sRNA-dependent control of curli biosynthesis in Escherichia coli

McaS directs endonucleolytic cleavage of csgD mRNA

Andreassen, Patrick Rosendahl; Pettersen, Jens Sivkær; Szczerba, Mateusz; Valentin-Hansen, Poul; Møller-Jensen, Jakob; Jørgensen, Mikkel Girke

Published in:
Nucleic Acids Research

DOI:
10.1093/nar/gky479

Publication date:
2018

Document version:
Final published version

6746–6760 *Nucleic Acids Research, 2018, Vol. 46, No. 13* *Published online 14 June 2018*
doi: 10.1093/nar/gky479

sRNA-dependent control of curli biosynthesis in *Escherichia coli*: McaS directs endonucleolytic cleavage of *csgD* mRNA

Patrick Rosendahl Andreassen[1], Jens Sivkær Pettersen[1], Mateusz Szczerba[2,3], Poul Valentin-Hansen[1], Jakob Møller-Jensen[1] and Mikkel Girke Jørgensen[1,*]

[1]Department of Biochemistry and Molecular Biology, University of Southern Denmark. Campusvej 55, 5230 Odense M. Denmark, [2]Ira A. Fulton Schools of Engineering and School of Life Sciences, Arizona State University, Tempe, AZ, USA and [3]Biodesign Center for Immunotherapy, Vaccines, and Virotherapy (B-CIVV), Biodesign Institute, Arizona State University. 727 East Tyler Street, Tempe, AZ 85287-5001, USA

Received January 12, 2018; Revised April 30, 2018; Editorial Decision May 15, 2018; Accepted May 25, 2018

ABSTRACT

Production of curli, extracellular protein structures important for *Escherichia coli* biofilm formation, is governed by a highly complex regulatory mechanism that integrates multiple environmental signals through the involvement of numerous proteins and small non-coding RNAs (sRNAs). No less than seven sRNAs (McaS, RprA, GcvB, RydC, RybB, OmrA and OmrB) are known to repress the expression of the curli activator CsgD. Many of the sRNAs repress CsgD production by binding to the *csgD* mRNA at sites far upstream of the ribosomal binding site. The precise mechanism behind sRNA-mediated regulation of CsgD synthesis is largely unknown. In this study, we identify a conserved A/U-rich region in the *csgD* mRNA 5′ untranslated region, which is cleaved upon binding of the small RNAs, McaS, RprA or GcvB, to sites located more than 30 nucleotides downstream. Mutational analysis shows that the A/U-rich region as well as an adjacent stem–loop structure are required for McaS-stimulated degradation, also serving as a binding platform for the RNA chaperone Hfq. Prevention of McaS-activated cleavage completely relieves repression, suggesting that endoribonucleolytic cleavage of *csgD* mRNA is the primary regulatory effect exerted by McaS. Moreover, we find that McaS-mediated degradation of the *csgD* 5′ untranslated region requires RNase E.

INTRODUCTION

Biofilms are communities of microorganisms that aggregate and adhere to a surface (usually solid–liquid) encased in a self-produced matrix consisting of extracellular polymeric substances (EPS). This EPS is composed of polysaccharides, proteins, lipids and DNA (1). The sessile growth phenotype is entirely different from the independent single-cell growth mode of planktonic bacteria. The biofilm structure provides mechanical stability and protects the bacterial cells from various environmental conditions such as UV radiation, dehydration, and provides strong resilience against antibiotics (2,3). Biofilms provide an important bacterial reservoir that may serve as nidus for recolonization of infected hosts. This feature is believed to be the leading cause of recalcitrant chronic bacterial infections (4).

The main proteinaceous component of the extracellular matrix produced by *Escherichia coli* biofilms is curli (5). Curli are amyloid fibers that mediate surface adhesion and cell aggregation of *E. coli* and *Salmonella* spp. species and have been implicated in both intestinal and extraintestinal infections (6,7). The curli fibers are composed of the main structural protein, CsgA, and the nucleator protein CsgB, which are encoded by the *csgBAC* operon (8,9). CsgC acts as a periplasmic chaperone that prevents CsgA from forming toxic intracellular aggregates (10). A second operon, *csgDEFG*, encodes proteins required for the fiber assembly and secretion (CsgE, CsgF and CsgG) as well as a positive transcriptional regulator of curli synthesis CsgD, which in turn activates transcription of the *csgBAC* operon (11). The regulation of curli biosynthesis is extraordinarily complex. Multiple regulatory cascades process and integrate signals into the curli synthesis. The master regulator, CsgD, is crucially controlled by the stationary sigma-factor (RpoS) of RNA polymerase (12). Furthermore, the bacterial signalling molecule bis-(3′–5′)-cyclic di-GMP (c-di-GMP) is also required for activating CsgD expression. C-di-GMP is produced by diguanylate cyclase enzymes (DGCs) and the molecule is degraded by phosphodiesterase enzymes (PDEs). C-di-GMP is an important biofilm messenger in the motile-to-adhesive 'lifestyle' switch of *E. coli* (13,14).

*To whom correspondence should be addressed. Tel: +45 65502334; Email: mikkelj@bmb.sdu.dk

In addition to the RpoS/c-di-GMP cascades leading to curli expression, at least three two-component systems regulate the expression of *csgD*. OmpR/EnvZ responds to changes in osmolarity and is essential for *csgD* expression. CpxA/R is a cell envelope stress response pathway responding to misfolded proteins repressing *csgD* expression. Similarly, the Rcs phosphorelay system is an envelope stress-controlled pathway that negatively affects the expression of *csgD* (15–20).

Besides the abundant transcriptional control of the curli operons, the *csgD* transcript serves as a hub for post transcriptional regulation by small non-coding RNAs (sRNAs). With the help from Hfq no less than seven sRNAs negatively influence the expression of *csgD* transcription—these are OmrA, OmrB, McaS, RprA, GcvB, RydC and RybB (21–26). The most common mechanism by which sRNAs regulate protein synthesis is by specific base pairing with sequences overlapping with the target mRNA ribosomal binding site (RBS), thereby interfering with translation initiation (27). Interestingly, the *csgD* mRNA contains an unusually long 5′ untranslated region (UTR) of 147 nucleotides (nt) and many of the sRNAs bind far upstream of the RBS. The exact mechanism by which many of these sRNAs inhibit CsgD synthesis and thereby curli production is not yet fully understood.

OmrA and OmrB are two redundant sRNAs located adjacent to each other on the chromosome. Whereas their 5′ and 3′ ends contain identical sequences, the central regions differ, indicating that they have both common and distinct targets (28). Both OmrA and OmrB bind *csgD* mRNA on the 5′ side of a conserved stem–loop structure, and base-pairing to nucleotides in a bulge is essential for regulation. *In vitro*, binding of OmrA and OmrB to *csgD* mRNA inhibits formation of the translation initiation complex in a manner independent of the RBS itself (21).

RprA binds *csgD* mRNA in three different regions: two distal sites far from the RBS and one overlapping with the Shine-Dalgarno (SD) sequence. Mutations in either of these sites do not alleviate *csgD* repression completely, whereas double mutants do (23). RybB, a major RpoE dependent sRNA known to downregulate multiple outer membrane porins, has two binding regions in the *csgD* UTR where one site overlaps with the SD sequence (26). Additionally, RydC has base pair complementarity to the SD region of *csgD* (25).

The precise mechanism by which the small RNAs McaS and GcvB regulate CsgD production is unknown. Both have *csgD* mRNA binding sites upstream of the RBS. Moreover, an *in vitro* structural probing assay suggests that McaS has extended base-paring with *csgD* mRNA. This extended base-paring allows McaS binding in proximity of the RBS and thereby directly prevents assembly of the ribosomal initiation complex (22).

Here, we present new insight into the mechanism by which curli-repressing sRNAs inhibit CsgD biosynthesis and thus biofilm formation. We show that McaS, and to some extent GcvB and RprA, induce endonucleolytic cleavage in a conserved A/U-rich region. This region is important for both binding of the RNA chaperone Hfq and induced degradation of *csgD* mRNA. In addition, we show that McaS-induced cleavage dependent on RNase E.

MATERIALS AND METHODS

Strains, plasmids and oligonucleotides

Bacterial strains and plasmids used in this study are listed in s*upplementary data* Supplementary Table S1 and S2; also their construction is described in *supplementary data*. Primers used for the construction of mutant strains, plasmids and DNA templates are listed in Supplementary Table S3.

Growth conditions

Cells of *E. coli* K12 were grown in standard Luria–Bertani (LB) broth at 37°C. When required, the medium was supplemented with antibiotics: ampicillin (30 μg/ml), chloramphenicol (30 μg/ml). The expression of curli was visualized on Congo red agar plates without the addition of NaCl (Bacto-Tryptone 10 mg/ml, yeast extract 5 mg/ml, Congo red 40 mg/ml, Coomassie Brilliant Blue G 10 mg/ml). The plates were incubated at 28°C for 48 h. Expression from the P_{lac} promoter was induced by addition of 1 mM IPTG and the expression from the P_{BAD} promoter was induced by addition of 0.2% arabinose.

Transmission electron microscopy

To visualise extracellular curli fibers, samples were placed on 200-mesh carbon coated grids for 1 min. The samples were washed with deionized water and negatively stained with 3% (w/v) uranyl acetate for 30 s. Images were acquired using a JEM-1200EX II (Joel Germany) microscope at an acceleration of 120 kV.

RNA purification

Total RNA was extracted from 10 ml culture. Each sample was spun, and pellet was re-suspended in 150 μl solution 1 (10 mM Na-citrate, 10 mM Na-acetate pH 4.5 and 2 mM EDTA), transferred to a phase-lock tube from 5 PRIME containing 150 μl solution 2 (10 mM Na-acetate pH 4.5 and 2% SDS), 300 μl chloroform and 700 μl acidic phenol pH 4.5. The tube was heated at 80°C for 5 min and spun at 14 000 rcf for 5 min to separate the phases. The aqueous phase was transferred to 1 ml 96% ethanol to precipitate the RNA. The RNA pellet was resuspended in RNase-free H_2O.

Northern blot experiments

Ten microgram total RNA was supplemented with 2 μl formamide loading buffer and separated on 6% polyacrylamide gels for 1.5–2.5 h at 300 V. Separated RNA was semi-dry blotted onto a Zeta-probe nylon membrane (Bio Rad) for 1 h at 400 mA, followed by cross-linking of RNA to the membrane with UV radiation. Probes listed in *supplementary data* were 5′-labeled with ^{32}P-ATP using T4-polynucleotide kinase (NEB). Membranes were pre-hybridized for 10 min at 42°C before probing with 5′-labeled oligos overnight (ON). Probed membranes were washed three times in 2× SCC and 0.1% SDS for 10 min at 42°C, dried and visualized by phosphorimaging on a Typhoon Trio scanner (GE Healthcare).

Western blot experiments

Samples were spun and cell pellets were re-suspended in 1× SDS loading buffer (3% SDS, 10% glycerol, 50 mM Tris–HCl pH 6.8, 0.1% bromophenol blue, 12.5 mM EDTA, 100 mM DTT) to a concentration of approximately 10^6 cells/μl and boiled at 95°C for 5 min. 10 μl of each sample was separated in 4–12% NuPAGE® Bis-Tris gels (Invitrogen) at 190 V for 45 min, and blotted onto a nitrocellulose membrane (Invitrogen) for 1 h at 150 mA. Blocking and antibody addition was done using the SNAP i.d.® 2.0 system (Merck Millipore). α-FLAG (Sigma-Aldrich) and α-GroEL (Sigma-Aldrich) monoclonal antibodies were diluted 1:20 000 and 1:50 000, respectively. Mouse and rabbit horseradish peroxidase-conjugated secondary antibodies were diluted 1:2000 (Dako Cytomation).

