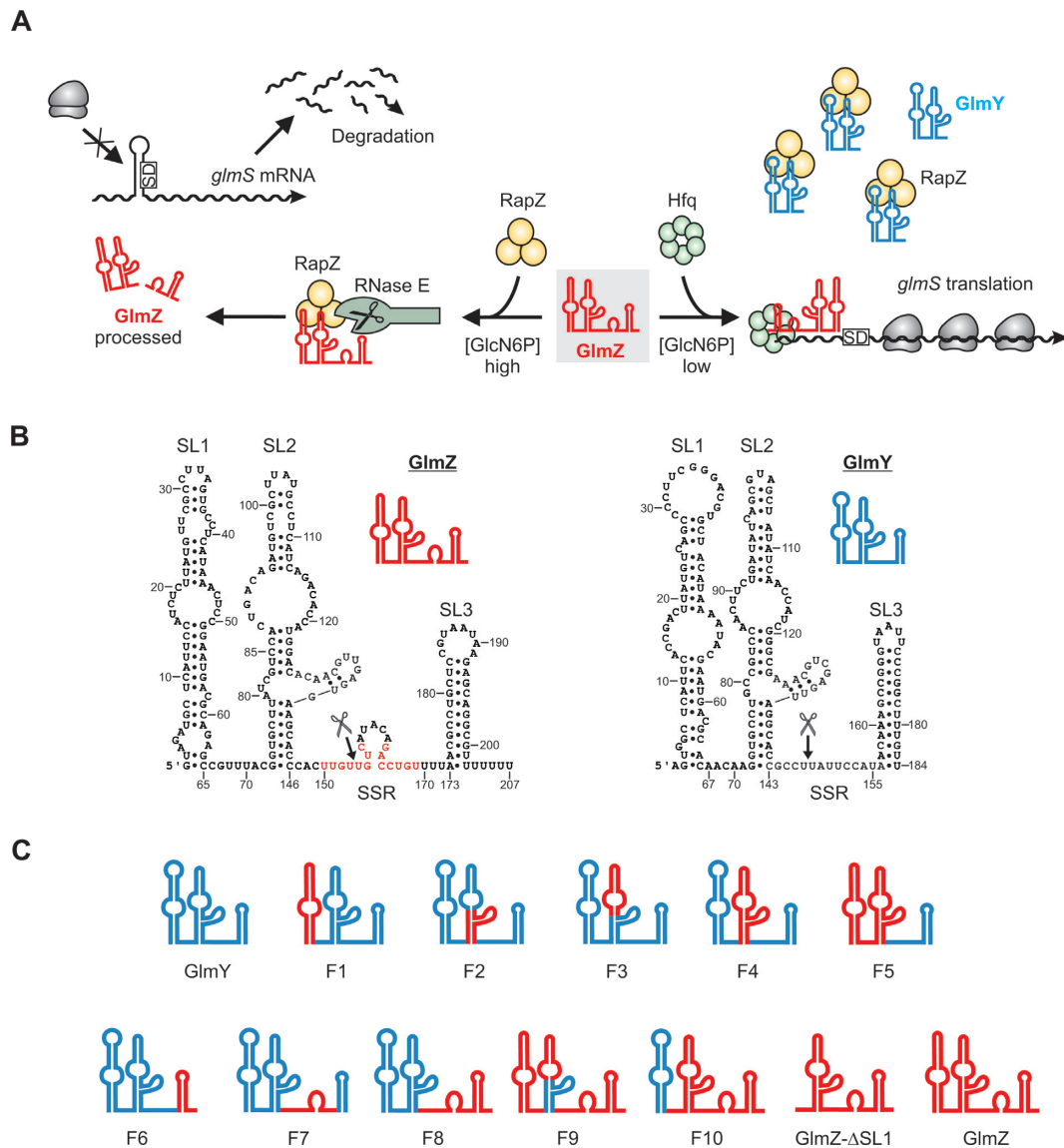
Figure 1. The homologous small RNAs GlmY and GlmZ in *E. coli*. (**A**) Alternative fates of sRNA GlmZ in response to the intracellular GlcN6P level as revealed by combined results obtained in the present and a previous study (30). Under conditions of GlcN6P sufficiency, GlmZ is rapidly bound by adaptor protein RapZ. RapZ recruits RNase E, which cleaves GlmZ in the *glmS* base-pairing region. As a result, GlmZ is unable to activate translation of *glmS* mRNA, which is therefore rapidly degraded (left). Under condition of GlcN6P deficiency, RapZ is sequestered by the homologous sRNA GlmY, thereby inhibiting processing of GlmZ. Accordingly, Hfq can bind the 3′ end of GlmZ, additionally contributing to protection of GlmZ from cleavage by RNase E. Hfq promotes base-pairing of GlmZ with *glmS* leading to GlmS synthesis. (**B**) Secondary structures of GlmZ and GlmY according to Mfold predictions and structure probing (30). The three stem loops (SL1–SL3) and the single stranded regions (SSR) in the sRNAs are denoted. The *glmS* base-pairing site in GlmZ is labelled in red. Processing sites in GlmZ and GlmY are marked with vertical arrows. For GlmZ, only the most frequently used processing site is indicated. (**C**) Schematic drawing depicting the composition and nomenclature of the various GlmYZ chimeras and derivatives used in this study. Sequence elements derived from GlmY are shown in blue and those derived from GlmZ are labelled in red. Exact sequence coordinates can be retrieved from Supplementary Table S2.

Isolation of total RNA and Northern analysis

Northern analysis of total RNA was carried out as described previously (34). Details are described in Supplementary Materials and Methods.

RNase E cleavage assays

RNase E cleavage assays were performed as previously described (30). Briefly, 40 nM of heat denatured α-^{32}P UTP labelled RNA was re-natured for 5 min at 30°C in 1x reaction buffer (25 mM Tris–HCl pH 7.5, 50 mM NaCl, 50 mM KCl, 10 mM MgCl$_2$, 1 mM DTT) containing 0.1 mg/ml yeast tRNA (Ambion). Reactions were either incubated with 150 nM Strep-RapZ or the equivalent volume of 1x reaction buffer for additional 10 min. Cleavage was started by addition of the N-terminal catalytic domain of His-RNase E (RNase E-N) in the assigned concentrations. The reaction was stopped after 20 min by the addition of 4 u of Proteinase K (Thermo Scientific) in 1 volume of Proteinase K buffer (100 mM Tris–HCl pH 7.5, 12.5 mM EDTA, 150 mM

NaCl, 1% SDS) and an additional incubation at 50°C for 10 min. For RNase E cleavage assays addressing the impact of Hfq on GlmZ processing, time course experiments were performed. Here, cleavage of GlmZ was assessed using fixed concentrations of RNase E-N (10 nM) and Strep-RapZ (38 nM) in the absence and presence of 150 nM Hfq. Reactions were prepared in 80 μl volume, 10 μl samples were taken at indicated times, and reactions were stopped by Proteinase K treatment as above. Subsequently 2x RNA loading dye was added and the samples were separated on 7 M urea/TBE/8% PAA gels. Gels were dried and analysed by Phosphoimaging.

3′ RACE

The 3′ RACE analysis was carried out using a previously published protocol (34). Strain Z106 harbouring the respective plasmids encoding GlmZ (pYG84), CMR4 (pYG177), CMR5 (pYG178) and CMR6 (pYG180), as well as the wild-type strain R1279 were grown in LB-medium to mid-log phase ($OD_{600} \approx 0.5$–0.6) and harvested after addition of RNAprotect^{TM} Bacteria Reagent (Qiagen). Total RNA was prepared using the RNeasy Mini Kit (Qiagen) and treated with DNase I (Thermo Scientific). Subsequently, 2.5 μg RNA was ligated with 100 pmol RNA oligonucleotide RIBOLI using 20 units T4 RNA ligase (NEB) in 1x reaction buffer (50 mM Tris–HCl pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 10% DMSO) at 37°C for 2 h. After Ethanol/LiCl precipitation, the RNA pellet was resuspended in 20 μl RNase free water pH 8.0. Five microlitres of the resolved RNA were incubated with 100 pmol DNA oligonucleotide DEOXYLI and 10 mM dNTPs at 65°C for 5 min and shocked on ice. First strand cDNA synthesis was carried out for 1 h at 42°C. The reaction was started by addition of 200 u ProtoScript II Reverse Transcriptase (NEB) in ProtoScript II RT Reaction buffer (NEB) containing 40 u RNase Inhibitor (Thermo Scientific) and 10 mM DTT. Subsequently, 2 μl of the RT reaction were used together with primers BG946 and DEOXYLI to amplify the cDNA products. PCR products were digested with PstI/BamHI and inserted into the PstI/BamHI sites of vector pLDR10. Recombinants were analysed by colony-PCR using primers BG946 and BG1286. PCR products corresponding in size to the processed form of the respective sRNA species were sequenced.

Western blotting

Western blot analyses were performed as previously described (30). Polyclonal rabbit anti-FLAG antiserum (antibodies-online) was used at a dilution of 1:20 000 to detect RpsT-3xFLAG. GFP was detected using GFP (D5.1) XP® rabbit monoclonal antibody (Cell Signaling Technology) diluted 1:10 000. The antibodies were visualized using goat anti-rabbit IgG secondary antibodies conjugated to alkaline phosphatase (Promega) at a dilution of 1:100 000. Signals were detected by chemiluminescence using CDP* as substrate (Novagen).

