ACS SyntheticBiology

RESEARCH ARTICLE

pubs.acs.org/synthbio

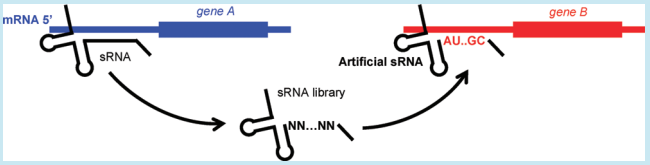
Engineering Artificial Small RNAs for Conditional Gene Silencing in *Escherichia coli*

Vandana Sharma, Asami Yamamura, and Yohei Yokobayashi*

Department of Biomedical Engineering, University of California, Davis, 451 Health Sciences Drive, Davis, California 95616, United States

Supporting Information

ABSTRACT: It has become increasingly evident that noncoding small RNAs (sRNAs) play a significant and global role in bacterial gene regulation. A majority of the *trans*-acting sRNAs in bacteria interact with the 5′ untranslated region (UTR) and/or the translation initiation region of the targeted mRNAs via imperfect base pairing, resulting in reduced translation efficiency and/or mRNA stability. Additionally, bacterial sRNAs often contain distinct scaffolds that recruit RNA chaperones such as Hfq to facilitate gene regulation. In this study, we describe a strategy to engineer artificial sRNAs that can regulate desired endogenous genes in *Escherichia coli*. Using a fluorescent reporter gene that was translationally fused to a native 5′ mRNA leader sequence, active artificial sRNAs were screened from libraries in which natural sRNA scaffolds were fused to a randomized antisense domain. Artificial sRNAs that posttranscriptionally repress two endogenous genes *ompF* and *fliC* were isolated and characterized. We anticipate that the artificial sRNAs will be useful for dynamic control and fine-tuning of endogenous gene expression in bacteria for applications in synthetic biology.



KEYWORDS: noncoding RNA, small RNA, gene regulation, gene knockdown, RNA engineering, translational regulation

Because of its predictable and programmable nature of gene silencing based on the target gene sequence, RNA interference (RNAi) has emerged as one of the most powerful tools to probe and manipulate gene expression in mammalian cells.[1,2] Unfortunately, analogous technology for use in bacteria that is as reliable and efficient as mammalian RNAi has remained elusive. First discovered in 1984,[3] noncoding small RNAs (sRNAs) that posttranscriptionally modulate gene expression via imperfect Watson−Crick base pairing with the 5′ untranslated region (UTR) and/or translation initiation region of bacterial mRNAs have attracted great interest as a major class of global bacterial gene regulators.[4] Recent investigations on the natural sRNAs have revealed the modular architecture of sRNAs. For example, it has been shown that the ∼60 nt 5′ domain of *Escherichia coli* Spot42 engages with the *gal*K translation initiation region via imperfect base pairing, while the ∼50 nt 3′ domain is responsible for Hfq recruitment and transcription termination.[5,6] Interestingly, it has been shown that an antisense sequence from one sRNA can be fused to a Hfq-binding motif of another sRNA to yield a functional hybrid sRNA.[7,8]

Inspired by the modular architecture of bacterial sRNAs, some researchers have sought to mimic the bacterial riboregulators to control noncognate genes. Coleman *et al.* constructed hybrid MicF sRNAs preserving the two stem loops at both 5′ and 3′ termini while inserting an artificial antisense sequence in between.[9] The antisense sequences they used, however, were long (∼50 to 250 nt) and perfectly complementary to various parts of the targeted mRNA, which are uncharacteristic of the antisense domains of the natural bacterial sRNAs.[4] Other earlier efforts have addressed the use of long antisense RNAs containing perfectly complementary sequences to various positions within the targeted mRNAs in *E. coli*.[10−14] However, the efficacies of these antisense RNAs have been variable, and no robust design guidelines have emerged. Moreover, specificity of these antisense RNAs has not been thoroughly investigated.