Primer extension analysis

Primer extension analysis was carried out using the PXT primer listed in *supplementary data*. The primer was 5′-labeled with ^{32}P-ATP using T4-polynucleotide kinase (NEB). Labeled primer was hybridized to 10 μg RNA, and extended with reverse transcriptase (Promega). Sanger sequencing was carried out with labeled PXT primer, PCR amplified *csgD* promoter and UTR as template and one of the ddNTPs (G, A, T, C) mixed with all four dNTPs in four different PCR reactions. The labeled cDNA and sequence of *csgD* were separated in 8% polyacrylamide gels, which were dried and visualized by phosphorimaging on a Typhoon Trio scanner (GE Healthcare).

In vitro synthesis and labeling of RNA

csgD RNAs were prepared by *in vitro* transcription (Megascript, Ambion) using 5 ug of DNA from a purified PCR product as template using the T7 oligos listed in Supplementary Table S3. The RNA transcripts were DNase I-treated before PAGE separation on a 4.5% 7 M urea denaturing gel. The RNA was extracted by electro elution using GeBAflex tubes (GerardBiotech). RNA transcripts were 5′-labeled with ^{32}P-ATP using T4-polynucleotide kinase (NEB).

Electrophoretic mobility shift assays

The *E. coli* Hfq protein was expressed and purified using the Impact System (New England Biolabs) as described in (29).

Binding reactions were carried out in 10 μl volumes in 1× binding buffer (20 mM Tris, pH 8, 100 mM KCl, 1 mM MgCl2, 1 mM DTT) and incubated at 37°C for 10 min before addition of 5 μl of loading buffer (50% glycerol). Hfq-*csgD* mRNA interactions were carried out in 1× binding buffer using 2 nM 5′ end-labeled *csgD* mRNA, 10 ug tRNA in the presence of 0, 0.25, 0.5, 0.75, 1 or 2 μM Hfq (monomer concentration). Samples were separated by native electophoresis on 5% polyacrylamide gels at 4°C in 0.5× TBE. Samples were visualized by phosphoimaging on a Typhoon Trio scanner (GE Healthcare). The relative intensity of the shifted bands was quantified and fitted to a sigmoid curve, and dissociation constant (K_d) values were determined with SigmaPlot.

Structural probing assay

2 nM of 5′ end-labeled csgD 5′ UTR RNA was mixed with increasing concentrations of Hfq (0, 0.25, 0.5 and 2 μM), McaS (10, 20, 40 nM) or both (0, 0.25, 0.5, 2 μM Hfq mixed with 10, 20, 40, 80 nM McaS) in 1× RNA Structure Buffer (Ambion) and incubated at 37°C for 10 min. Samples were cooled at room temperature before addition of 2.5 nM lead(II)acetate or 1 U of Nuclease S1 (Promega). After incubation for 5 min at 37°C, 10 μl of the samples were transferred to 10 μl of cold formamide loading dye. RNase T1 digestions were carried out in 1× Sequence Buffer with 0.1 U of RNase T1 (Ambion) for 5 min at 37°C before transferring it to cold formamide loading dye. The samples were run at 55 W on 8% denaturing PAGE gel along with an alkaline hydrolysis ladder of the 5′ end-labeled *csgD* 5′ UTR RNA.

RESULTS

Induction of McaS inhibits the synthesis of curli amyloid fibers

We previously showed that ectopical co-expression of McaS with *csgD* mRNA rapidly downregulates the cellular levels of *csgD* mRNA. Furthermore, expression of McaS in *E. coli* K-12 wild type-cells showed a white morphotype when streaked on Congo red agar plates, indicative of a curli deficient phenotype (22). To see the direct effect of McaS on curli fiber biosynthesis we inspected a wild-type strain carrying the *mcaS* gene under control of an IPTG-inducible *lac* promoter on a low-copy-number R1 plasmid (pNDM-mcaS) by high-resolution transmission electron microscopy (TEM). When grown under curli inducing conditions, ectopic production of McaS in wild-type cells resulted in a complete loss of curli fibers and visibly naked cells (Figure 1). Moreover, these curli deficient cells showed an increase in flagella production. In contrast, wild-type cells harboring the empty vector pNDM220 produced curli fibers and only few visible flagella. These findings are consistent with the observation that McaS production leads to increased motility by activating FlhD the master regulator of flagella synthesis (24).

The secondary structure of 5′ UTR of csgD mRNA is conserved among bacterial species

To identify regions important for sRNA-mediated regulation, we aligned the 5′ untranslated region (UTR) of *csgD* including the promoter regions from bacterial species encoding the sRNA McaS (22). The *csgD* promoter and +1 start site (identified with primer extension, Figure 8), Shine-Dalgarno sequence and the target sequence for McaS binding are conserved (Figure 2A). The alignment further revealed that three predicted stem–loop elements (I, II and III) were preserved, suggesting that the mRNA 5′ UTR structure is conserved too. This is supported by *in vitro* structural probing of the *E. coli csgD* mRNA (Figures 2B and 7) (21,25). Other than stem–loop I, II and III, the 5′end of the *csgD* UTR was not predicted to take up any consensus secondary structure (Figure 2C). Interestingly, however, this region contained a conserved A/U-rich region immediately upstream of stem–loop I (Figure 2A).

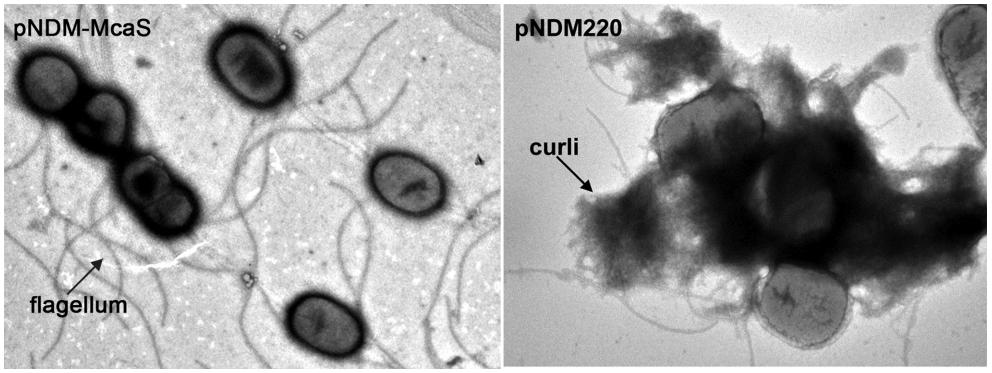

Figure 1. McaS producing cells lack extracellular curli fibers. Close inspection of the inhibitory effect of McaS on curli fibers. Wild-type cells of SØ928Δ*mcaS* producing McaS from a P_{lac} containing low-copy plasmid (pNDM-mcaS) and SØ928Δ*mcaS* cells carrying the empty vector (pNDM220) were grown in curli stimulating conditions on LB-agar plates (without NaCl) at 28°C. High-resolution EM images were taken after 48 hours of growth in the presence of 100 μM IPTG inducer.

Identification of sRNA-stimulated cleavage sites in *csgD* mRNA

Commonly, *trans*-encoded sRNAs induce degradation of their target mRNAs, and McaS co-degrades with *csgD* (22,30). We decided to identify the exact *csgD* 5′ UTR cleavage sites during *in vivo* induction of the regulatory sRNAs McaS, RprA, GcvB and OmrA by primer extension analysis. We constructed a minimal target molecule consisting of the *csgD* 5′ UTR plus 87 nucleotides of the coding region fused to the small RNA ChiX for stabilization of the 3′ end. This short target, termed $csgD^{chiX}$, allowed for better separation of cleavage products in polyacrylamide gels. Expression of sRNAs were induced from P_{lac} by addition of IPTG from a low-copy number mini R1 plasmid (pNDM220), and the target $csgD^{chiX}$ RNA was expressed from a pBAD33 derivative by addition of arabinose. To reduce interference from the chromosomally encoded sRNAs, these experiments were performed in a strain deleted for the sRNAs McaS, RprA, GcvB, OmrA, and OmrB. For simplicity we call this strain SØ928Δ*5*. McaS strongly induced RNA degradation while OmrA did not stimulate degradation. RprA and GcvB also stimulated cleavage but to a lesser extent than McaS. The most apparent cleavage sites were identified between residues G-132 and U-131, and U-131 and U-130, which correspond to the 5′ end of the conserved A/U-rich region (Figure 3A and B). Several other less prominent cleavage sites were identified, all located within the A/U-stretch (A-127 to U-123, A-116 to U-114 and U-102 to A-101), except for two cleavage sites induced specifically by GcvB (A-93 to A-91) (Figure 3A). Thus, McaS, RprA and GcvB appear, at least in part, to regulate *csgD* expression by a common mechanism – one that is different from that of OmrA and OmrB.

Only the distally located McaS-binding site within the *csgD* 5′UTR is necessary *in vivo*

Previously, we identified two McaS-binding sites on *csgD* 5′UTR. Mutational analysis and structural probning experiments showed that McaS directly interacts with *csgD* 5′UTR at two sites (22). The first site (*Binding Site 1*) is located far upstream the AUG start codon and the second site (*Binding Site 2*) is located close to, and overlapping with, the ribosomal binding site (Figure 4A). Moreover, a toeprinting assay revealed that $tRNA^{fMet}$-dependent 30S complex at the *csgD* TIR (translation initiation region) was lost upon McaS pre-incubation *in vitro*. This suggested that McaS binding to *Binding Site 2* on *csgD* inhibits translation initiation. Thus, two possibilities might explain the regulatory mechanism by which McaS regulates CsgD synthesis. One mechanism relies on the conventional pattern where the sRNA binds to the translational initiation region of the target mRNA and prevents translation initiation. The other mechanism results in coupled degradation of sRNA and the target mRNA to control CsgD expression.

We decided to test the importance of the two McaS base-paring sites on CsgD production *in vivo*. We examined the ability of McaS to downregulate the expression of two mutant *csgD* alleles by tagging the very end of the *csgD* gene with a sequence encoding a triple FLAG epitope. The mutant alleles each carried a deletion, which interrupted the McaS binding site with *Binding Site 1* or *Binding Site 2,* respectively (Figure 4A). We transiently induced expression of McaS from pNDM220 for 10 min followed by short-term expression of $csgD^{FLAG}$ and mutant alleles. Western analysis showed that McaS regulated the $CsgD^{FLAG}$ expression with a *Binding Site 2* truncation as well as wild-type *csgD*, whereas disruption of *Binding Site1* completely disrupted regulation, indicating that *Binding Site 2* plays little role in McaS dependent CsgD regulation (Figure 4B).