Liquid culture fluorescence measurements

Details are described in Supplementary Materials and Methods.

RESULTS

Construction of GlmYZ chimeras by domain swapping

GlmY and GlmZ are similar sRNAs (Supplementary Figure S1, Figure 1B). According to *in silico* analysis and previous structure probing experiments, these sRNAs fold into three stem loops SL1, SL2 and SL3 (Figure 1B; (30,33,34)). RapZ binds the sRNAs primarily by interaction with residues located in SL2, more specifically in the lateral bulge protruding from its basal part (30). SL3 is located at the 3′ end of the sRNAs and mediates transcription termination. Accordingly, both sRNAs carry 3′-terminal oligo(U) sequences. Single stranded regions (SSRs) are present between SL2 and SL3. The 26 nt long SSR of GlmZ includes the RNase E cleavage site as well as residues involved in base-pairing with *glmS*. In contrast, the SSR of GlmY comprises only 12 nt and exhibits no complementarity to *glmS*. GlmY is also processed in the SSR, but by a yet unknown enzyme (Figure 1B).

In conclusion, GlmY and GlmZ are composed of four major modules: SL1, SL2, SL3 and SSR. Their similarity offers no explanation for the different Hfq binding and RNase E cleavage properties of the sRNAs. In order to pinpoint the molecular requirements for these activities we swapped modules and subdomains between both sRNAs. Ultimately, we constructed 10 different GlmYZ chimeras (designated F1–F10), the composition of which is illustrated schematically in Figure 1C. We reasoned that the various sRNA modules may likely fold into their correct structures when present in context of the homologous sRNA. Indeed, *in silico* analysis using Mfold (38) indicated that all hybrid sRNAs fold in the schematically depicted manner (Figure 1C). As the only exception, hybrid F10 may adopt alternative structures with lower free energy (ΔG) values. To confirm these predictions experimentally, we assessed binding of the various GlmYZ hybrids by RapZ using EMSA (Supplementary Figure S2). Efficient binding should rely on proper folding of SL2, which is contacted by RapZ. The EMSAs confirmed previous data (30) showing that RapZ specifically binds GlmY (app. K_d ~30 nM) and GlmZ (app. K_d ~75 nM). The results indicated that RapZ binds most hybrid sRNAs similarly efficient (Supplementary Figure S2). The only exception was hybrid F10, which required significantly higher RapZ concentrations for complex formation. In conclusion, the EMSA results support the Mfold predictions indicating that except for hybrid F10 the various chimeras form the expected secondary structures.

The 3′ end comprising the base-pairing region and the intrinsic terminator enables GlmZ to bind Hfq with high affinity.

We previously observed that Hfq interacts with GlmZ with high affinity *in vitro* (app. K_d ~10 nM), whereas its affinity for GlmY is much lower (30), i.e. GlmY requires ~10-fold higher Hfq concentrations for complex formation (Figure
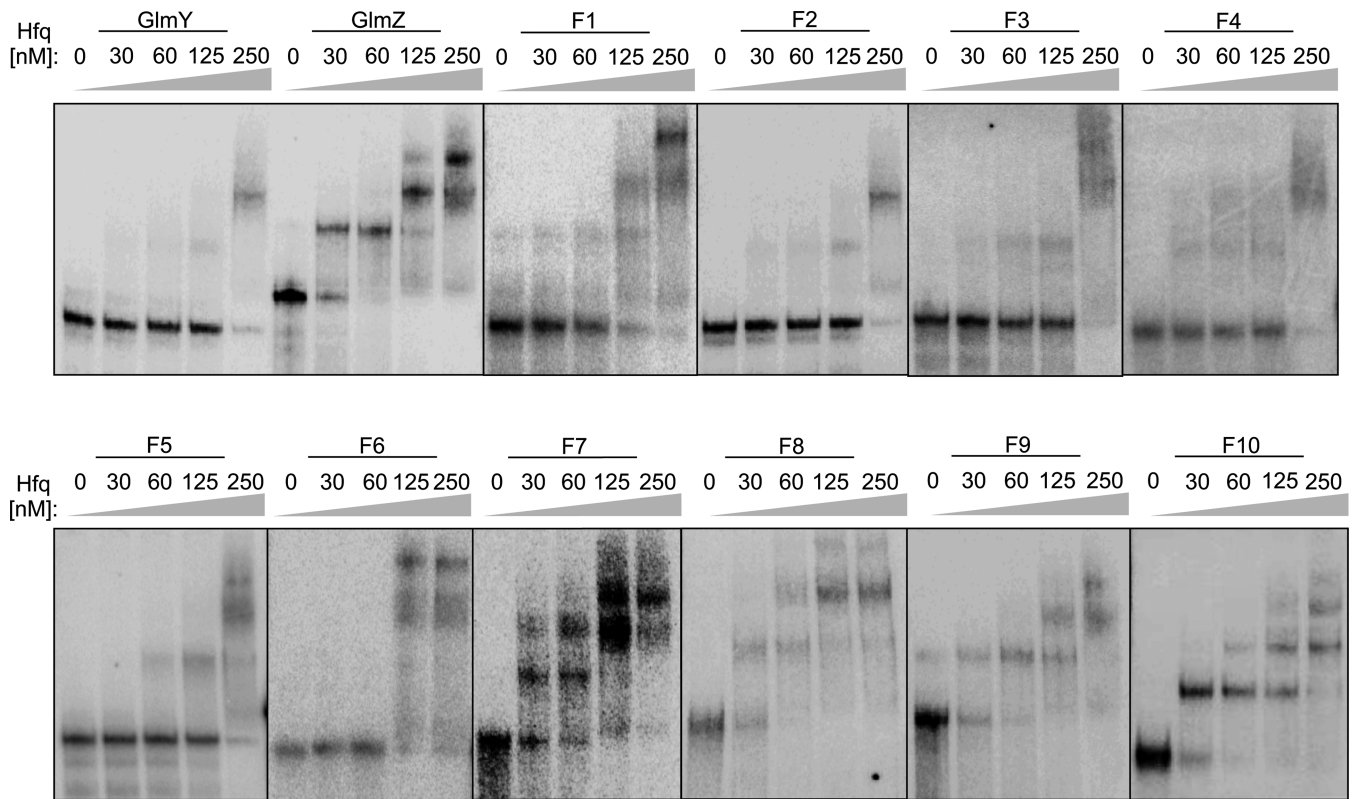
Figure 2. Interaction of GlmY, GlmZ and the various chimeras with Hfq *in vitro*. EMSAs using α-^{32}P-UTP labelled sRNAs and increasing concentrations of purified Hfq protein as indicated. Binding reactions were separated by non-denaturing gel electrophoresis and RNAs were visualized by phosphorimaging. It should be noted that assays using wild-type GlmZ and/or GlmY were carried out in parallel and analysed alongside (data not shown) the respective chimeric sRNA(s) on the same gel to allow for direct comparison of the binding affinities.

2). To dissect the sequence elements responsible for these drastically different Hfq binding properties, we tested interaction of the chimeric sRNAs with Hfq by EMSA. This analysis generated clear-cut results: All chimeras carrying the sequence of GlmZ downstream of SL2, i.e. F8, F9 and F10 interacted with Hfq with high affinity, indistinguishable from GlmZ (Figure 2). In contrast, chimeras possessing the sequence of GlmY downstream of SL2, i.e. F1, F2, F3, F4 and F5 showed a much lower affinity for Hfq, comparable to that of GlmY. These results indicate that the sequence following SL2 enables GlmZ to interact with Hfq with high affinity. To further dissect this module, we swapped the SSR and SL3 of GlmY individually with the corresponding modules from GlmZ. Swapping SL3 (hybrid F6) only mildly improved the Hfq binding potential when compared to GlmY (Figure 2). In contrast, swapping of the SSR resulted in a sRNA (hybrid F7) with a significantly enhanced affinity for Hfq, indicating that the SSR of GlmZ is a major determinant for interaction with Hfq. However, to obtain the full Hfq binding potential, both modules SSR and SL3 from GlmZ are required as exemplified by hybrid F8 (Figure 2; Supplementary Figure S3).

Recent studies demonstrated a role of the 3′ terminal oligo(U) tail for interaction of some sRNAs with Hfq (7,8). Since SL3 of GlmZ contributes to interaction with Hfq, we determined whether the oligo(U) tail at the 3′ end is responsible for this function. To this end, we tested a GlmZ variant lacking the 5 unpaired 3′-terminal uridines by EMSA. In-

deed, absence of these uridines clearly impaired interaction although considerable Hfq binding potential was retained as compared to GlmY (Supplementary Figure S4; Figure 2). In conclusion, there are two distinct regions in GlmZ required for specific interaction with Hfq, at least *in vitro*: A major interaction site is provided by the SSR of GlmZ, and the oligo(U) tail following the terminator hairpin also contributes to interaction.