More recently, Man *et al.* reported a semirational strategy to design artificial sRNAs.[15] They generated a pool of synthetic sRNA sequences *in silico* by randomly fusing an antisense sequence, a Hfq-binding sequence, and a transcription terminator and computationally screened them according to qualitative structural criteria. Candidate sRNAs were individually synthesized and evaluated in *E. coli*. Although there should be hundreds of potential sRNA designs that meet their criteria, the need for individual synthesis and evaluation severely limited their throughput. Out of the 16 initially designed sRNAs targeting two genes, only two clones repressed the target gene expression by 70% or greater.

Although predictable and rational design of artificial sRNAs is a desirable long-term goal, our current incomplete understanding of the design principles of bacterial sRNAs still limits our ability to rationally design highly active artificial sRNAs. Numerous factors that are likely to influence sRNA efficacy, such as target mRNA structures, reaction kinetics, and off-target effects, have not been incorporated in the design criteria of the previous efforts. Unlike the mammalian microRNAs that share common and intuitive sequence characteristics,[16] bacterial sRNAs are highly diverse in their sequences and in their interactions with

Received: July 15, 2011
Published: August 25, 2011

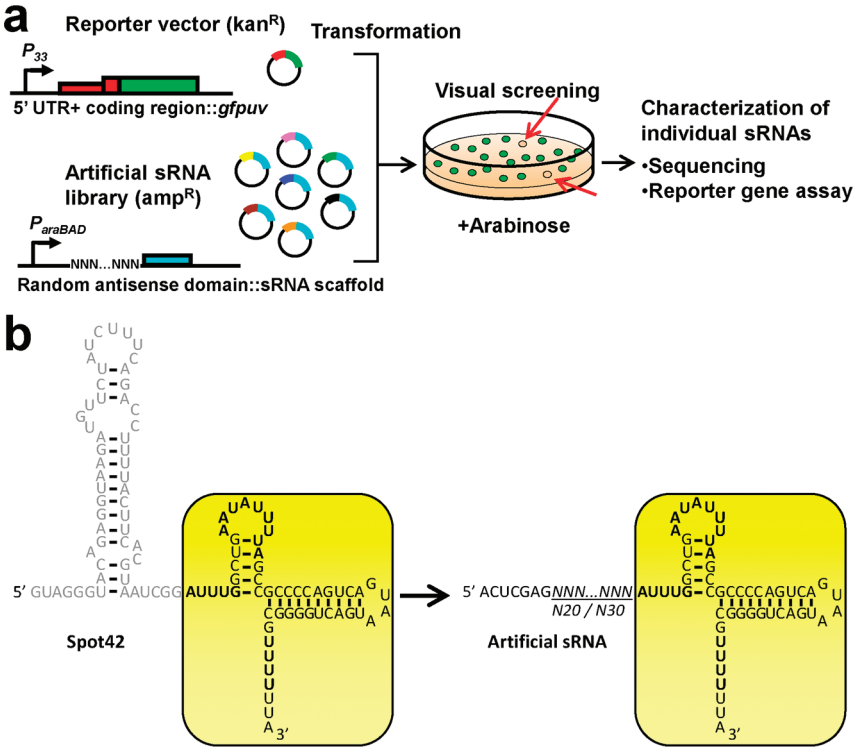


Figure 1. Artificial sRNA screening strategy and library design. (a) Schematic illustration of the artificial sRNA screening strategy. A reporter vector with the target mRNA leader sequence fused to *gfpuv* is cotransformed with a partially randomized artificial sRNA expression library and plated on agar plates. Colonies with weaker fluorescence are picked and characterized. (b) Artificial sRNA library based on the Spot42 sRNA scaffold (yellow box). The antisense domain in Spot42 (identified for *galK*) is shown in gray, and the bases that were shown to interact with Hfq are indicated in bold.[5] Degenerate bases (N) were inserted between the vector-derived sequence (5′-ACUCGAG-3′) and the sRNA scaffold.

the targets. For example, the number of critical bases for antisense function can be as few as six bases[17] to as extensive as several dozen bases with internal mismatches.[6,18] Moreover, sRNAs are known to interact with cognate mRNAs at an untranslated region (5′ or intergenic), coding region, or both.[4] Therefore, it seemed prudent not to restrict ourselves to a rational design rule that may later reveal unforeseen limitations.