McaS induces *csgD* mRNA cleavage at the conserved A/U-rich region

To investigate the importance of sequence and structure motifs for the action of McaS, we constructed a set of five *csgD* mutants and evaluated target RNA stability by northern blotting using a ChiX-specific probe (Figure 5A). The $csgD^{chiX}$ probe binds the 3′ end of the *csgD-chiX* fusion and thus only the full length and 3′end degradation products of *csgD* are able to hybridize to the probe. In a parallel set of experiments, we monitored $CsgD^{FLAG}$ protein levels (Figure 5C). We performed transient expression assays where McaS was turned on by addition of IPTG for 10 min followed by short-term induction of *csgD* by addition of ara-

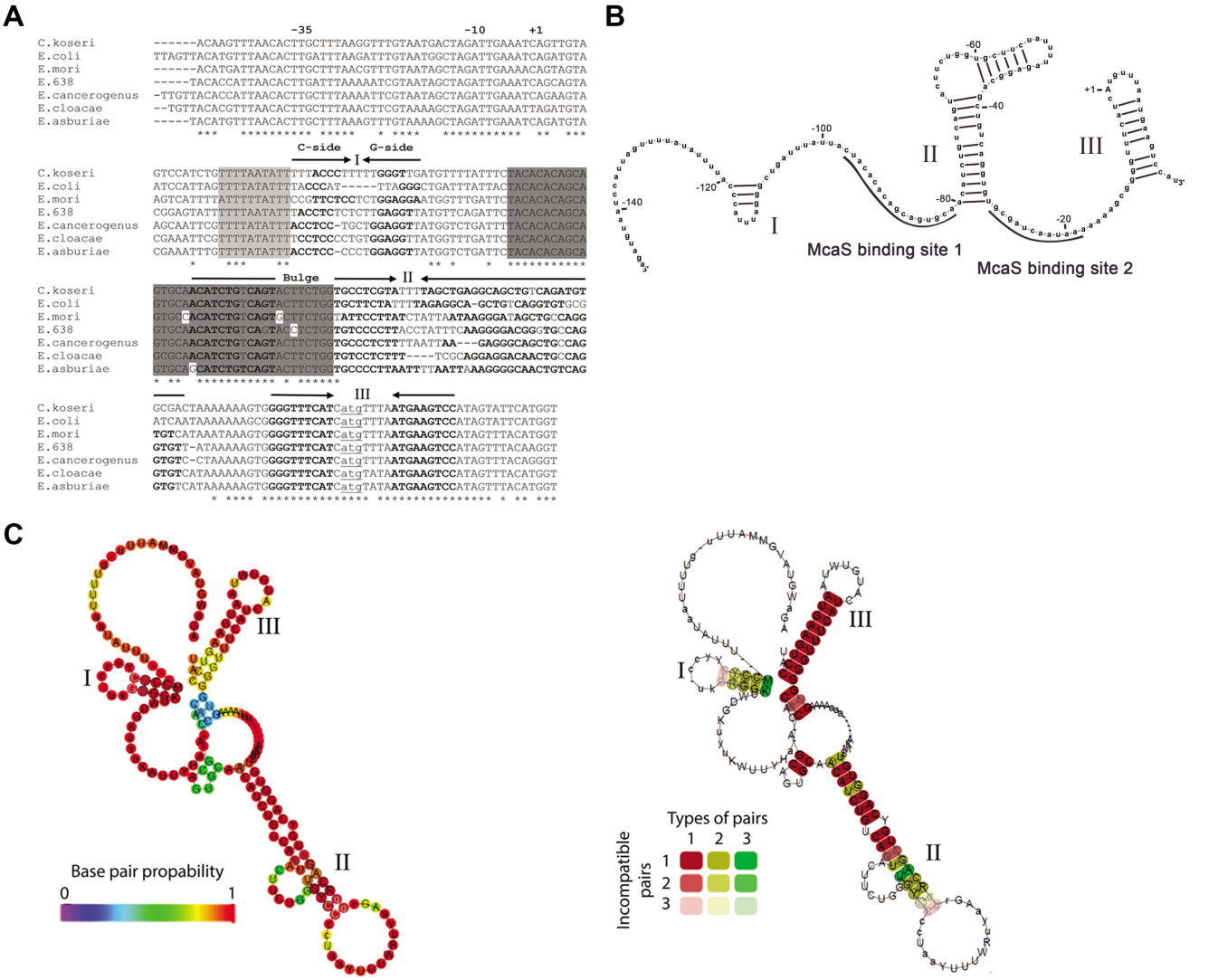

Figure 2. An A/U-rich stretch followed by a small stem–loop structure is conserved in the *csgD* mRNA. *csgD* from seven different species, all having a conserved *mcaS* gene, was aligned using Clustal Omega Multiple Sequence Alignment tool. (**A**) The promoter elements (−35 and −10) and transcriptional start site (+1) are highly conserved, as well as the McaS-binding site (dark gray), the RBS and stem–loop II. Furthermore, both an A/U-rich region (light gray) and a C and a G rich stretch (C-side and G-side, respectively) just downstream the A/U-rich region are conserved. Arrows indicate possible stem–loop structures (numbered I, II and III), bold letters indicate nucleotide residues engaging in the base-pairing of the stem–loops, and underlined lower-case letters indicate the translation start codon. (**B**) The *E. coli csgD* 5′UTR used in this study adopts an overall similar fold as the consensus folding shown in C. The *E. coli* structure is based on the structural probing presented in Figure 7 and supplementary data. The two McaS-binding sites are highlighted (**C**) RNAalifold was used to predict a generic consensus secondary structure of the *csgD* alignment obtained from Clustal Omega. The consensus sequence predicts *csgD* to contain three stem–loop structures: the first just downstream of the A/U-rich region (I), a second central large stem–loop structure upstream of the RBS containing a bulge (II), and a third stem–loop structure containing the RBS (III).

binose for 5 min. As observed for full-length *csgD* 5′UTR, short-term induction of McaS resulted in clear *csgD* degradation products and a corresponding decrease in CsgD protein levels (Figure 5B and C, lanes 2 and 3). Deletion of the entire conserved A/U-rich region (ΔA/U) completely prevented cleavage at this site and made the target RNA less responsive to McaS-mediated degradation and consequently, western analysis of CsgD protein levels showed that downregulation by McaS was lost (Figure 5B and C, lanes 4 and 5). Mutations in the conserved A/U-rich region by substitution of four nucleotides with four guanines (GGGG), changed the target RNA degradation pattern and decreased translational inhibition, but to a lesser extent than ΔA/U (Figure 5B and C, lanes 6 and 7). Deletion of stem–loop I (Δstem) inhibited McaS-dependent cleavage in the conserved A/U-rich region and decreased its translational inhibition, however to a lesser degree than ΔA/U (Figure 5B and C, lanes 8 and 9). The two remaining mutants were created in order to force an alternative secondary structure within the *csgD* 5′ UTR (Forced). In both mutants, the 3′ half of the conserved A/U-rich region and the 5′ half of the conserved stem–loop was replaced with a stable tetraloop to induce formation of a large stem masking the cleavage sites. Further substitution of a single uracil for cytosin in the 5′ end of the conserved A/U-rich region (Forced+C) completely prevented cleavage at that site and

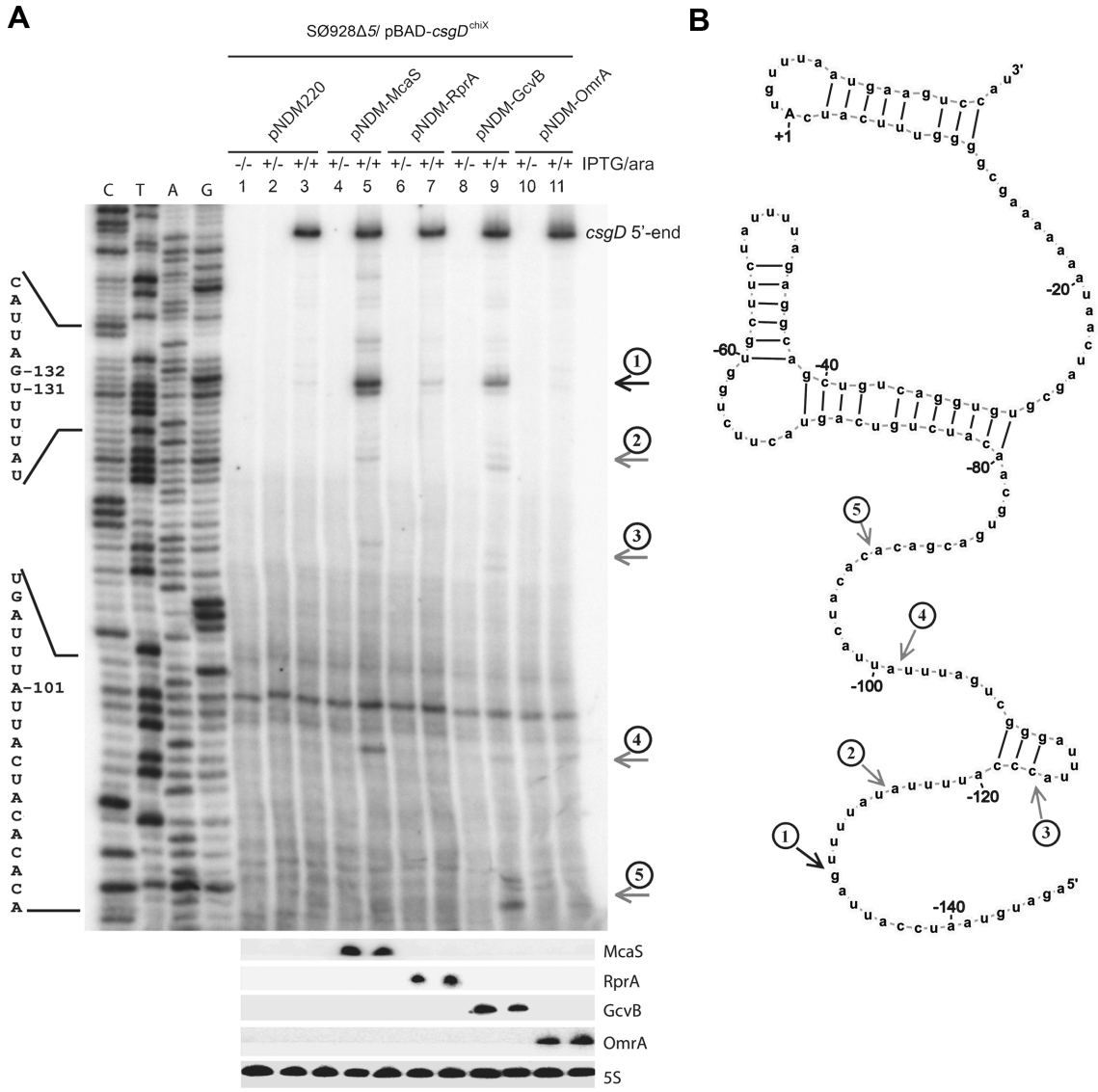

Figure 3. McaS, RprA, and GcvB stimulate *csgD* mRNA cleavage in the conserved A/U-rich region. (**A**) Primer extension and northern blot with RNA purified from SØ928Δ5 carrying pBAD-csgD^{chiX} and either pNDM-mcaS, pNDM-rprA, pNDM-gcvB, pNDM-omrA or the empty vector pNDM220. Cultures were grown in LB broth at 37° to an OD_{600} of 0.6 (−/−). The cultures were induced with 1 mM IPTG for 10 min (+/−), followed by 5 min induction with 0.2% arabinose (+/+). Primer extension and northern blot was performed as described in methods. (**B**) *csgD* mRNA sequence with cleavage sites indicated. The most apparent cleavage sites (①) are marked with black arrows and the remaining (②③④⑤) with grey arrows. ④is specific to McaS, while ⑤is specific to GcvB. McaS, RprA and GcvB all induce cleavage in the conserved A/U-rich region (gray bar).