The 3′ end of GlmZ provides protection against degradation in an Hfq-dependent manner

Next, we aimed to obtain *in vivo* evidence supporting our conclusion that the 3′ end of GlmZ is involved in interaction with Hfq. *In vivo*, GlmZ is destabilized in the absence of Hfq, similar to other Hfq dependent sRNAs. Accordingly, lower amounts of full-length GlmZ are observed in *hfq* mutants. In contrast, the amount of GlmY is unaffected (30). To characterize the modules of GlmZ providing protection against degradation in an Hfq-dependent manner, we compared steady state levels of the various GlmYZ chimeras in isogenic *hfq⁺* versus *Δhfq* strains. To this end, the various sRNAs were expressed from plasmid pBR–pLac, allowing transcription of the sRNAs to start at their genuine +1 position from the IPTG inducible $P_{LlacO-1}$ promoter (39). Since expression of all hybrids is driven from isogenic plasmids by the same synthetic promoter, differences in the sRNA amounts likely reflect differences in stabili-
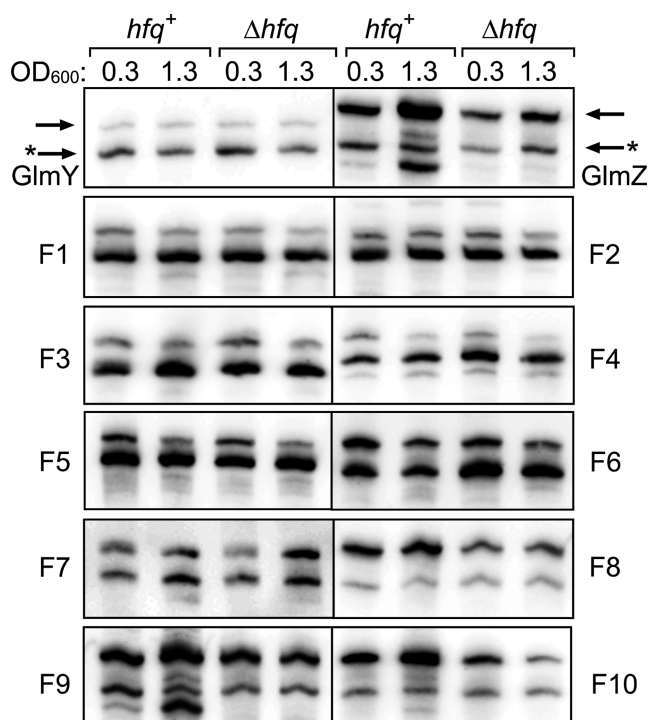
Figure 3. Effect of a *Δhfq* mutation on cellular levels of GlmY, GlmZ and the various GlmYZ chimeras. Northern blot analyses addressing the amounts of the hybrid GlmYZ sRNAs present in strains Z106 (*hfq⁺*) and Z865 (*Δhfq*). Both strains lack chromosomal *glmY* and *glmZ* to allow for detection of the plasmid encoded sRNAs using a mixture of probes specific for GlmY and GlmZ. The sRNAs were transcribed from plasmid pBR–plac under control of the LacI-repressible $P_{LlacO-1}$ promoter (see Supplementary Table S2 for plasmid names). In addition, the compatible plasmid pFDX500 delivering LacI was present. Bacteria were grown in LB and total RNA was extracted from samples harvested at an OD_{600} of ~0.3 and ~1.3, respectively. Different sRNA species are separated by lines and their names are given alongside. Arrows denote the full-length and processed (additionally labelled with asterisks) forms of GlmY and GlmZ. The membranes were re-probed against 5S rRNA to obtain loading controls (Supplementary Figure S7).

ties rather than expression. Pilot experiments revealed that GlmY and GlmZ were strongly overproduced from plasmid pBR–pLac as compared to their chromosomally encoded copies (Supplementary Figure S5). Despite their overproduction plasmid-encoded GlmY and GlmZ are functional in regulation of *glmS* translation (Supplementary Figure S6) as predicted by the model in Figure 1. In order to obtain more physiological sRNA levels, a compatible plasmid producing LacI for repression of the $P_{LlacO-1}$ promoter was additionally introduced. The resulting transformants exhibited reduced expression of the sRNAs. As demonstrated for GlmY and GlmZ, the levels of plasmid and chromosomally encoded sRNAs were comparable, at least during exponential growth (Supplementary Figure S5).

Northern analysis showed that the amount of full-length GlmZ was significantly reduced in the *hfq* mutant, whereas GlmY levels remained constant (Figure 3). Comparable results were previously obtained when addressing the chromosomally encoded sRNAs (30). From the various GlmYZ chimeras only F8, F9 and F10 were affected by the *hfq* mutation. The amounts of the unprocessed forms of these sR-

NAs were significantly reduced in the *Δhfq* mutant indicating that Hfq stabilizes these hybrids. In contrast, all other chimeras behaved as GlmY and were not affected by the *hfq* mutation. As a common feature, hybrids F8, F9 and F10 carry the SSR as well as SL3 of GlmZ. Taken together, the *in vivo* data (Figure 3) are in perfect agreement with the EM-SAs (Figure 2) showing that the complete 3′ end of GlmZ is required for high affinity interaction with Hfq. In contrast, the origin of SL1 and SL2 plays no role for this interaction. An additional variant, which is shorter than the normally processed species, was visible in the *hfq⁺* strain for GlmZ and also for hybrid F9. This species has not been observed before for chromosomally encoded GlmZ (30,32–34). Control experiments show that appearance of this novel GlmZ variant strictly correlates with GlmZ overproduction (Supplementary Figure S5B), indicating that it might result from aberrant cleavage due to an overload of the canonical processing machinery (Supplementary Figure S8).

Hfq counteracts cleavage of GlmZ by RNase E

Our analyses indicate that Hfq specifically binds the 3′ end of GlmZ consisting of SL3 and the SSR, which is also cleaved by RapZ/RNase E (Figure 1B), raising the possibility that interaction with Hfq protects GlmZ from cleavage by RNase E. To obtain evidence for this hypothesis, we tested the impact of Hfq on cleavage of GlmZ by RNase E *in vitro*. Our previous data showed that presence of RapZ and the catalytic domain of RNase E (subsequently designated RNase E-N) are essential and sufficient to reconstitute processing of GlmZ *in vitro* (30). Here, we used protein concentrations triggering partial cleavage of GlmZ within 20 min reaction time (30). The fate of GlmZ was followed over a period of 60 min in absence and presence of 150 nM Hfq (Figure 4). In absence of Hfq, almost complete processing of GlmZ could be observed. No cleavage occurred in the absence of RapZ or RNase E-N or both (Figure 4, lanes 1–3), confirming that both proteins are required for this function. Presence of Hfq considerably retarded cleavage of GlmZ supporting the idea that binding of GlmZ by Hfq counteracts processing (Figure 4).

To confirm this conclusion, additional assays with varying RapZ concentrations and a constant reaction time of 20 min were performed (Supplementary Figure S9). In agreement with previous data (30), an ~3-fold molar excess of RapZ over RNase E-N was required to achieve complete cleavage of GlmZ. However, in additional presence of 150 nM Hfq, complete cleavage of GlmZ was observed only at the highest RapZ concentration corresponding to a 15-fold molar excess over RNase E-N. In sum, the data indicate that Hfq counteracts cleavage of GlmZ by RapZ/RNase E, most likely by masking the RNase E cleavage site upon binding to the SSR.