To address the challenges in engineering artificial sRNAs with limited understanding of the design principles of bacterial sRNAs, we developed a high-throughput screening strategy that can efficiently identify synthetic sRNAs capable of regulating endogenous genes in *trans*. To demonstrate our strategy, we searched for and isolated artificial sRNAs that repress two endogenous genes in *E. coli*. Artificial sRNAs targeting *ompF* were found to exhibit characteristics similar to those of the natural sRNA MicF that is known to regulate the same gene.[3,19] Additionally, artificial sRNAs that repress flagellin (encoded by *fliC*),[20] a key structural component of the flagella,[21] were engineered and found to induce an expected phenotype (reduced cell motility).[22]

RESULTS AND DISCUSSION

Our experimental strategy is illustrated in Figure 1a. A library of artificial sRNAs was constructed by fusing a randomized antisense domain to scaffolds from natural sRNAs that are known to interact with RNA chaperones. The sRNA plasmid library was cotransformed with a reporter vector in which GFPuv was translationally fused to the 5′ leader sequence of the targeted mRNA. *E. coli* colonies harboring promising artificial sRNAs were visually screened by reduced fluorescence and further characterized individually.

Design of Artificial sRNA Libraries. We arbitrarily selected the four previously characterized *E. coli* sRNAs DsrA,[23−25] GcvB,[26,27] MicF,[18] and Spot42[5,6] to be used as scaffolds to which 20 and 30 degenerate bases were fused as putative antisense domains. These scaffolds are known to interact with RNA chaperones such as Hfq and are predicted to be important for sRNA mediated gene regulation.[5,28] The library based on the Spot42 scaffold[6] is depicted in Figure 1b, and those based on the other scaffolds are shown in Supplementary Figure S1. Plasmid libraries based on the four scaffolds were prepared individually and mixed in equal amounts. The combined library contained $>1 \times 10^7$ unique artificial sRNA clones, which is greater than the typical number of sRNA clones screened in this study ($\sim 10^5$). The artificial sRNAs were cloned downstream of an arabinose-inducible promoter (P_{araBAD}) on a plasmid with a low-copy origin of replication (pMB1 ori derived from pBR322) and an ampicillin-selectable marker (pBadXH) (Supplementary Figure S6a).

Screening of Artificial sRNAs Targeting *ompF* 5′ UTR. The artificial sRNA library was first used to screen for activity against the 5′ leader sequence of the mRNA encoding the outer membrane porin OmpF. *E. coli* OmpF is known to be translationally regulated by the sRNA MicF, which is the first *trans*-acting bacterial sRNA discovered by Inouye and colleagues.[3] Existence of the natural sRNA made the *ompF* 5′ UTR an attractive first target of artificial sRNAs for validating our screening strategy. The reporter plasmid pKP33-OmpF::GFPuv was