greatly decreased McaS-dependent translational inhibition (Figure 5B and C, lanes 10 and 11). Mutation of two adjacent uracils located on the descending strand of the stem just upstream of the sRNA binding site (corresponding to cleavage site ④ in Figure 3) into guanines, inhibited cleavage in the conserved A/U-rich region and decreased McaS-dependent translational inhibition greatly (Figure 5B and C, lanes 12 and 13). Taken together, the Northern- and Western blots demonstrate that prevention of cleavage in the conserved A/U-rich region alleviates McaS-dependent repression of CsgD biosynthesis, suggesting that McaS induces *csgD* mRNA cleavage at the conserved A/U-rich region thereby inhibiting translation. Furthermore, we conclude that the conserved A/U-rich region itself as well as the local secondary structure are important for efficient McaS repression of CsgD expression *in vivo*.

csgD mRNA degradation is Hfq-dependent

The function of most sRNAs in *E. coli* is greatly affected by Hfq, which binds and stabilises sRNA molecules and mediates their base-pairing with cognate target mRNAs (29). Hfq further exhibits RNA chaperone activity affecting RNA secondary structure (31). Hfq recognizes single-stranded A/U-rich or ARN_n motifs (where A is adenine, R is a purine, and N is any nucleotide) sequences, which are often preceded or followed by stem–loop structures (32,33). It interacts directly with RNase E and PNPase and has there-

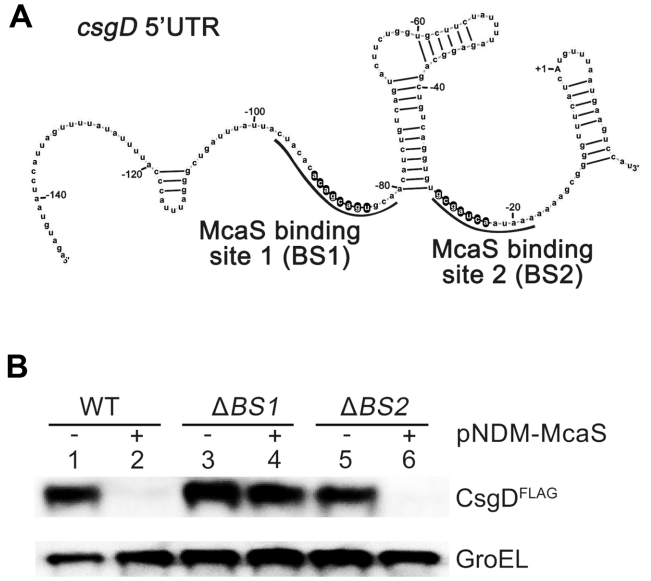

Figure 4. Only the upstream binding site is necessary for McaS-mediated regulation of CsgD protein levels. (**A**) Schematic presentation of the *csgD* 5′UTR indicating the two sites of *csgD* complementary to McaS (underlined). Deletion mutants in binding sites 1 and 2 respectively are highlighted by black circles. (**B**) *E. coli* strain SØ928 carrying either a wild-type $csgD^{FLAG}$ construct or mutant alleles on a pBAD plasmid vector. Expression from the empty vector control pNDM220 (−) or the isogenic plasmid borne McaS (+) was induced by addition of 1 mM IPTG. After 10 min of incubation, 0,2% arabinose was added to induced expression of the $csgD^{FLAG}$ genes. After 5 min of induction samples for western blot analysis was taken. GroEL was probed as loading control.

fore been suggested to recruit RNases to cleave the sRNA–mRNA complexes (34,35).

We used electrophoretic mobility shift assays (EMSA) to investigate the interaction between Hfq and the *csgD* 5′UTR *in vitro*. To examine the interaction in detail, we measured the dissociation constant (K_d) to describe the affinity between Hfq and *csgD*. Deletion of the conserved A/U-rich region as well as forcing an alternative secondary structure (ΔA/U, Forced+C and Forced+GG described above) reduced the affinity between Hfq and *csgD* mRNA molecules (Figure 6). Deletion of the stem–loop I structure (Δstem) did not impact complex formation and substitution of the conserved A/U-rich region with four guanine nucleotides (GGGG) slightly reduced complex formation. Taken together with the results from Figure 5, McaS-stimulated cleavage in the conserved A/U-rich region and the efficiency of McaS-dependent translational inhibition of CsgD does not correlate directly with changes in Hfq binding affinities. However, CsgD synthesis inhibition is at least in part dependent on Hfq binding to a region close to or at the conserved A/U-rich region. This is consistent with Hfq involvement in the recruitment of the RNA degradation machinery to the cleavage sites at mRNA substrates (30). Interestingly, at high Hfq concentrations, we see higher order multimers of *csgD* and Hfq complexes suggesting that Hfq interacts with *csgD* at multiple sites.

Another explanation for the induction of cleavage in the A/U-rich sequence could be a conformational change within *csgD* upon McaS and Hfq binding (36). To pinpoint the exact Hfq-binding sites on *csgD* 5′UTR we performed structural probing experiments using Nuclease S1, which specifically degrades single-stranded RNA (or DNA). This simultaneously allowed Hfq and McaS binding site identifications as well as probing the overall structure of the *csgD* 5′UTR. Figure 7 shows the probing of *csgD* 5′UTR alone and with increasing concentrations of Hfq, McaS, or both. The probing was visualized on different acrylamide gels to improve the resolution of both the very 5′end and the far upstream located translation initiation region (Figure 7A). We observed several Hfq-binding sites within the *csgD* 5′ UTR (Figure 7B). We identified one site at the A/U rich sequence in the 5′end extending to include the AUUUA hairpin (Site I). This site is located at nucleotides –128 to –112 between the cleavage region and the McaS base-paring region. This site is close to, but not overlapping, the strong cleavage sites identified in Figure 3. Moreover, Hfq cooperates in complex formation and facilitate *csgD*-McaS interaction (Figure 7, lanes 11–14). In addition, we observed two additional sites located close to the translational initiation region, Site II is positioned in the single-stranded A-tract region preceding the Shine-Dalgarno sequence, and Site III is located right after the AUG start codon in the open reading frame (Figure 7B). Indeed, Hfq binding Site I and II were recently suggested as functional Hfq binding sites (33). The protection observed with Nuclease S1 at Site II and Site III were weak, so we decided to validate and confirm these sites with Pb^{2+}-probing (Supplementary data Figure S4). We also detected protection in a bulge in the conserved stem–loop structure (*). However, we note that the protection is weak and only present at high Hfq concentrations, and the protected sequence does not constitute the typical Hfq recognition sequences (i.e. an A/U tract or an ARN motif) (33). In addition, the Nuclease S1 probing suggested that Hfq binds to the hairpin of the conserved stem–loop (**), however, we could not confirm this site by Pb^{2+}-probing. Furthermore, binding of Hfq and McaS to the *csgD* 5′UTR did not result in an overall structural rearrangement (Figure 7).

RNase E is required for McaS-dependent degradation of *csgD*

McaS, RprA and GcvB all stimulate endoribonucleolytic cleavage of *csgD* mRNA. To further our understanding of the sRNA dependent cleavage of *csgD* mRNA we examined several ribonucleases as specific mediators of McaS-*csgD* decay. Several sRNA-mRNA complex responsive ribonucleases exist in *E. coli* including RNase III and RNase E (37). In addition, a recently identified RNase, YbeY, was considered as a candidate for sRNA-stimulated *csgD* mRNA cleavage since it modulates Hfq dependent and independent sRNA-mRNA interactions (38).

RNase III exists as a dimer and cuts specifically at double-stranded RNA, which allows further processing of the resulting fragments (39,40). RNase III and YbeY are not essential in *E. coli*, and McaS-stimulated regulation of curli synthesis could therefore be assayed in the corresponding mutant backgrounds on agar plates containing Congo red, which binds to the amyloid curli fibers to produce a red color. RNase III- and YbeY-deficient strains were cultured alongside their respective isogenic strains on Congo

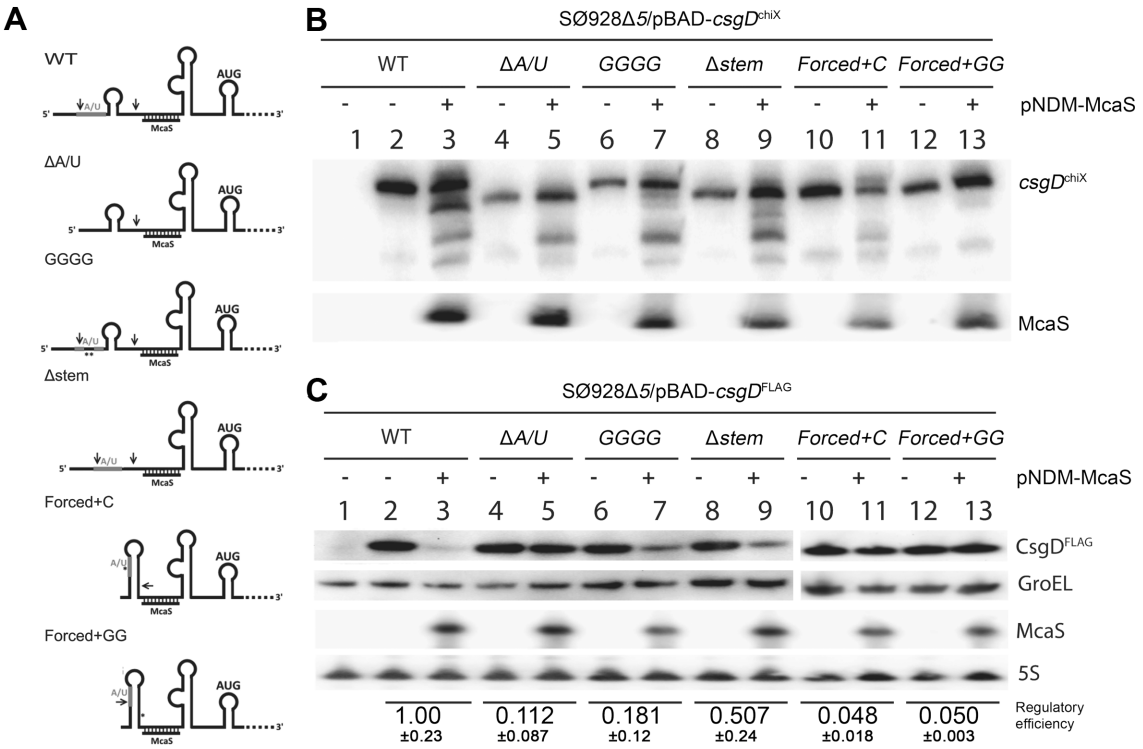

Figure 5. The A/U-rich region and secondary structure is important for efficient mRNA cleavage and translational inhibition. (**A**) Graphical representation of wild-type (WT) and mutant *csgD* mRNAs, shown with consensus secondary structure (Figure 2) indicating the conserved A/U-rich region (grey line), nucleotide substitutions (asterisk) and cleavage site ①and ④from Figure 3 (arrows). (**B**) Northern blot with RNA purified from SØ928Δ5 carrying either wild-type or mutant pBAD-csgD^{chiX} derivatives (this plasmid contains the *csgD* 5′UTR and the first 87 nucleotides of the coding region and the small RNA ChiX for stabilisation of the construct) and either pNDM-mcaS or the empty vector pNDM220. Cultures were grown in LB media at 37°C to an OD_{600} of 0.6 at which point a sample was taken from the culture with wild type *csgD* and empty pNDM220 vector as a negative control. The cultures were induced with 1 mM IPTG for 10 min, followed by 5 min induction with 0.2% arabinose at which point samples were taken from all cultures. Northern blot was performed as described in methods. (**C**) Northern and Western blot with RNA and proteins purified from SØ928Δ5 carrying either wild type or mutant pBAD-csgD^{FLAG} and either pNDM-mcaS or the empty vector pNDM220. Cultures were grown and samples were taken as for panel B and Western and northern blots were performed as described in methods. The CsgD-FLAG levels were quantified from three experiments and normalized to the loading control (GroEL). The efficiency of McaS-mediated repression was determined by dividing the signal from samples without McaS by the signal from samples with McaS. Efficiency of McaS on wild type *csgD* was defined as 1.

red plates under McaS-inducing conditions. In all strains McaS stimulated down-regulation of curli production (Supplementary Figure S2). The *ybeY* mutant is characterised by a slow growth phenotype and poor red colouring on Congo red agar plates. We therefore validated the plate assay with western blot analysis (Supplementary Figure S3).