GlmZ is a cofactor-dependent 'direct entry' substrate for RNase E

RNase E preferentially cleaves RNAs with 5′ monophosphorylated ends. However, initial cleavage of a distinct group of substrates occurs via a 'direct entry' mechanism that is independent of the 5′ phosphorylation status of the
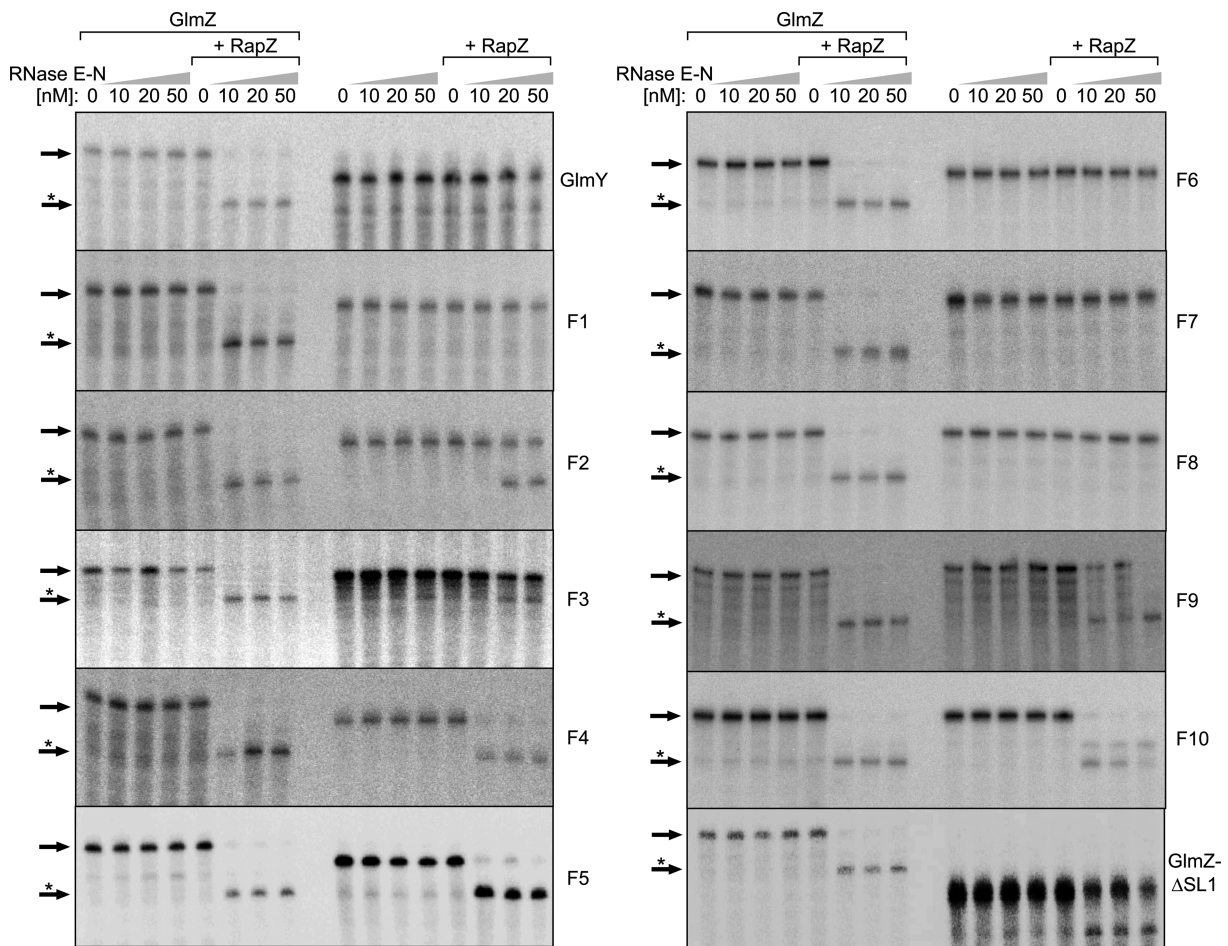
Figure 4. Hfq antagonizes cleavage of GlmZ by RapZ/RNase E *in vitro*. Temporal course of GlmZ processing by RNase E and RapZ *in vitro* in absence or presence of 150 nM Hfq. Reaction mixtures containing 40 nM α-^{32}P-UTP-labelled GlmZ and the proteins designated with plus signs were incubated at 30°C. Samples were removed and reactions were stopped at the indicated times. In the first four lanes GlmZ was incubated for 60 min either alone (lane 1) or together with only one of the three proteins (lanes 2–4) showing that none of these proteins is capable to trigger GlmZ decay when provided alone. Reaction mixtures were separated on denaturing polyacrylamide gels, which were analysed by phosphorimaging. Full-length and processed (additionally labelled with an asterisk) forms of GlmZ are denoted with arrows.

RNA. Cleavage of GlmZ by RNase E strictly depends on the presence of RapZ, but the mechanistic role of this protein is essentially unknown. Removal of the 5′-terminal triphosphate from RNA substrates by RppH requires at least two unpaired nucleotides at the RNAs 5′ end (40). However, the first two nucleotides at the 5′ end of GlmZ are presumably involved in intramolecular base-pairing (Figure 1B) suggesting that GlmZ is resistant to 5′ end dephosphorylation catalysed by RppH. In agreement, GlmZ levels are unaffected by a $\Delta rppH$ mutation (Supplementary Figure S10). Hence, one possibility is that RapZ serves as pyrophosphohydrolase specific for GlmZ. To address this question, we tested cleavage of GlmZ, carrying either a 5′ triphosphate or a 5′ monophosphate, by RNase E-N *in vitro*. However, even when carrying a monophosphate at its 5′ end, GlmZ was resistant to cleavage by RNase E-N (Supplementary Figure S11). Cleavage only occurred in presence of RapZ as observed for 5′-triphosphorylated GlmZ. These observations suggest that the 5′ phosphorylation state of GlmZ plays no role for its cleavage. Accordingly, GlmZ is processed by RNase E via a 'direct entry' mechanism that depends on the additional presence of RapZ.

The central stem loop of GlmZ is crucial for RapZ-mediated cleavage by RNase E

Assisted by RapZ, RNase E cleaves GlmZ in the SSR independent of its 5′ phosphorylation status. However, binding of an RNA substrate by RapZ does not necessarily lead to cleavage since GlmY, although perfectly bound by RapZ (Supplementary Figure S2), is not cleaved by RNase E (Figure 5, Supplementary Figure S12; (30)). Consequently, elements triggering cleavage by RNase E must be present in GlmZ but absent in GlmY.

To map these elements, *in vitro* RNase E cleavage assays using the chimeras were performed. The various sRNA species were incubated with increasing concentrations of RNase E-N in the absence or presence of RapZ (Figure 5). For direct comparison of cleavage efficiencies, cleavage of wild-type GlmZ was assessed in parallel. None of the sRNAs was cleaved by RNase E in the absence of RapZ, reemphasizing the essential role of this protein for cleavage. In the presence of RapZ, complete cleavage by RNase E could be observed only for hybrid sRNAs possessing SL2 of

GlmZ (i.e. F4, F5, F10; Figure 5). In contrast, all chimeras carrying SL2 of GlmY were completely resistant to cleavage (i.e. F1, F6, F7, F8). Further, the origin of SL1 appears to be irrelevant for cleavage. Hybrids F4 and F10 were cleaved by RNase E even though they carry SL1 of GlmY (Figure 5). To determine whether SL1 is at all required for cleavage, we tested a 5′-truncated GlmZ variant lacking SL1. The resulting molecule GlmZ–ΔSL1 (Figure 1C) was still cleaved by RapZ/RNase E, but inefficiently (Figure 5). GlmZ-ΔSL1 is only very weakly bound by RapZ, which might explain this result (Supplementary Figure S2). It remains to be clarified whether SL1 directly contributes to interaction with RapZ or indirectly by triggering correct folding of SL2, which is contacted by RapZ (30).

Next, we tried to dissect requirements within SL2 for RNase E cleavage in more detail. We previously showed that mutation of residues G132, G137 or G139, located in the lateral bulge of SL2, strongly inhibited cleavage suggesting that this element is important for processing of GlmZ by RNase E (30). The single nucleotide deviating between GlmY and GlmZ in the apical part of this bulge is U134 in GlmZ, corresponding to residue C131 in GlmY (Figure 1B and Supplementary Figure S1). To determine whether this position is decisive for RNase E cleavage, we swapped the corresponding residues between GlmY and GlmZ. However, we observed no changes in performance when comparing the mutants with the respective wt-sRNAs (Supplementary Figures S2 and S12). This result suggested that residue U134 is either not important or that multiple elements uniquely present in SL2 of GlmZ are required for RNase E cleavage. The latter possibility is supported by additional experiments in which only parts of SL2 were swapped between GlmY and GlmZ. Hybrids F2 and F3 correspond to GlmY molecules carrying either the lower or the upper portion of SL2 of GlmZ (Figure 1C). Hybrid F3 was resistant to cleavage and for hybrid F2 only partial cleavage could be observed (Figure 5). In a complementary experiment the lower portion of SL2 in GlmZ was swapped for the corresponding element of GlmY generating hybrid F9. Complete cleavage of this hybrid was only observable at highest RNase E concentrations indicating inefficient processing (Figure 5). Collectively, these results show that SL2 of GlmZ is the only element that is specifically required

Figure 5. The central stem loop of GlmZ is the decisive element required for RNA cleavage by RNase E and RapZ. *In vitro* cleavage assays of α-^{32}P-UTP-labelled sRNAs using varying concentrations of RNase E-N and a fixed concentration of 150 nM RapZ. sRNAs were generated by *in vitro* transcription of PCR fragments obtained using the corresponding plasmids (see Supplementary Table S2) as templates. The assays were incubated for 20 min at 30°C, stopped and subsequently separated by denaturing gel electrophoresis and analysed by phosphorimaging. The various sRNA species to be tested were separated on the right half of the gels and their names are given alongside, respectively. Assays using wild-type GlmZ were carried out in parallel and separated alongside (left half of each gel) to allow for direct comparison of cleavage efficiencies. Full-length and processed forms of GlmZ are indicated by arrows.

from GlmZ to achieve efficient cleavage of the RNAs by RNase E in a RapZ dependent manner. Surprisingly, the RNase E processing site itself, which is located in the SSR of GlmZ, is dispensable for cleavage since hybrids composed of SL2 from GlmZ and the SSR from GlmY are perfectly cleaved (i.e. hybrids F4 and F5).