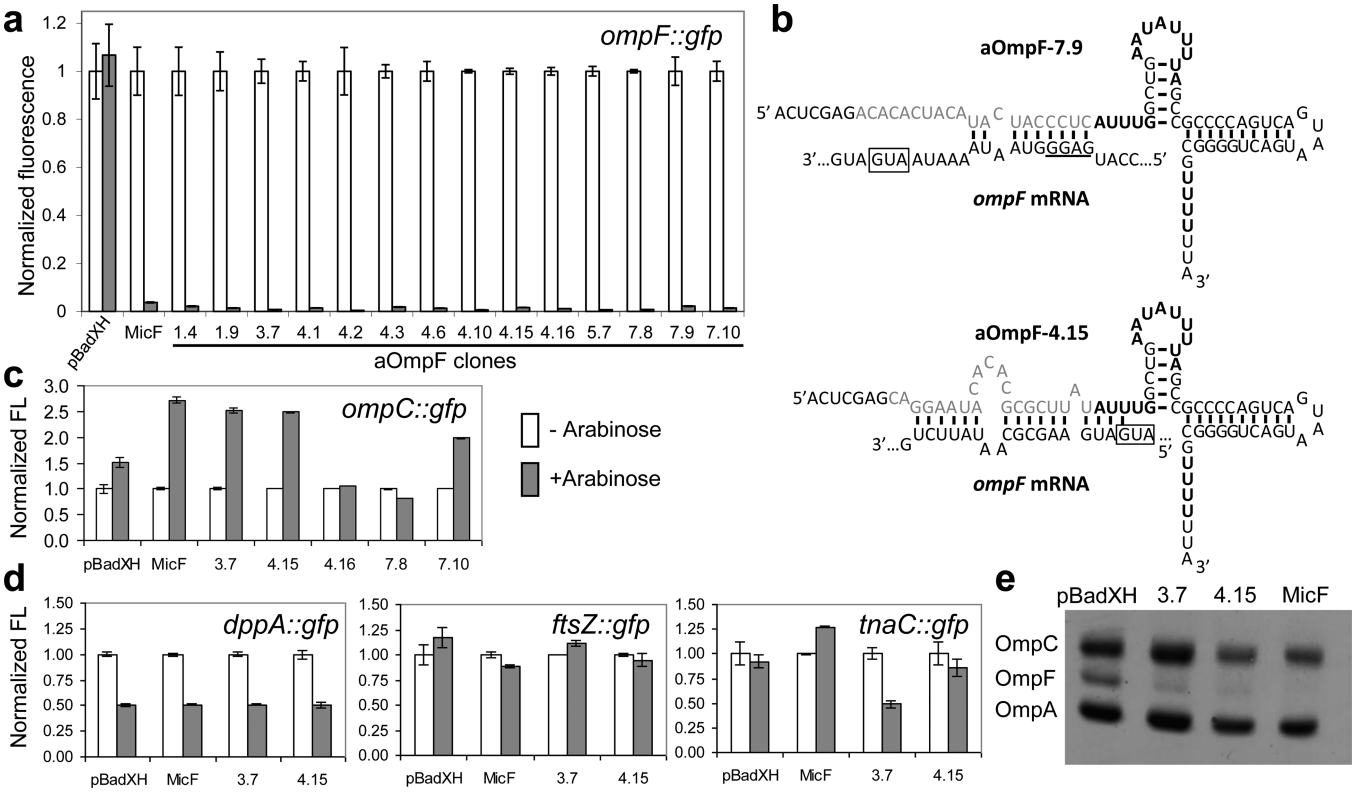


Figure 2. Artificial sRNAs targeting *ompF*. (a) Reporter gene assay (*ompF::gfp*) of the anti-*ompF* artificial sRNAs. Normalized cellular fluorescence in the absence (white bars) and presence (dark bars) of arabinose. (b) Putative models of the interactions between the artificial sRNAs and *ompF* mRNA predicted by IntaRNA. aOmpF-7.9 is representative of the artificial sRNAs that contain the 5′-CCCUC-3′ motif that is likely to engage with the SD sequence (underlined). aOmpF-4.15 represents the artificial sRNAs that are likely to target the *ompF* coding region. The AUG start codon is boxed, and the sequences corresponding to the randomized region are shown in gray. (c) Effects of the sRNAs on *ompC::gfp* reporter gene expression. (d) Effects of the selected sRNAs on additional mRNA leader sequences. (e) Expression levels of the outer membrane porins (OmpA, OmpC, OmpF) in TOP10 cells expressing the sRNAs (pBadXH: empty vector used as a negative control). (a, c, d) The data are averages of triplicate overnight cultures, and the error bars indicate SD. Cellular fluorescence in the absence of arabinose was normalized to 1.0.