RNase E is essential in *E. coli,* and therefore a primer extension assay was set up using strain N3431, which carries the *rne-3071* allele, a temperature-sensitive mutation of RNase E, and its isogenic wild type strain N3433 (41). Since pNDM220 exhibits greatly elevated copy numbers due to run-away replication at 42°C (42) (the non-permissive temperature for RNase E in N3431), the arabinose inducible pBAD33 vector derivative was used for sRNA induction, while the chromosomal *csgD* transcript level was assayed. Natural levels of *csgD* proved insufficient for primer extension (data not shown) and therefore a hyperactive OmpR mutant with a point mutation, OmpR234 (43), was introduced to N3431 and N3433 to increase transcription of *csgD* from the chromosome (Supplementary Figure S3). We also deleted RNAse G, encoded by *rng,* in this strain background for several reasons: *First*, RNase G is a paralog of RNase E with a close catalytic domain resemblance; *Second*, RNase E is not essential in a strain overproducing RNase G indicating that they have redundant targets; *Third*, McaS co-purifies with RNase G in a recent immunoprecipitation study (40,44,45). Finally, we introduced the triple FLAG epitope tag sequence to the 3′end of chromosomally encoded *csgD* to monitor protein levels as well as transcript abundance.

The strains were grown to late-log phase in the presence of McaS inducer (0.2% arabinose) to keep *csgD* mRNA levels to a minimum. The cultures were shifted to the non-permissive temperature of RNase E for 20 min and the samples were assayed for CsgD protein and transcript accumulation (Figure 8). A WT strain and the corresponding *rne-3071, rng,* and the double mutants harbouring the empty pBAD33 control vector all showed a clear 5′end *csgD* product and a corresponding clear CsgD^{FLAG} protein band when shifted to 42°C (Figure 8, lanes 1–4). However, *csgD* degradation products disappeared in the *rne* mutant allele (Figure 8, lanes 2 and 4). These degradation products presumably stem from the chromosomally encoded sRNAs McaS, RprA and GcvB (Figure 3). The co-expression of McaS from P_{BAD} showed significant downregulation of *csgD* 5′end mRNA and a dramatic decrease in protein levels

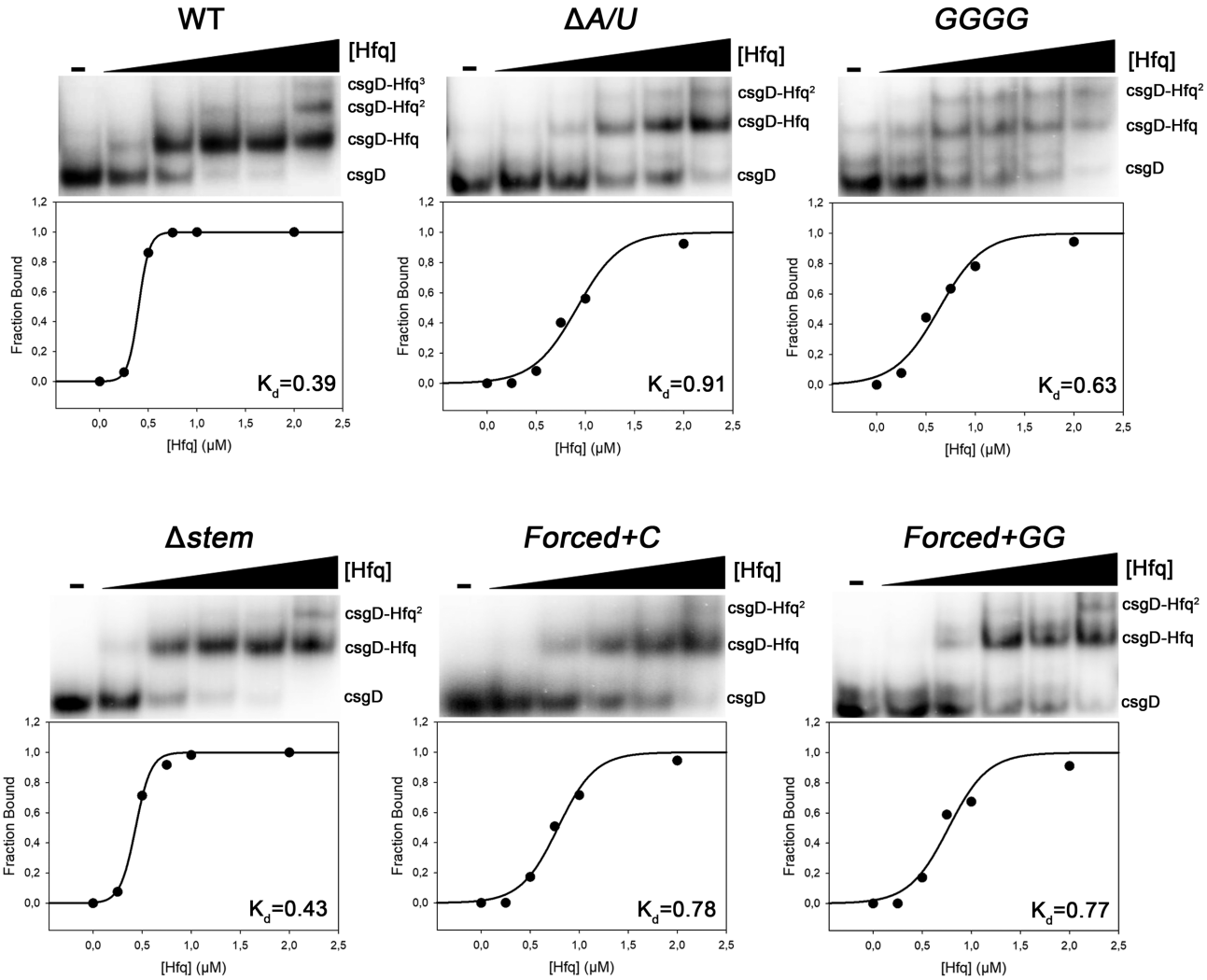

Figure 6. Hfq binds the *csgD* 5′UTR constructs with different affinities assisted by the adjacent conserved stem–loop. Electrophoretic mobility shift assays of Hfq binding to *csgD* mRNA. Samples containing 5′ end-labeled transcripts of 2 nM *csgD* wild type or *csgD* mutant mRNAs (Figure 4A) were incubated with increasing amounts of Hfq. Final monomeric concentrations of Hfq were 0, 0.25, 0.5, 0.75, 1 and 2 μM from left to right.

(Figure 8, lanes 5–8). Interestingly, $CsgD^{FLAG}$ protein levels accumulated only in strains lacking RNase E (Figure 8, lanes 6 and 8). Consistently, these strains also showed elevated 5′end transcript levels. Taken together we conclude that McaS dependent inhibition of CsgD synthesis depend on RNase E. Furthermore, RNase E cuts at A/U rich single stranded regions and we observe that the degradation products are located in the single stranded A/U rich region of *csgD* 5′UTR. Finally, we note that RNase G does not contribute to McaS mediated degradation of *csgD* (Figure 8, lanes 7–8 and Supplementary Figure S2).

DISCUSSION

The switch from a motile single cell to a sessile lifestyle in a biofilm is a highly regulated process and a prominent adaptation programme. This lifestyle decision must be strictly coordinated, because once the decision to form a biofilm is made, the individual cells cannot easily reverse to single planktonic cells again. It is therefore not a surprise that the master regulator of curli biosynthesis, CsgD, is controlled at many levels. Regulation takes place at both the transcriptional and post-transcriptional level by a cohort of regulatory proteins and small non-coding RNAs. The *csgD* transcript has an unusually long 5′ UTR, which is largely conserved. No less than seven sRNAs bind the *csgD* 5′UTR to regulate the expression of curli biosynthesis. Each sRNA relays a signal to the genetic network thereby cross-linking individual genetic regulons and adaptation strategies. Consistent with this concept, many of the sRNAs show largely non-redundant expression patterns (46).

We sought to dissect the mechanism by which several sRNAs downregulate the expression of CsgD. Our results demonstrate that at least two different mechanisms for CsgD repression are used by the sRNAs; *one* prevents ribosomes from translating CsgD, either through direct binding to the RBS or via structural rearrangement of a preceding stem–loop (21,25), and the *second* mechanism induces endoribonucleolytic cleavage of the *csgD* mRNA in a conserved A/U-rich region located in the 5′ UTR.