RNase E cleaves any sequence when fused to the 3′ end of the RapZ/RNase E recognition module of GlmZ

Our results indicated that SL2 from GlmZ serves as a recognition module for RapZ/RNase E cleavage, whereas presence of the SSR containing the actual cleavage site is dispensable. This observation suggested the possibility that every single-stranded RNA-stretch regardless of its sequence composition might be cleaved by RNase E/RapZ when fused to the 3′ end of SL2 of GlmZ. To test this idea, we created six artificial sRNAs, CMR1–CMR6 (CMR = cleavage module for RNase E), which carried randomly created sequences with varying GC content fused to the 3′ end of the GlmZ–SL1/SL2 module (Figure 6A). In all CMRs,

SL3 was replaced by the λt$_0$ transcriptional terminator. First, we performed RNase E cleavage assays using the various CMRs. Intriguingly, all CMRs were efficiently cleaved by RNase E/RapZ as evaluated from direct comparison with GlmZ (Figure 6B). For some CMRs, in particular for CMR6, a small fraction remained unprocessed, which might be attributed to hairpin formation in the SSR. Next, we tested whether the various CMRs are also cleaved in a RapZ-dependent manner *in vivo* (Figure 7A). We performed Northern analysis of isogenic *rapZ*$^+$ and *ΔrapZ* strains carrying the various *cmr* sequences on plasmids. Indeed, all CMRs were efficiently cleaved in a RapZ-dependent manner, i.e. full-length as well as processed CMR species were detectable in the *rapZ*$^+$ strain, whereas in the *ΔrapZ* mutant processing was abolished. Interestingly, the ratio of processed versus unprocessed forms appeared to be higher for the various CMRs when compared to GlmZ. Thus, the CMRs are processed even more efficiently *in vivo*, which might be attributed to an absence of interaction with Hfq,
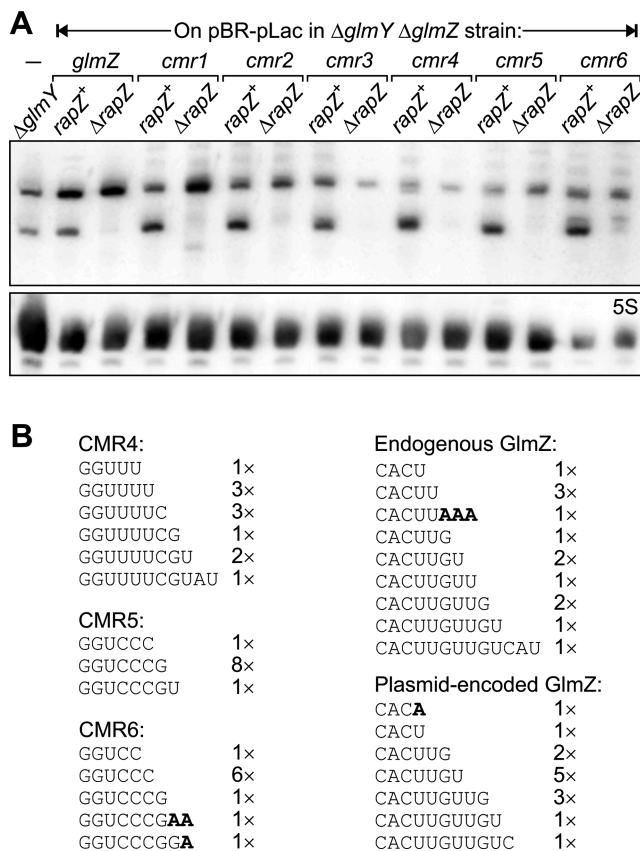
Figure 6. Cleavage of artificial RNA substrates by RNase E and RapZ *in vitro*. (**A**) Schematic presentation of the artificial RNA substrates CMR1–CMR6 tested for cleavage by RNase E and RapZ. All CMRs were composed of the GlmZ–SL1/SL2 module, a variable SSR and the λt0 transcriptional terminator. Of the variable SSRs only the first 6 (CMR1 and CMR2) or first 3 (CMR3–CMR6) residues were preassigned to generate MfeI or AgeI restriction sites on the DNA level. Similarly, the last three residues were specified to obtain an XbaI restriction site together with the first three residues of the λt0 terminator. The sequence between these sites (depicted in bold letters) was randomly created. The GC content (given at the right) was gradually varied from 12 to 100%. Cleavage events mapped by 3′ RACE are indicated by vertical arrows. Bold arrows indicate preferred cleavage site. For comparison, the sequence of the SSR of GlmZ is shown. The *glmS* base-pairing site is underlined. (**B**) *In vitro* assays addressing cleavage of the artificial RNA substrates CMR1–CMR6 by RNase E and RapZ. See legend to Figure 5 for additional information.

which in case of GlmZ protects the SSR from cleavage (Figure 4).

To determine whether the CMRs are cleaved by RNase E/RapZ within their SSRs, in a manner analogous to GlmZ, we mapped the 3′ ends of the processed variants of CMR4, CMR5 and CMR6 by 3′ RACE (Figure 7B). Indeed, all cleavages took place within the arbitrarily composed SSRs. Although processing events were somewhat scattered between positions 151–158, the majority of cleavages occurred between positions 152/153 and 153/154, i.e. 6 and 7 residues, respectively, downstream of SL2 (Figures 6A and 7B). A length of 155 nt was previously reported for the processed species of GlmZ (41) suggesting that it is cleaved two residues further downstream as compared to the CMRs. To clarify this point, we also mapped the 3′ ends of the processed form of chromosomal and plasmid encoded GlmZ (Figures 6A and 7B). Taken together, RNase E cleaves GlmZ preferentially between positions 153/154 (7 clones) and 155/156 (5 clones). These results

are in agreement with previous Northern analysis showing that two processed GlmZ species slightly differing in length were detectable when separated on high resolution acrylamide gels (34). Collectively, our data demonstrate that RNase E/RapZ cleave any foreign sequence regardless of its GC content, when fused to the 3′ end of the GlmZ–SL1/SL2 module. This cleavage preferentially occurs between residues 6/7 or 7/8 within the foreign sequence.

Modulating mRNA and protein abundance using the conditional module for RNase E cleavage

Upon cleavage, the foreign RNA that is fused to the GlmZ–SL1/SL2 module is released in a 5′ monophosphorylated state. Such RNAs are expected to become rapidly degraded due to their interaction with the 5′ phosphate binding pocket of RNase E. Therefore, we reasoned that this tool may be used to modulate mRNA and accordingly protein abundance in a RapZ-dependent manner. That is, an

A

B

CMR4:		Endogenous GlmZ:	
GGUUU	1×	CACU	1×
GGUUUU	3×	CACUU	3×
GGUUUUC	3×	CACUU**AAA**	1×
GGUUUUCG	1×	CACUUG	1×
GGUUUUCGU	2×	CACUUGU	2×
GGUUUUCGUAU	1×	CACUUGUU	1×
		CACUUGUUG	2×
CMR5:		CACUUGUUGU	1×
GGUCCC	1×	CACUUGUUGUCAU	1×
GGUCCCG	8×		
GGUCCCGU	1×	Plasmid-encoded GlmZ:	
		CAC**A**	1×
CMR6:		CACU	1×
GGUCC	1×	CACUUG	2×
GGUCCC	6×	CACUUGU	5×
GGUCCCG	1×	CACUUGUUG	3×
GGUCCCG**AA**	1×	CACUUGUUGU	1×
GGUCCCGG**A**	1×	CACUUGUUGUC	1×

Figure 7. Foreign RNAs fused to the SL1/SL2 module of GlmZ are cleaved at defined positions *in vivo* in a RapZ-dependent manner regardless of their sequence. (**A**) Northern analysis addressing RapZ-dependent cleavage of artificial RNAs CMR1–CMR6 *in vivo*. GlmZ and the various CMR RNAs were expressed from plasmid pBR-plac (see Supplementary Table S2 for plasmid names) in strains Z106 (*rapZ*⁺) and Z864 (*ΔrapZ*), respectively. Both strains lack endogenous *glmY* and *glmZ* to allow for specific detection of the plasmid encoded sRNAs. Additionally plasmid pFDX500 producing LacI was present. Total RNA was isolated from exponentially grown cells and analysed by Northern blotting using a mixture of probes directed against GlmZ and the various CMRs, respectively. Strain Z96 was analysed in parallel (lane 1) to compare the length of the various RNAs with chromosomally encoded GlmZ. The membranes were re-probed against 5S rRNA (bottom). (**B**) Results of the 3′ RACE analysis mapping the 3′ ends of processed CMR4–6 and of chromosomally and plasmid-encoded GlmZ. Amplified cDNA products were cloned and at least 10 inserts corresponding in size to the processed variants were analysed by sequencing. The frequency of occurrence of the individual species is indicated. Adenosine residues in bold are not represented in the sRNA gene and were post-transcriptionally added.

mRNA fused to GlmZ–SL1/SL2 should be relatively stable and susceptible to translation in absence of RapZ. Upon RapZ availability, RNase E would release the mRNA from the fusion in a 5′ monophosphorylated state, leading to its rapid decay and consequently to low levels of the mRNA product (Figure 8A). To test this idea, we first used the well-studied *rpsT* mRNA, which encodes ribosomal protein S20 (18). We created a fusion gene comprising the GlmZ–SL1/SL2 module fused the 5′ end of the sequence corresponding to the *rpsT* mRNA started from its P2 promoter (Figure 8A). As a difference, λt₀ rather than the endogenous *rpsT* terminator is used to stop transcription. Moreover, we added the sequence encoding the 3xFLAG epitope in frame

to the 3′ end of the *rpsT orf* to facilitate detection of the encoded protein.