constructed by cloning the 110 bases containing the 5′ UTR and the first 12 codons from the coding region of *ompF* in frame with *gfpuv* (*ompF::gfpuv*) (Supplementary Figure S7).[18] The *ompF::gfpuv* fusion mRNA was transcribed from a synthetic constitutive promoter P_{33} (Supplementary Figure S6b and Methods). The reporter mRNAs in this study contained a 21 nt leader sequence corresponding to the *lac* operator and a restriction site (*Eco*RI) derived from the cloning vector at the 5′ end of the native sequence, unless noted otherwise. The reporter plasmids were constructed on a plasmid backbone with a low-copy origin of replication (p15a ori) and a kanamycin-selectable marker and, therefore, were compatible with the artificial sRNA expression plasmids (Supplementary Figure S6b). Using this system, we investigated if it is possible to isolate artificial MicF-like sRNAs based on the noncognate scaffolds.

The artificial sRNA library consisting of the scaffolds based on the sRNAs mentioned above with the exception of MicF was transformed into TOP10 *E. coli* cells harboring pKP33-OmpF::GFPuv and plated on LB agar plates containing 0.1% arabinose. Of the approximately 3×10^5 colonies screened, ~90 clones (0.03%) were found to exhibit reduced fluorescence when visually examined over a UV transilluminator (365 nm). Fourteen clones were further characterized after isolating the artificial sRNA plasmids and retransformed into the fresh host strain to rule out mutations in the host or the reporter plasmid. As shown in Figure 2a, the artificial sRNAs exhibited remarkable downregulation of the *ompF::gfpuv* reporter gene expression in overnight cultures ranging from 45- to 145-fold repression, which is comparable to or even exceeds that of MicF (27-fold repression) under the same conditions. Qualitatively similar results were observed in log-phase and plate cultures (Supplementary Figure S2). Overall, the dual plasmid-based visual colony screening of artificial sRNAs proved to be reasonably efficient with no significant false positives.

Characterization of Artificial sRNAs Targeting *ompF* 5′ UTR. In addition to the 14 clones shown in Figure 2a, 12 artificial sRNA clones (a total of 26 clones) that were confirmed to repress *ompF::gfpuv* reporter gene expression were sequenced, which identified 23 unique anti-*ompF* sRNA clones (Table 1). It was found that 19 sRNA clones contained the Spot42 scaffold, while the remaining four clones harbored the DsrA scaffold. The reason for the observed bias in the sRNA scaffold is not clear, but it is possible that some scaffolds are intrinsically more tolerant to the fused antisense sequences compared to others. Another notable observation was that 11 of the 23 artificial sRNAs contained a consensus sequence 5′-CCCUC-3′, which is complementary to the region that includes the Shine-Dalgarno (SD) sequence of *ompF* mRNA (5′-GAGGG-3′). As expected, at least some of these artificial sRNAs were predicted to hybridize with the sequence around the SD region of *ompF* 5′ UTR when

Table 1. Sequences of Artificial anti-*ompF* sRNAs

aOmpF sRNA	scaffold	randomized region[a,b]
1.4	DsrA	GACUAACACAUCCCUCUAUGGCCGCUGCGC
1.9	DsrA	CCUGCUGCCG
2.2	Spot42	CAAAACCAUACCCUC
2.13	Spot42	GGAAUAAACUGCCAGGCUACCCCCCUUGUG
3.7[c]	DsrA	CCAAGACAUAUUGUG
4.1	Spot42	UAUAUAUAAGUUCCCUCGUG
4.2	Spot42	UGAAAAAACCCUCUCAAAA
4.3	Spot42	UACCCCUCAUGAGACAUCC
4.5	Spot42	AGGUAAAAUGCGCUUCAGU
4.6	Spot42	UCACGAAAUUCUUCCCUCAU
4.10	DsrA	UCAAAAAACACCCUCAC
4.12	Spot42	CAUUUUCAGCUCACCAUACCCUCCAUGCGC
4.15	Spot42	CAGGAAUACACACGCGCUUAU
4.16	Spot42	UGCGAUC
5.4	Spot42	GACUCGGUUACAUUGCGC
5.6	Spot42	CAAAGUUAACAUAGACCCUC
6.7	Spot42	GUUUAUUGCGUCAUAAACAU
7.2	Spot42	UAAUUAUCUGCCAUUAGAUG
7.3	Spot42	CCCUCAUGUC
7.7	Spot42	AUCACAUGACAGAGCGCAUC
7.8	Spot42	GUUAAAAGUGGACGACACUA
7.9	Spot42	ACACACUACAUACUACCCUC
7.10	Spot42	GUUACAUUGCGC