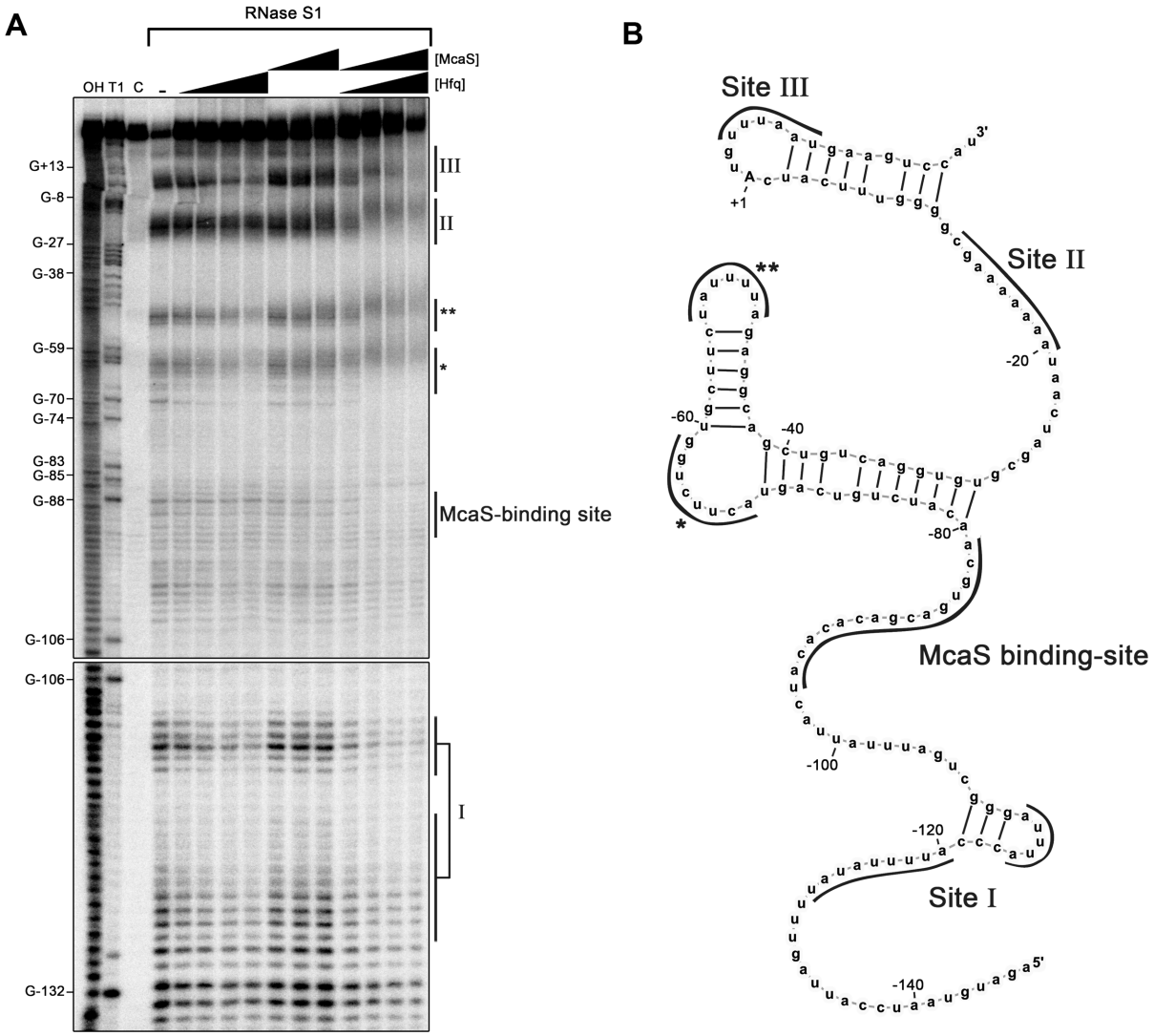

Figure 7. Hfq interacts with the *csgD* 5′UTR at multiple positions. (**A**) Structural probing assays of *csgD* 5′ UTR RNA using Nuclease S1 with increasing concentrations of Hfq, McaS or both. Samples containing 5′end-labeled transcripts of 2 nM *csgD* 5′UTR RNA were pre-incubated at 37°C with Hfq, McaS or both for 10 min before incubation with Nuclease S1 for 5 min. An alkaline hydrolysis ladder ('OH') and a G-ladder ('T1', RNase T1 digestion) was used to determine nucleotide position. Untreated *csgD* 5′ UTR RNA was used as a control ('C'). (**B**) Protected Hfq and McaS binding sites are underlined in the *csgD* 5′ UTR structure.

Short transient ectopic expression of OmrA does not result in any visible *csgD* degradation products. According to our data, the OmrA-mediated regulation of curli biosynthesis is therefore not dependent on rapid cleavage of the *csgD* transcript. This is consistent with previous studies, showing that OmrA and OmrB bind the *csgD* mRNA *in vitro* and ectopic expression of the sRNAs regulate CsgD production *in vivo*. The two small RNAs bind upstream the translational binding site (Figure 9A) to inhibit the access of initiating ribosomes (21).

Previously, we showed that GcvB downregulates the expression of *csgD*. Here, we discovered a specific degradation pattern of *csgD* mRNA following short-term GcvB expression (Figure 3A). The degradation pattern resembles that caused by McaS, except for one cleavage product that is characteristic to GcvB. It is presently not known whether the regulatory effect of GcvB is exerted primarily through translation inhibition or transcript degradation. Other studies show that GcvB prevents ribosomal binding on target mRNAs (47). However, based on the almost identical *csgD* mRNA breakdown products caused by McaS and GcvB induction, we speculate that GcvB may affect CsgD expression through a similar mechanism as McaS. However, the precise mechanism behind GcvB-mediated control of curli biosynthesis and the origin of the specific degradation product, remains to be determined (Figure 3A).

In contrast, RprA-mediated repression has been investigated previously. Hengge suggested in 2012 that two RprA binding sites of *csgD* are involved in a partially redundant manner. This was based on the observation that a single deletion mutation in the common sRNA binding site (shared with McaS) in the proximity of the RBS did not relieve RprA-mediated repression of CsgD biosynthe-

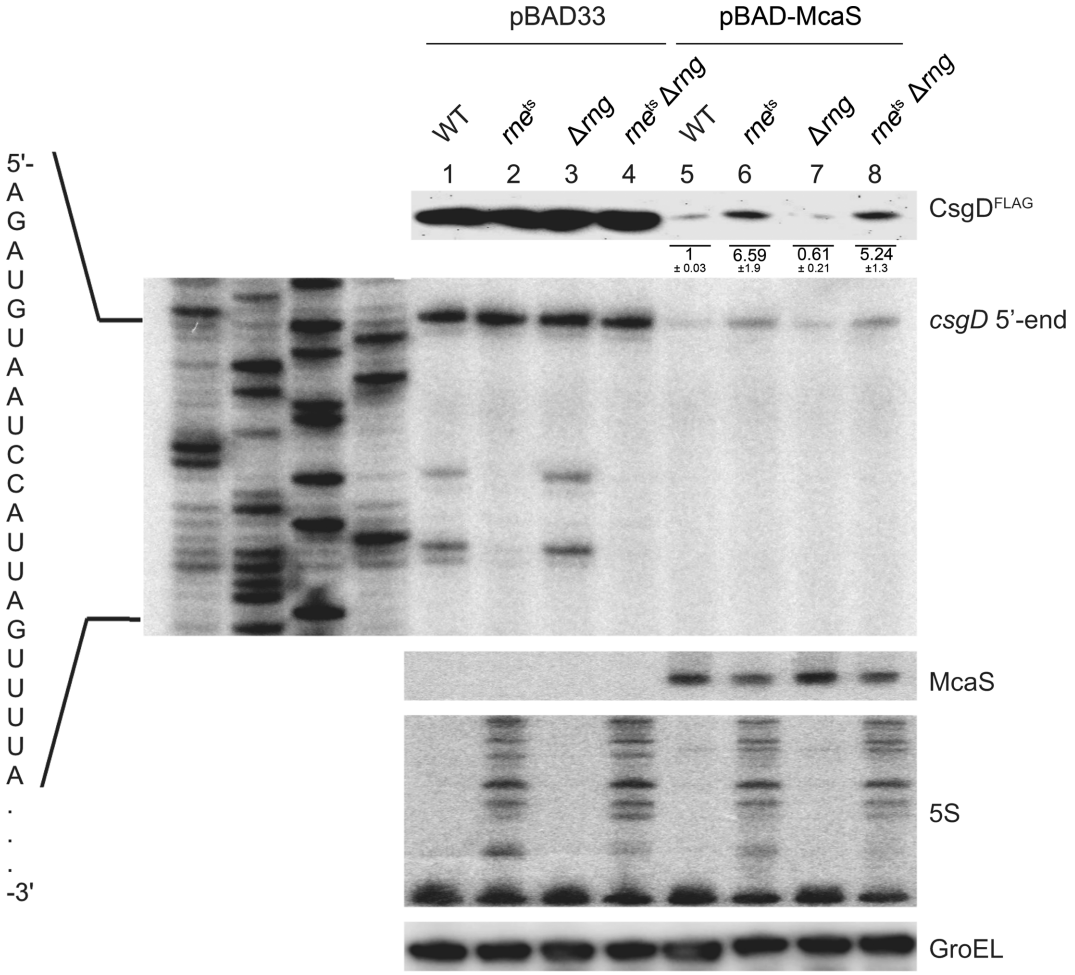

Figure 8. McaS mediated *csgD* mRNA cleavage is RNase E dependent. Western blot of $CsgD^{FLAG}$ and primer extension analysis of *csgD* 5′end. Cultures of N3431 (*rne-3071*), the isogenic N3433 WT strain, containing the OmpR234 point mutation and a single or double knock-outs of *rng,* the gene encoding RNase G. The strains carry either the empty vector control pBAD33 or pBAD-mcaS. All cultures were grown in LB medium at 30°C in the presence of inducer (0.2% arabinose) to an OD_{600} of 1.2. At this point, the cultures were shifted to 42°C for 20 min before sampling for proteins and RNA. GroEL was probed for protein loading control and 5S, showing characteristic unprocessed bands in the *rne-3071* allele, was probed as RNA loading control. Fold-change are of $CsgD^{FLAG}$ protein levels from three experiments.

sis downregulation, while double mutations did (23). However, we have previously shown, in a different experimental setup, that a deletion mutation in the common sRNA binding site completely relieved RprA-mediated repression (22). The study conducted by Hengge used constitutively expressed *csgD* and RprA from low-copy and high-copy number plasmids, respectively, while we used short-term ectopic expression of *csgD* and RprA from medium-copy and low-copy number plasmids, respectively. Furthermore, the deletion mutants used by the two studies were slightly different. Due to the differences in experimental setup in these studies, the exact nature and the primary regulatory mechanism of RprA-controlled repression remain to be resolved. However, given that RprA also induces cleavage in the A/U-rich region of *csgD*, we suspect that the mechanism of RprA-mediated control of *csgD* is similar to that of McaS (see below).

Short-term ectopic expression of McaS leads to degradation of *csgD* mRNA expressed ectopically as a $csgD^{chiX}$ fusion transcript with specific degradation patterns (Figure 5). Primer extension analysis identified sRNA mediated cleavage sites within the 5′ UTR of the *csgD* mRNA (Figure 3). We performed mutational analysis of these sites via nucleotide substitution, deletion, or forced secondary structure formation (Figure 5A). We cannot exclude the possibility that the changes in cleavage patterns in the mutants are at different sites than in the wild-type construct. However, we observe that specific cleavage products disappear in the mutant constructs and other products remain, indicating that the change in cleavage patterns reflect the introduction of specific mutations. Combined we find that the single stranded A/U-rich tract and the adjacent stem loop at the 5′ end of *csgD* mRNA are necessary for McaS-mediated degradation *in vivo*.