The *glmZ'-rpsT* fusion gene was placed on a plasmid under control of the $P_{LlacO-1}$ promoter and subsequently introduced into a *rapZ*⁺ strain, a *ΔrapZ* mutant and a *rapZ*⁺ strain additionally overexpressing *rapZ*. First, we determined the levels of the *glmZ'-rpsT* fusion mRNA present in the various strains by Northern analysis (Figure 8B). For comparison, the same strains expressing GlmZ rather than the *glmZ'-rpsT* fusion were tested in parallel. With a GlmZ-specific probe, two transcripts were detected in the *rapZ*⁺ strain encoding the *glmZ'-rpsT* mRNA: A signal corresponding to the size of processed GlmZ representing the 5′ cleavage product and an additional faint band corresponding to unprocessed *glmZ'-rpsT* mRNA (Figure 8B, compare lane 2 with lanes 1 and 5). Intensity of the latter signal increased in the *ΔrapZ* mutant while the 5′ cleavage product disappeared as observed for sRNA GlmZ (Figure 8B, lanes 2, 3, 5, 6). Upon *rapZ* overexpression, full-length *glmZ'-rpsT* was almost undetectable, whereas the 5′ cleavage product strongly accumulated (Figure 8B, lane 4). Comparable results were obtained with an *rpsT*-specific probe. This probe detected the two chromosomally encoded *rpsT* transcripts, which originate from two alternative promoters (P1 or P2) and therefore differ in length (447 nt versus 357 nt; Figure 8, lanes 8–11). An additional band corresponding to the 555 nt long *glmZ'-rpsT* fusion mRNA was detected in the *rapZ*⁺ strain and this signal increased in the *ΔrapZ* mutant, but was absent in the RapZ overproducer (Figure 8, lanes 9–11).

To determine whether the detected differences in *glmZ'-rpsT* mRNA levels also impact on the amounts of RpsT-3xFLAG protein, Western analysis was conducted (Figure 8C). Indeed, highest RpsT-3xFLAG protein amounts were present in the *ΔrapZ* mutant. In contrast, only low RpsT-3xFLAG amounts were observed in the *rapZ* overexpressing strain, whereas intermediary levels were detected in the *rapZ*⁺ strain (Figure 8C, compare lanes 2–4 and 5–7). In sum, these results suggest that the GlmZ-SL1/SL2 module allows modulating abundance of a fused mRNA and its encoded protein product at the post-transcriptional level by RapZ availability. To support this conclusion, we performed an additional experiment using a fusion gene comprising the GlmZ-SL1/SL2 module fused to the 5′ end of a *gfp*mut3* allele rather than the *rpsT* gene. The corresponding Gfpmut3* variant is provided with a SsrA protein degradation-tag resulting in a half-live of ∼40 min (42), which is likely shorter than the notoriously long half-live of ribosomal proteins such as RpsT. The construct encoding the *glmZ'-gfp*mut3* fusion gene was introduced into the *rapZ*⁺ strain, the *ΔrapZ* mutant and the strain overexpressing *rapZ*. Subsequently, we measured the fluorescence produced by these transformants and assessed in parallel their Gfpmut3* protein levels by Western blotting (Figure 8D). Indeed, fluorescence was detectable in the *ΔrapZ* mutant, whereas fluorescence yields produced by the *rapZ*⁺ strains were close to background levels (Figure 8D, left). In agreement, Western blotting detected the Gfpmut3* protein exclusively in the *ΔrapZ* mutant, whereas it was undetectable
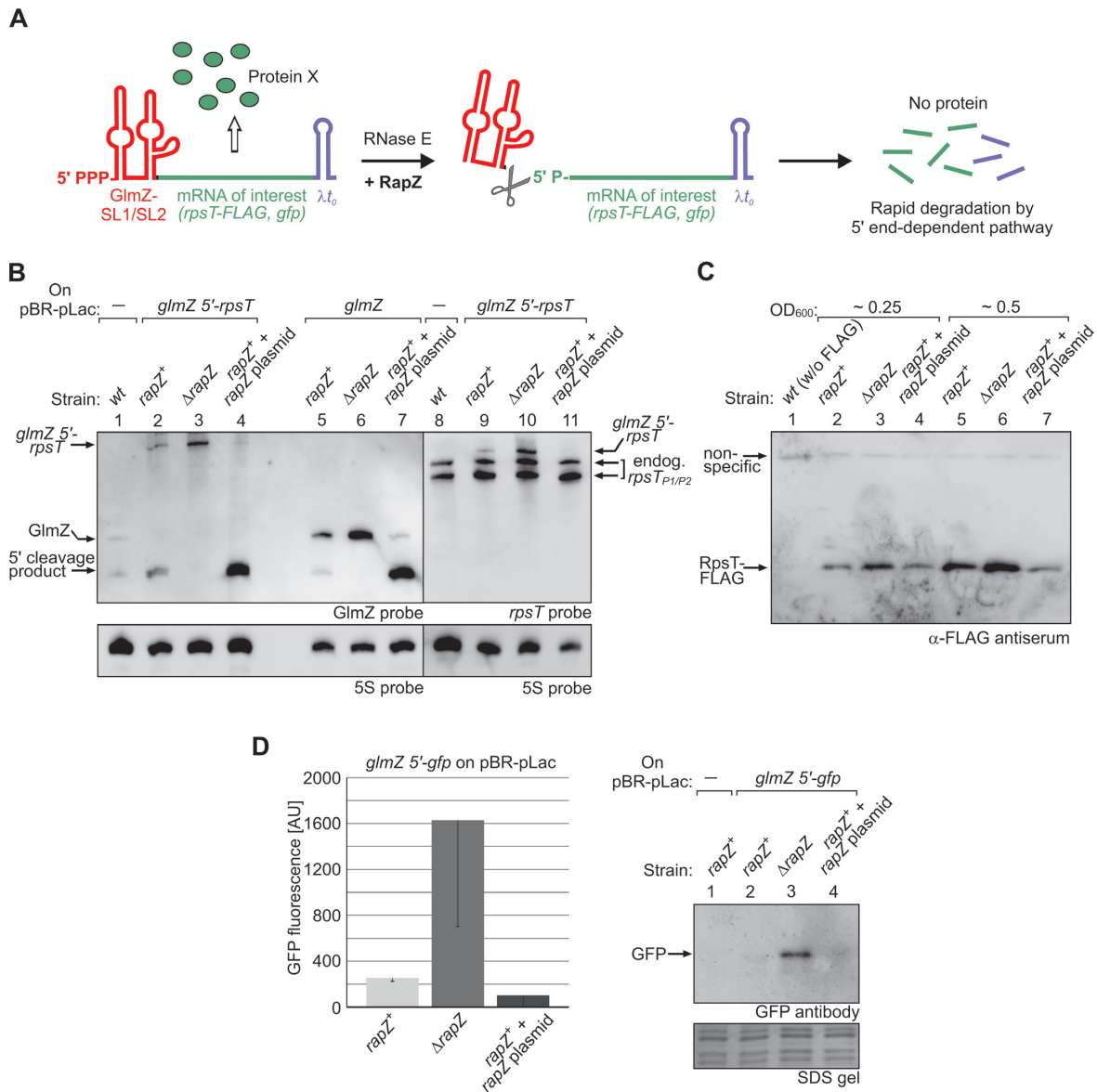
Figure 8. The SL1/SL2 module of GlmZ allows posttranscriptional regulation of mRNA and protein abundance through modulation of RapZ availability. (**A**) Outline of the experimental approach to address RapZ-dependent cleavage of artificial fusion mRNAs by RNase E. One tested fusion RNA consists of the *rpsT*-P$_2$ transcript (*rpsT* −43 to +261; in green) fused to the 3′ end of the GlmZ–SL1/SL2 module (in red). The sequence encoding the 3xFLAG epitope was fused in frame to the 3′ end of the *rpsT orf* to enable detection of the encoded RpsT protein. The λt$_0$ terminator at the 3′ end (in blue) serves to stop transcription. The second tested fusion RNA carries the sequence encoding an unstable Gfpmut3* variant including a 31 nt long 5′ UTR fused to the GlmZ–SL1/SL2 module. The fusion RNAs are expected to be stable and expressed in the absence of RapZ protein (left). RapZ availability leads to recruitment of RNase E, which cleaves the transcripts a few residues after the SL1/SL2 module leading to release of the *rpsT* and *gfp*mut3* mRNAs, respectively, bearing monophosphate groups at their 5′ ends (middle). Accordingly, the latter transcripts are expected to become rapidly degraded via the 5′ end dependent pathway suppressing synthesis of the encoded proteins (right). (**B**) Northern blot analysis of cells transcribing the *glmZ′-rpsT* fusion mRNA from the $P_{LlacO-1}$ promoter on plasmid pYG183 (lanes 2–4 and 9–11). For comparison, cells expressing sRNA GlmZ from plasmid pYG84 were tested in parallel (lanes 5–7). The plasmids were introduced into strains Z106 (*rapZ*$^+$; lanes 2, 5, 9), Z864 (*ΔrapZ*; lanes 3, 6, 10) and strain Z106 overexpressing *rapZ* from plasmid pFDX4324 (lanes 4, 7, 11), respectively. These strains lacked chromosomally encoded *glmY* and *glmZ*. In addition, the compatible plasmid pFDX500 delivering LacI for repression of the $P_{LlacO-1}$ promoter was present. Total RNA extracted from exponentially growing cells was hybridized with probes specific for GlmZ (lanes 1–7) or *rpsT* (lanes 8–11). RNA extracted from wild-type strain R1279 was included for detection of chromosomally encoded endogenous GlmZ (lane 1) and *rpsT* transcripts (lane 8). (**C**) Western analysis addressing the impact of RapZ on synthesis of the RpsT-FLAG protein encoded by the *glmZ′-rpsT* fusion mRNA. The strains and transformants used in Figure 8B, lanes 1–4 were tested. Protein extracts were prepared from cells harvested at cell densities corresponding to OD$_{600}$ readings of ∼0.25 (lanes 2–4) and ∼0.5 (lanes 1 and 5–7) and subsequently separated on 15% SDS polyacrylamide gels and blotted. Application of the α-FLAG antiserum detected the RpsT-FLAG protein, which was absent in the untransformed wild-type strain R1279 (lane 1; negative control). A non-specifically detected protein band served as loading control (marked with an arrow). (**D**) Fluorescence measurements (left) and Western analysis (right) of strains transcribing the *glmZ′-gfp*mut3* fusion mRNA. The same strains as described for Figure 8B, lanes 2–4 were used, but as a difference the strains carried plasmid pYG213 encoding the *glmZ′-gfp*mut3* mRNA rather than plasmid pYG183. Fluorescence yields above background levels are expressed as arbitrary units [AU]. In the Western blotting analysis, strain Z106 carrying the empty plasmid pBR–pLac was included as a negative control (lane 1). As a loading control, the SDS PAA gel stained after blotting is provided at the bottom.