Further analysis by structural probing show that the A/U-rich region and the adjacent hairpin in the *csgD* 5′-end most likely serve as a platform for Hfq binding to both sites (Figure 7). Disruption of Hfq binding site I by deletion of the stem–loop structure, reduced McaS dependent control of CsgD synthesis *in vivo* (Figure 5) but did not change the

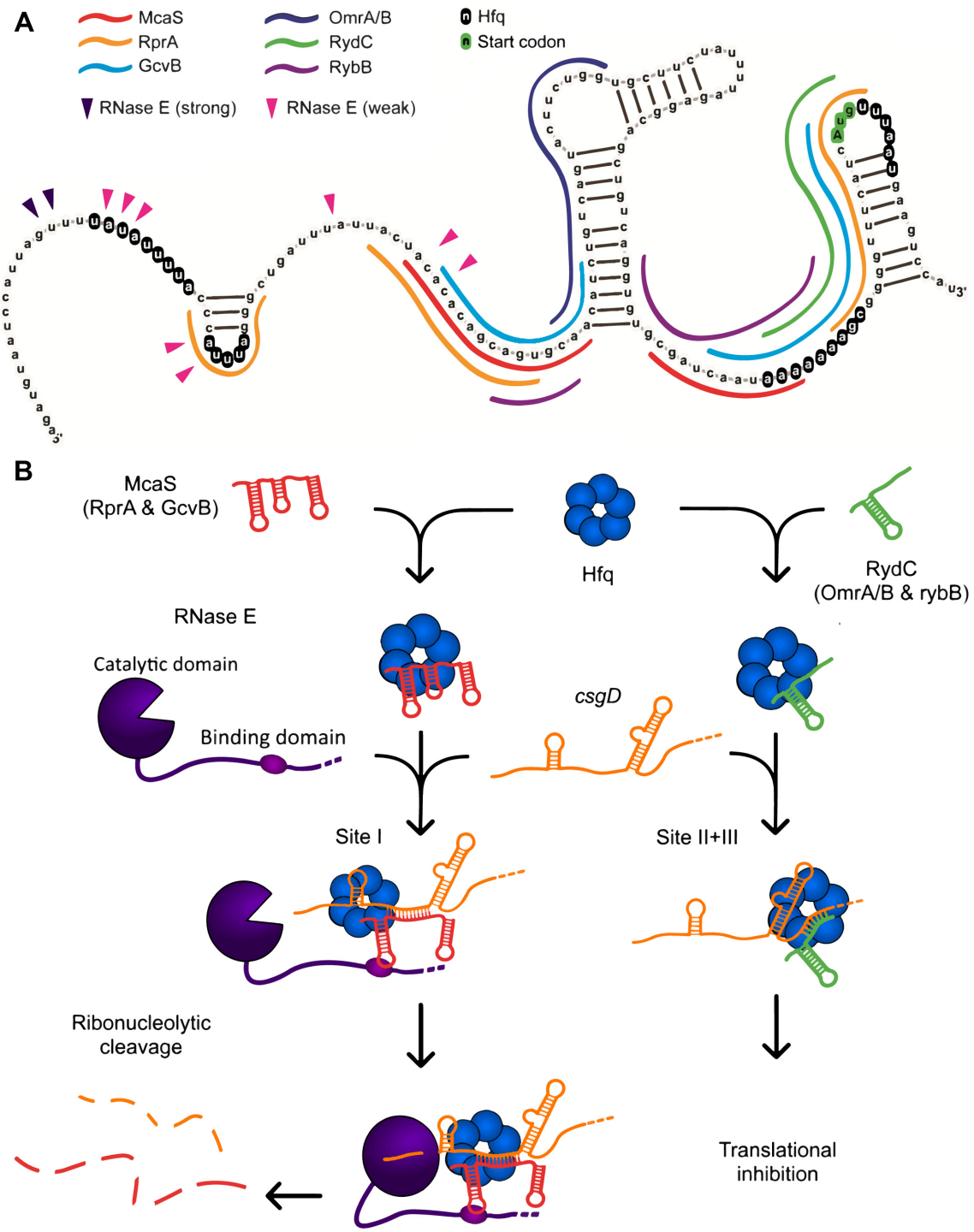

Figure 9. Model for sRNA-mediated regulation of *csgD*. (**A**) Binding/cleavage sites of the sRNAs, Hfq and RNase E on the *csgD* mRNA. McaS, RprA, GcvB and OmRA/B all have extensive base pairing located distally from the ribosomal binding site. While OmrA and OmrB cause translational inhibition, McaS, RprA and GcvB induce ribonucleolytic cleavage of the mRNA. The most significant cleavage site is located near the 5′-end, just upstream of an Hfq binding site. Mutations of this Hfq site inhibits RNase E cleavage and alleviates McaS induced repression of CsgD expression. RydC and RybB bind *csgD* at the ribosomal binding site to inhibit the formation of the translational initiation complex. (**B**) Model for the two sRNA mediated mechanisms for posttranscriptional regulation of *csgD*. Left) McaS (RprA and GcvB) is first bound and stabilized by Hfq. The sRNA-Hfq binary complex interacts with RNase E of the degradosome and its substrate mRNA through base-pairing between the RNAs as well as protein-RNA and protein-protein interactions, in which Hfq binds to Site I of *csgD*. Complex formation leads to ribonucleolytic cleavage of *csgD* and McaS. Right) RydC (omrA/B and rybB) is bound by Hfq. This sRNA-Hfq complex will bind near the ribosomal binding site by base-pairing between the two RNAs as well as Hfq interactions with *csgD* at site II and site III. Formation of this complex leads to translational inhibition.

dissociation constant (Kd) between Hfq and *csgD in vitro* (Figure 6). However, we no longer observe a higher mobility complex corresponding to *csgD*-Hfq^3 in electric mobility shift assays. In general, we do not know the exact identity of the nucleoprotein complexes in our mobility shift assays. Hfq interacts with *csgD* at multiple sites making the interpretation of the results more difficult. It is entirely plausible that the observed mobility shifts are a mixture of Hfq binding at different *csgD* binding sites. However, given that we have only made mutations close to Hfq-binding Site I, suggests that the changes in Kd values are a consequence of Hfq binding to this site.

Further inspection of the structural probing experiment (Figure 7) suggests that McaS and Hfq cooperates at Hfq-binding Site II and Site III, and not just to Site I. We have previously seen that McaS forms extendend base-paring with *csgD* (McaS-binding site 2, Figure 4A) (22). This extendend base-paring brings McaS close to Hfq-binding Site II and the RBS. Thus, the additional protection at these sites probably reflects McaS association to this region *in vitro.* However, this region is not important *in vivo* since deletion of McaS-binding Site 2 has no effect on McaS mediated repression of CsgD protein levels (Figure 4). Whilst McaS inhibits translation initiation *in vitro* (22), our new data does not support this observation but suggests that the extended base-paring at McaS-binding Site 2 is not functional *in vivo*.

Interestingly, the A/U tract is cleaved by RNase E. McaS producing cells have no or only minor CsgD production in an *ompR234* background strain, however, upon shift to the non-permissive temperature of a temperature sensitive mutation of RNase E, *csgD* mRNA accumulates and consequently CsgD protein level increases. The fact that McaS-mediated regulation of CsgD is lost in the absence of RNase E support our view that inhibition of translation initiation has no effect *in vivo* (Figures 4 and 8). Introducing GGGG residues in the A/U rich region, as well as deleting it, decrease cleavage efficiency and thus McaS mediated control of CsgD biosynthesis (Figure 5, Δstem and ΔA/U). Replacing the adjacent stem–loop structure with a stable tetraloop and simultaneous masking the A/U region by base-paring (Figure 5, Forced C and Forced+GG) completely eliminates McaS control of CsgD *in vivo*. Combined, we therefore suggest that the primary regulatory mechanism relies on endoribonucleolytic cleavage in the 5′end of *csgD* mRNA.

We have schematically summarized our findings in Figure 9A. To complete the overview, we have included the known small RNAs regulating curli production via the CsgD transcriptional regulator. From the illustration, it is clear that the *csgD* 5′UTR serves as a hub for post-transcription regulation with the involvement of many small non-coding RNAs and with multiple Hfq binding sites. Several of the small RNAs have multiple base-paring regions along the *csgD* sequence. In this study, we focused on the small RNA McaS and the mechanism by which it controls CsgD expression. The McaS dependent RNase E induced cleavage sites are highlighted by arrows with intense and weak cleavage products, respectively.

In some cases, the sRNA-binding to *csgD* overlaps with our detected Hfq binding sites. A previous study using Spot 42 small RNA as a model and its association with the *entB* mRNA demonstrated that the Hfq-binding and sRNA targeting sites cannot overlap (48). How exactly the sRNAs and Hfq interact with *csgD* at these overlapping sites remain unclear. In addition, it is equally unclear how the different Hfq-binding sites serve as regulatory elements and how they individually contribute to sRNA recognition. We are currently investigating the importance of the individual Hfq-binding sites and how they each aid in regulating CsgD in detail.

It seems likely that the two downstream located Hfq-binding sites, at the TIR are involved in the base-paring of the small RNAs that inhibit translation initiation (e.g. OmrA/B, RydC and RybB) and that the first Hfq-binding site (located at the A/U rich sequence), together with the small RNAs that leads to *csgD* decay (e.g. McaS, RprA and GcvB), is involved in the recruitment of RNase E, to cleave the *csgD* substrate (Figure 9B). For simplicity, we have only added RNase E itself to our illustration, though it is part of the degradosome. The degradosome is a multienzyme complex essential for RNA metabolism and consists of RNase E, the helicase, RhlB, the polynucleotide posphorylase, PNPase, and enolase. The N-terminal half of RNase E is highly conserved and contain the endoribonuclease activity. The C-terminal domain is natively disordered, interacts with RNA and provide a scaffold for degradosome assembly (49). The C-terminal scaffold interacts with Hfq-sRNA binary complexes facilitating mRNA target recognition (50). Hfq is likely to dissociate from the complex at some point before or after the initial degradation process. In such a scenario, RNase E could induce cleavage at the A/U rich sequence after the dissociation of Hfq. Then, the weak overlapping cleavage sites in the A/U rich sequence does not exclude Hfq association to this site as well.

Another explanation for the induction of cleavage at the A/U-rich sequence could be a conformational change within the *csgD* mRNA upon McaS binding. It was previously reported that the binding of MicF deep within the coding region of *lprX* mRNA decreased mRNA stability by inducing a conformational change that induced RNase E mediated cleavage by exposing the mRNA sequence (36). Our data does not support a model where the *csgD* mRNA undergo structural rearrangement upon McaS/Hfq binding (Figure 7 and Supplementary Figure S4). We thus favour a model where McaS, together with Hfq, recruit RNase E directly to the A/U rich sequence to induce cleavage of the *csgD* mRNA.

Interestingly, RNase G does not contribute to *csgD* mRNA decay but RNase G co-purifies with McaS in pull-down experiments (45). McaS is a dual-functioning sRNA. It activates the production of exopolysaccharide β-1,6 N-acetyl-D-glucosamine (PGA) by interacting with the carbon storage regulator CsrA (51). It also serves as a canonical sRNA regulating gene expression in a Hfq dependent manner. In addition to regulating *csgD* expression, McaS also stimulates the synthesis of flagella by activating *flhD* in an Hfq dependent manner (24). However, it seems unlikely that RNase G is involved in these processes, suggesting that McaS has additional uncharacterized targets.

Tight control of *csgD* expression is an important and integral aspect in *E. coli* physiology. The switch from a motile single cell lifestyle to that of a sessile non-motile

lifestyle involves a complex cascade where many environmental conditions signal the rewiring of the regulatory network. At the centre of this network is the CsgD protein. Our study demonstrates that the McaS mediated control of CsgD biosynthesis depend on mRNA degradation. CsgD has an unusual long 5′UTR where McaS (but also RprA and GvcB) binds far upstream the RBS and induces mRNA cleavage at sites located additionally upstream. Thus, the long *csgD* 5′UTR is not just required for stabilisation of transcript but also serves as a conserved highly regulatory element. To begin to understand the significance of long untranslated leaders it is important to understand how other long UTRs are involved in transcript regulation. It would be interesting in the future to study if many other UTRs also serve as hubs for sRNA mediated signal integration. Indeed, it has been suggested that long 5′UTRs are important in the control of many virulence genes in *L. monocytogenes* and it highlights the importance of post-transcriptional control of key regulatory pathways (52).

DATA AVAILABILITY

RNAalifold: http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAalifold.cgi. Clustal Omega: https://www.ebi.ac.uk/Tools/msa/clustalo/.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Johannes Havelund from SDU, Denmark for strains containing knock-outs of sRNAs *omrAB* and *gcvB*.