in the *rapZ*⁺ strain and in the RapZ overproducing strain (Figure 8D, right).

DISCUSSION

In the current work, we dissected functional modules in a bacterial sRNA by domain swapping. Advantageously, this approach may ensure correct folding of modules within the context of the homologous sRNA, as opposed to mutational and deletion analyses that generate unpredictable constraints on sRNA structures. We identified two distinct regions in GlmZ, namely the 3′ terminal oligo(U) tail and the SSR preceding the transcriptional terminator, which bestow the ability to interact with Hfq. Consequently, Hfq binding inhibits GlmZ processing, most likely by preventing access of RNase E to the overlapping cleavage site. Intriguingly, the sequence comprising this site is not mandatory for RapZ-mediated cleavage by RNase E. In contrast, the central stem loop of GlmZ triggers cleavage within any RNA that is attached to its 3′ end suggesting a novel mechanism for RNase E 'direct entry' mediated by an RNA-binding protein. Finally, we showed that GlmZ provides an aptamer for generation of 5′ monophosphorylated RNAs by request, through modulation of RapZ availability. This tool might not only be valuable for RNA research in general, but also permits to control gene expression at the post-transcriptional level through modulation of RNA decay.

GlmZ binds Hfq with an affinity of ∼10 nM. GlmY requires much higher Hfq concentrations indicating that it is not a *bona fide* Hfq-binding sRNA (Figure 2; (30)). Inconsistent with our data, another study reported comparable Hfq binding affinities for both sRNAs in the range of 4–5 nM (4). Yet, our current data unequivocally show that high affinity for Hfq is an exclusive feature of GlmZ and determined by its 3′ end. By swapping this element, the Hfq binding potential of GlmZ is transferred to GlmY (Figures 2 and 3; Supplementary Figure S3). Our findings are supported by global Hfq pulldown studies, which identified GlmZ and the *glmS* mRNA as being associated with Hfq, whereas GlmY could not be retrieved (43–45). Taken together, it appears unlikely that GlmY relies on Hfq for function as proposed by Salim *et al.* (4). GlmY requires Hfq, neither for turnover (Figure 3) nor for recruitment of RapZ (Supplementary Figure S2), as it perfectly counteracts cleavage of GlmZ *in vitro* in absence of Hfq (30). GlmZ uses two distinct elements, the 3′ terminal oligo(U) tail and the SSR, to contact Hfq. The latter element is also consistent with a recent Hfq–RNA crosslinking study proposing that the consensus Hfq binding site may include U–U dinucleotide(s) close to or overlapping the sRNAs seed region (45). The SSR of GlmZ contains three U–U dinucleotides, which overlap the base-pairing site (Figure 1B). GlmZ competes with other sRNAs for binding to Hfq and is accordingly destabilized upon overproduction of Hfq-associated sRNAs (11,29). This suggests that GlmZ contacts Hfq in a manner similar to other sRNAs (7–9,12): the 3′ terminal oligo(U) tail may bind the proximal face, while the SSR might contact the rim, although this remains to be proven. In contrast, GlmY is an inefficient competitor (11), consistent with its weak Hfq binding potential observed here (Figures 2 and 3). GlmY lacks an accessible 3′ oligo(U) tail and

its SSR might be too short or structurally constrained to undergo high affinity interactions with Hfq.

Binding of Hfq protects GlmZ from processing by RNase E, presumably by blocking access to the cleavage site (Figure 4). Thus, the fate of GlmZ is ultimately determined by competition between Hfq and RapZ (Figure 1A). When the concentration of free RapZ is high, it will rapidly bind GlmZ and target it to cleavage before Hfq may associate. When RapZ is sequestered by GlmY, Hfq may bind and protect GlmZ, thereby steering it to base-pairing. Inhibition of cleavage by Hfq due to overlapping RNase E cleavage and Hfq binding sites has also been observed for sRNAs DsrA and RprA (14,28). Interestingly, degradation of RprA by RNase E is regulated by osmolarity through a yet unknown mechanism (27). It is tempting to speculate that mechanisms reminiscent of GlmZ also control the fates of these sRNAs: when superfluous, they are degraded by RNase E, which might be recruited by an adapter protein or by a base-pairing RNA sponge as reported for sRNA GcvB (29). When required, Hfq binding may protect and license the sRNAs to base-pairing.

Attempts have been made to define a consensus sequence for RNase E target sites. These studies concluded that RNase E cleaves single-stranded AU-rich sequences, preferably those carrying a G residue 2 nt upstream of the scissile bond (18,22). This assumption appears to be foiled by our observation that RNase E (assisted by RapZ) even cleaves 100% GC-rich sequences when fused to SL2 of GlmZ. Moreover, cleavage did not require a G residue at the -2 position (Figures 6 and 7). In contrast, RNase E apparently cleaved at a fixed distance 6 or 7 nt downstream of SL2. Using this device RNase E can be redirected to cleave within any RNA sequence, but how does it work? GlmZ does not fulfil the requirements for either of the two known paths of RNase E cleavage. The phosphorylation state of its 5′ end plays no role for cleavage (Supplementary Figures S10 and S11), and there is only one SSR to which RNase E may bind. The need for a second handhold in GlmZ may be bypassed by RapZ, which binds GlmZ at SL2 and also interacts with the RNase E N-terminus. Consequently, GlmZ is completely resistant to cleavage in absence of RapZ (Supplementary Figure S11; (30)). However, RNA binding by RapZ *per se* is not sufficient for its cleavage, since substrates containing SL2 of GlmY are not cut (Supplementary Figure S2 and Figure 5). The lateral bulge in SL2 is not a discriminating element since its apical part is conserved in both sRNAs (Supplementary Figure S1) and the only deviating residue (U134 in GlmZ) has no impact (Supplementary Figure S12). We speculate that only substrates containing SL2 of GlmZ can be sterically orientated by RapZ allowing RNase E to access the adjacent SSR, which is ultimately cleaved. Alternatively, RapZ might melt SL2 of GlmZ providing an additional handhold for RNase E. Melting may be impaired in case of GlmY due to stronger base-pairings (Figure 1B). This mechanism could provide an explanation for our unsuccessful attempts to further dissect SL2 of GlmZ for elements conferring efficient cleavage (Figure 5). Irrespective of the mechanistic details, GlmZ apparently represents a novel type of RNase E substrates that uses an adaptor protein to provide entry for RNase E in a regulated manner. This concept might also hold true for other tran-

scripts. At least two further sRNAs, CsrB and CsrC, require an auxiliary protein for degradation by RNase E (46).