FUNDING

Lundbeck Foundation [R100-A9549 to M.G.J.]. Funding for open access charge: Publication charges are funded through the Lundbeck Foundation (to M.G.J.); University of Southern Denmark open access policy grants.
Conflict of interest statement. None declared.

REFERENCES

1. Flemming,H.C. and Wingender,J. (2010) The biofilm matrix. *Nat. Rev. Microbiol.*, **8**, 623–633.
2. Hall-Stoodley,L., Costerton,J.W. and Stoodley,P. (2004) Bacterial biofilms: from the natural environment to infectious diseases. *Nat. Rev. Microbiol.*, **2**, 95–108.
3. Espeland,E.M. and Wetzel,R.G. (2001) Complexation, stabilization, and UV photolysis of extracellular and surface-bound glucosidase and alkaline phosphatase: implications for biofilm microbiota. *Microb. Ecol.*, **42**, 572–585.
4. Hoiby,N., Ciofu,O., Johansen,H.K., Song,Z.J., Moser,C., Jensen,P.O., Molin,S., Givskov,M., Tolker-Nielsen,T. and Bjarnsholt,T. (2011) The clinical impact of bacterial biofilms. *Int. J. Oral Sci.*, **3**, 55–65.
5. White,A.P., Gibson,D.L., Collinson,S.K., Banser,P.A. and Kay,W.W. (2003) Extracellular polysaccharides associated with thin aggregative fimbriae of Salmonella enterica serovar enteritidis. *J. Bacteriol.*, **185**, 5398–5407.
6. Kai-Larsen,Y., Luthje,P., Chromek,M., Peters,V., Wang,X., Holm,A., Kadas,L., Hedlund,K.O., Johansson,J., Chapman,M.R. *et al.* (2010) Uropathogenic Escherichia coli modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37. *PLoS Pathog.*, **6**, e1001010.
7. Lloyd,S.J., Ritchie,J.M. and Torres,A.G. (2012) Fimbriation and curliation in Escherichia coli O157:H7: a paradigm of intestinal and environmental colonization. *Gut Microbes*, **3**, 272–276.
8. Hammar,M., Arnqvist,A., Bian,Z., Olsen,A. and Normark,S. (1995) Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in Escherichia coli K-12. *Mol. Microbiol.*, **18**, 661–670.
9. Hammar,M., Bian,Z. and Normark,S. (1996) Nucleator-dependent intercellular assembly of adhesive curli organelles in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 6562–6566.
10. Evans,M.L., Chorell,E., Taylor,J.D., Aden,J., Gotheson,A., Li,F., Koch,M., Sefer,L., Matthews,S.J., Wittung-Stafshede,P. *et al.* (2015) The bacterial curli system possesses a potent and selective inhibitor of amyloid formation. *Mol. Cell*, **57**, 445–455.
11. Romling,U., Rohde,M., Olsen,A., Normark,S. and Reinkoster,J. (2000) AgfD, the checkpoint of multicellular and aggregative behaviour in Salmonella typhimurium regulates at least two independent pathways. *Mol. Microbiol.*, **36**, 10–23.
12. Arnqvist,A., Olsen,A., Pfeifer,J., Russell,D.G. and Normark,S. (1992) The Crl protein activates cryptic genes for curli formation and fibronectin binding in Escherichia coli HB101. *Mol. Microbiol.*, **6**, 2443–2452.
13. Lindenberg,S., Klauck,G., Pesavento,C., Klauck,E. and Hengge,R. (2013) The EAL domain protein YciR acts as a trigger enzyme in a c-di-GMP signalling cascade in E. coli biofilm control. *EMBO J.*, **32**, 2001–2014.
14. Pesavento,C., Becker,G., Sommerfeldt,N., Possling,A., Tschowri,N., Mehlis,A. and Hengge,R. (2008) Inverse regulatory coordination of motility and curli-mediated adhesion in Escherichia coli. *Genes Dev.*, **22**, 2434–2446.
15. Romling,U., Sierralta,W.D., Eriksson,K. and Normark,S. (1998) Multicellular and aggregative behaviour of Salmonella typhimurium strains is controlled by mutations in the agfD promoter. *Mol. Microbiol.*, **28**, 249–264.
16. Ogasawara,H., Hasegawa,A., Kanda,E., Miki,T., Yamamoto,K. and Ishihama,A. (2007) Genomic SELEX search for target promoters under the control of the PhoQP-RstBA signal relay cascade. *J. Bacteriol.*, **189**, 4791–4799.
17. Ogasawara,H., Yamada,K., Kori,A., Yamamoto,K. and Ishihama,A. (2010) Regulation of the Escherichia coli csgD promoter: interplay between five transcription factors. *Microbiology*, **156**, 2470–2483.
18. Barnhart,M.M. and Chapman,M.R. (2006) Curli biogenesis and function. *Annu. Rev. Microbiol.*, **60**, 131–147.
19. Ferrieres,L. and Clarke,D.J. (2003) The RcsC sensor kinase is required for normal biofilm formation in Escherichia coli K-12 and controls the expression of a regulon in response to growth on a solid surface. *Mol. Microbiol.*, **50**, 1665–1682.
20. Vianney,A., Jubelin,G., Renault,S., Dorel,C., Lejeune,P. and Lazzaroni,J.C. (2005) Escherichia coli tol and rcs genes participate in the complex network affecting curli synthesis. *Microbiology*, **151**, 2487–2497.
21. Holmqvist,E., Reimegard,J., Sterk,M., Grantcharova,N., Romling,U. and Wagner,E.G. (2010) Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J.*, **29**, 1840–1850.
22. Jorgensen,M.G., Nielsen,J.S., Boysen,A., Franch,T., Moller-Jensen,J. and Valentin-Hansen,P. (2012) Small regulatory RNAs control the multi-cellular adhesive lifestyle of Escherichia coli. *Mol. Microbiol.*, **84**, 36–50.
23. Mika,F., Busse,S., Possling,A., Berkholz,J., Tschowri,N., Sommerfeldt,N., Pruteanu,M. and Hengge,R. (2012) Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in Escherichia coli. *Mol. Microbiol.*, **84**, 51–65.
24. Thomason,M.K., Fontaine,F., De Lay,N. and Storz,G. (2012) A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in Escherichia coli. *Mol. Microbiol.*, **84**, 17–35.

25. Bordeau,V. and Felden,B. (2014) Curli synthesis and biofilm formation in enteric bacteria are controlled by a dynamic small RNA module made up of a pseudoknot assisted by an RNA chaperone. *Nucleic Acids Res.*, **42**, 4682–4696.
26. Serra,D.O., Mika,F., Richter,A.M. and Hengge,R. (2016) The green tea polyphenol EGCG inhibits E. coli biofilm formation by impairing amyloid curli fibre assembly and down-regulating the biofilm regulator CsgD via the sigma -dependent sRNA RybB. *Mol. Microbiol.*, **101**, 136–151.
27. Storz,G., Vogel,J. and Wassarman,K.M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell*, **43**, 880–891.
28. Guillier,M. and Gottesman,S. (2006) Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. *Mol. Microbiol.*, **59**, 231–247.
29. Moller,T., Franch,T., Hojrup,P., Keene,D.R., Bachinger,H.P., Brennan,R.G. and Valentin-Hansen,P. (2002) Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol. Cell*, **9**, 23–30.
30. Aiba,H. (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.*, **10**, 134–139.
31. Moll,I., Leitsch,D., Steinhauser,T. and Blasi,U. (2003) RNA chaperone activity of the Sm-like Hfq protein. *EMBO Rep.*, **4**, 284–289.
32. Brennan,R.G. and Link,T.M. (2007) Hfq structure, function and ligand binding. *Curr. Opin. Microbiol.*, **10**, 125–133.
33. Schu,D.J., Zhang,A., Gottesman,S. and Storz,G. (2015) Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J.*, **34**, 2557–2573.
34. 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.
35. Mohanty,B.K., Maples,V.F. and Kushner,S.R. (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol. Microbiol.*, **54**, 905–920.
36. Corcoran,C.P., Podkaminski,D., Papenfort,K., Urban,J.H., Hinton,J.C. and Vogel,J. (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol. Microbiol.*, **84**, 428–445.
37. Hui,M.P., Foley,P.L. and Belasco,J.G. (2014) Messenger RNA degradation in bacterial cells. *Annu. Rev. Genet.*, **48**, 537–559.
38. Pandey,S.P., Winkler,J.A., Li,H., Camacho,D.M., Collins,J.J. and Walker,G.C. (2014) Central role for RNase YbeY in Hfq-dependent and Hfq-independent small-RNA regulation in bacteria. *BMC Genomics*, **15**, 121.
39. Meng,W. and Nicholson,A.W. (2008) Heterodimer-based analysis of subunit and domain contributions to double-stranded RNA processing by Escherichia coli RNase III in vitro. *Biochem. J.*, **410**, 39–48.
40. 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.
41. Apirion,D. (1978) Isolation, genetic mapping and some characterization of a mutation in Escherichia coli that affects the processing of ribonuleic acid. *Genetics*, **90**, 659–671.
42. Larsen,J.E., Gerdes,K., Light,J. and Molin,S. (1984) Low-copy-number plasmid-cloning vectors amplifiable by derepression of an inserted foreign promoter. *Gene*, **28**, 45–54.
43. Klunk,W.E., Pettegrew,J.W. and Abraham,D.J. (1989) Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. *J. Histochem. Cytochem.*, **37**, 1273–1281.
44. Lee,K., Bernstein,J.A. and Cohen,S.N. (2002) RNase G complementation of rne null mutation identifies functional interrelationships with RNase E in Escherichia coli. *Mol. Microbiol.*, **43**, 1445–1456.
45. van Nues,R.W., Castro-Roa,D., Yuzenkova,Y. and Zenkin,N. (2016) Ribonucleoprotein particles of bacterial small non-coding RNA IsrA (IS61 or McaS) and its interaction with RNA polymerase core may link transcription to mRNA fate. *Nucleic Acids Res.*, **44**, 2577–2592.
46. Boehm,A. and Vogel,J. (2012) The csgD mRNA as a hub for signal integration via multiple small RNAs. *Mol. Microbiol.*, **84**, 1–5.
47. Sharma,C.M., Darfeuille,F., Plantinga,T.H. and Vogel,J. (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.*, **21**, 2804–2817.
48. Beisel,C.L., Updegrove,T.B., Janson,B.J. and Storz,G. (2012) Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J.*, **31**, 1961–1974.
49. Carpousis,A.J. (2007) The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.*, **61**, 71–87.
50. Bruce,H.A., Du,D., Matak-Vinkovic,D., Bandyra,K.J., Broadhurst,R.W., Martin,E., Sobott,F., Shkumatov,A.V. and Luisi,B.F. (2018) Analysis of the natively unstructured RNA/protein-recognition core in the Escherichia coli RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic Acids Res.*, **46**, 387–402.
51. Jorgensen,M.G., Thomason,M.K., Havelund,J., Valentin-Hansen,P. and Storz,G. (2013) Dual function of the McaS small RNA in controlling biofilm formation. *Genes Dev.*, **27**, 1132–1145.
52. Loh,E., Gripenland,J. and Johansson,J. (2006) Control of Listeria monocytogenes virulence by 5′-untranslated RNA. *Trends Microbiol.*, **14**, 294–298.