Finally, we report a tool allowing release of any RNA of interest in a completely 5′ monophosphorylated state. As the scissile phosphodiester bond in the RNA is determined by distance and not by sequence (Figures 6A and 7B), no foreign nucleotides are added to the 5′ cleavage products. Thus, RNA isoforms can be generated differing from naturally encoded transcripts only by bearing 5′ terminal monophosphates. So far, only a few similar devices exist. By modulating the RNA binding code of a PUF domain, a fused PIN nuclease domain can be reprogrammed to recognize a specific RNA sequence (47). However, these enzymes generate cleavage products of heterogeneous size and exhibit considerable off-target activity. Recently, RNA thermometers exposing an RNase E cleavage site upon temperature increase have been engineered (48). As a limitation, temperature has significant side effects on physiology including various processes controlled by RNase E (49). Other tools release downstream cleavage products with 5′ hydroxyl groups. This applies to ribozymes as well as many endoribonucleases used in biotechnology (50). As 5′ hydroxylated RNAs are highly stable in *E. coli* (50,51), these approaches are not suited to generate rapidly degradable cleavage products. Here, we provide preliminary evidence that the GlmZ–SL1/SL2 aptamer can be used to regulate mRNA and accordingly protein abundance by adjusting RapZ levels, which are limiting for cleavage (Figure 8). By using an inducible expression system for RapZ, target RNA cleavage and thereby expression rates may be dynamically modulated. This tool might be useful to study toxin/antitoxin systems and essential genes as it allows for rapid depletion of functional RNAs. In sRNA research it could serve to investigate whether sRNA activity or recruitment to Hfq is affected by the 5′ phosphorylation status (2). For instance, it can be used to clarify whether activation of RNase E through 5′ monophosphates provided by base-paired sRNAs is a more widespread phenomenon and is functional *in vivo* (24). Moreover, several sRNAs accumulate as processed 3′ species presenting a new 5′ end for target interactions (52,53). Using the GlmZ–SL1/SL2 module such shorter monophosphorylated sRNAs could be released on demand, facilitating their study *in vivo*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

FUNDING

Conflict of interest statement. None declared.

REFERENCES

1. Updegrove,T.B., Shabalina,S.A. and Storz,G. (2015) How do base-pairing small RNAs evolve? *FEMS Microbiol. Rev.*, **39**, 379–391.
2. Vogel,J. and Luisi,B.F. (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.*, **9**, 578–589.
3. 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.
4. Salim,N.N., Faner,M.A., Philip,J.A. and Feig,A.L. (2012) Requirement of upstream Hfq-binding (ARN)x elements in *glmS* and the Hfq C-terminal region for GlmS upregulation by sRNAs GlmZ and GlmY. *Nucleic Acids Res.*, **40**, 8021–8032.
5. Peng,Y., Soper,T.J. and Woodson,S.A. (2014) Positional effects of AAN motifs in *rpoS* regulation by sRNAs and Hfq. *J. Mol. Biol.*, **426**, 275–285.
6. Link,T.M., Valentin-Hansen,P. and Brennan,R.G. (2009) Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 19292–19297.
7. Ishikawa,H., Otaka,H., Maki,K., Morita,T. and Aiba,H. (2012) The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3′ poly(U) tail. *RNA*, **18**, 1062–1074.
8. Sauer,E., Schmidt,S. and Weichenrieder,O. (2012) Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 9396–9401.
9. Dimastrogiovanni,D., Fröhlich,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**, e05375.
10. Panja,S., Schu,D.J. and Woodson,S.A. (2013) Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. *Nucleic Acids Res.*, **41**, 7536–7546.
11. Moon,K. and Gottesman,S. (2011) Competition among Hfq-binding small RNAs in *Escherichia coli*. *Mol. Microbiol.*, **82**, 1545–1562.
12. Malecka,E.M., Strozecka,J., Sobanska,D. and Olejniczak,M. (2015) Structure of bacterial regulatory RNAs determines their performance in competition for the chaperone protein Hfq. *Biochemistry*, **54**, 1157–1170.
13. Fender,A., Elf,J., Hampel,K., Zimmermann,B. and Wagner,E.G. (2010) RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev.*, **24**, 2621–2626.
14. Henderson,C.A., Vincent,H.A., Casamento,A., Stone,C.M., Phillips,J.O., Cary,P.D., Sobott,F., Gowers,D.M., Taylor,J.E. and Callaghan,A.J. (2013) Hfq binding changes the structure of *Escherichia coli* small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of *rpoS*. *RNA*, **19**, 1089–1104.
15. Zhang,A., Schu,D.J., Tjaden,B.C., Storz,G. and Gottesman,S. (2013) Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets. *J. Mol. Biol.*, **425**, 3678–3697.
16. Laalami,S., Zig,L. and Putzer,H. (2014) Initiation of mRNA decay in bacteria. *Cell. Mol. Life Sci.*, **71**, 1799–1828.
17. Gorna,M.W., Carpousis,A.J. and Luisi,B.F. (2012) From conformational chaos to robust regulation: the structure and function of the multi-enzyme RNA degradosome. *Q. Rev. Biophys.*, **45**, 105–145.
18. Mackie,G.A. (2013) RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.*, **11**, 45–57.
19. 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.
20. 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.
21. Bandyra,K.J. and Luisi,B.F. (2013) Licensing and due process in the turnover of bacterial RNA. *RNA Biol.*, **10**, 627–635.
22. Ait-Bara,S. and Carpousis,A.J. (2015) RNA degradosomes in Bacteria and Chloroplasts: classification, distribution and evolution of RNase E homologs. *Mol. Microbiol.*, **97**, 1021–1035.

23. Morita,T. and Aiba,H. (2011) RNase E action at a distance: degradation of target mRNAs mediated by an Hfq-binding small RNA in bacteria. *Genes Dev.*, **25**, 294–298.

24. 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.

25. 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.

26. McCullen,C.A., Benhammou,J.N., Majdalani,N. and Gottesman,S. (2010) Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: pairing increases translation and protects *rpoS* mRNA from degradation. *J. Bacteriol.*, **192**, 5559–5571.

27. 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.

28. Moll,I., Afonyushkin,T., Vytvytska,O., Kaberdin,V.R. and Bläsi,U. (2003) Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA*, **9**, 1308–1314.

29. 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.

30. Göpel,Y., Papenfort,K., Reichenbach,B., Vogel,J. and Görke,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.

31. Göpel,Y., Khan,M.A. and Görke,B. (2014) Ménage à trois: Post-transcriptional control of the key enzyme for cell envelope synthesis by a base-pairing small RNA, an RNase adaptor protein and a small RNA mimic. *RNA Biol.*, **11**, 433–442.

32. Kalamorz,F., Reichenbach,B., März,W., Rak,B. and Görke,B. (2007) Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*. *Mol. Microbiol.*, **65**, 1518–1533.

33. Urban,J.H. and Vogel,J. (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol.*, **6**, e64.

34. Reichenbach,B., Maes,A., Kalamorz,F., Hajnsdorf,E. and Görke,B. (2008) The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of *glmS* expression and is subject to regulation by polyadenylation in *Escherichia coli*. *Nucleic Acids Res.*, **36**, 2570–2580.

35. Wilson,G.G., Young,K.Y., Edlin,G.J. and Konigsberg,W. (1979) High-frequency generalised transduction by bacteriophage T4. *Nature*, **280**, 80–82.

36. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 6640–6645.

37. Resch,M., Göpel,Y., Görke,B. and Ficner,R. (2013) Crystallization and preliminary X-ray diffraction analysis of YhbJ from Escherichia coli, a key protein involved in the GlmYZ sRNA regulatory cascade. *Acta Crystallogr. F Struct. Biol. Crystallization Commun.*, **69**, 109–114.

38. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.

39. Guillier,M. and Gottesman,S. (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.*, **59**, 231–247.

40. Foley,P.L., Hsieh,P.K., Luciano,D.J. and Belasco,J.G. (2015) Specificity and Evolutionary Conservation of the *Escherichia coli* RNA Pyrophosphohydrolase RppH. *J. Biol. Chem.*, **290**, 9478–9486.

41. Wassarman,K.M., Repoila,F., Rosenow,C., Storz,G. and Gottesman,S. (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev.*, **15**, 1637–1651.

42. Andersen,J.B., Sternberg,C., Poulsen,L.K., Bjorn,S.P., Givskov,M. and Molin,S. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl. Environ. Microbiol.*, **64**, 2240–2246.

43. 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.

44. Bilusic,I., Popitsch,N., Rescheneder,P., Schroeder,R. and Lybecker,M. (2014) Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. *RNA Biol.*, **11**, 641–654.

45. 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.

46. Suzuki,K., Babitzke,P., Kushner,S.R. and Romeo,T. (2006) Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. *Genes Dev.*, **20**, 2605–2617.

47. Choudhury,R., Tsai,Y.S., Dominguez,D., Wang,Y. and Wang,Z. (2012) Engineering RNA endonucleases with customized sequence specificities. *Nat. Commun.*, **3**, 1147.

48. Hoynes-O'Connor,A., Hinman,K., Kirchner,L. and Moon,T.S. (2015) De novo design of heat-repressible RNA thermosensors in *E. coli*. *Nucleic Acids Res.*, **43**, 6166–6179.

49. Afonyushkin,T., Moll,I., Bläsi,U. and Kaberdin,V.R. (2003) Temperature-dependent accessibility and translation of *Escherichia coli* ompA mRNA. *Biochem. Biophys. Res. Commun.*, **311**, 604–609.

50. Qi,L., Haurwitz,R.E., Shao,W., Doudna,J.A. and Arkin,A.P. (2012) RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.*, **30**, 1002–1006.

51. Celesnik,H., Deana,A. and Belasco,J.G. (2007) Initiation of RNA decay in *Escherichia coli* by 5′ pyrophosphate removal. *Mol. Cell*, **27**, 79–90.

52. 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.

53. 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.