[a] Sequence corresponding to the degenerate bases depicted in Figure 1b and Supplementary Figure S1. Note that all sRNAs start with 5′-ACUCGAG derived from the vector (not shown above) with the exception of aOmpF-1.9 (5′-ACUCGA-) and aOmpF-5.4 and aOmpF-7.10 (5′-ACUCG-). [b] Conserved motif (5′-CCCUC-3′) is underlined. [c] A partial deletion in DsrA scaffold was discovered. The retained scaffold sequence is 5′ CUUCU UGCUU AACGA AGUUU CAUCC CGACC CCCUC AGGGU CGGGA UUU 3′.

analyzed by IntaRNA[29] (Figure 2b). On the other hand, most (10 out of 12) of the remaining artificial sRNAs were predicted to interact mainly with the coding region of the *ompF* mRNA (Figure 2b). Interestingly, MicF has been reported to engage in a more extensive hybridization with the *ompF* mRNA spanning −16 to +10 bases relative to the translation initiation base.[18]

It is known that *E. coli* intricately regulates OmpF and another major outer membrane porin OmpC in response to various environmental stimuli such as changes in temperature and osmotic pressure.[30] OmpC is also regulated by its specific sRNA MicC.[31] Therefore, we investigated if the artificial sRNAs screened against the *ompF* leader sequence exhibit target specificity similar to that of MicF by examining their effects on the expression of the *ompC::gfpuv* reporter. We tested eight artificial sRNA clones that contain the 5′-CCCUC-3′ motif for their activity against *ompC::gfpuv*. Surprisingly, five of these artificial sRNAs showed significant downregulation of *ompC::gfpuv* (4- to 20-fold), while three showed modest (∼1.5-fold) repression (Supplementary Figure S3). This cross-reactivity may be explained by the presence of two 5′-GAGGG-3′ motifs located in the 5′ UTR of *ompC* mRNA, one of which overlaps with the SD sequence. In contrast, other artificial sRNAs that lack the 5′-CCCUC-3′ motif were found to have only moderate effects on *ompC::gfpuv* expression (Figure 2c). Moderate upregulation of *ompC::gfpuv* was observed with MicF and the empty vector (2.7- and 1.5-fold, respectively), indicating that this level of change in expression may be due to nonspecific factors or indirect effects (Figure 2c). Anti-*ompF* artificial sRNAs aOmpF-3.7 and aOmpF-4.15 were further tested against mRNA leader sequences from three additional arbitrarily selected genes from the *E. coli* genome (Figure 2d). No significant activities against these leader sequences were observed, suggesting that these artificial sRNAs are highly specific.

Finally, we examined if the artificial sRNAs can repress endogenous OmpF expression. *E. coli* cells transformed with the appropriate sRNA expression plasmid (pBAD-aOmpF-3.7, pBAD-aOmpF-4.15, or pBAD-MicF) or the empty vector (pBadXH) (Supplementary Figure S6a) were grown in LB medium supplemented with 0.1% arabinose to mid-log phase (OD_{600} ∼0.6) at 24 °C to induce OmpF expression.[32] As shown in Figure 2e, OmpF expression was clearly reduced in the cells expressing the artificial sRNAs compared to the control (empty vector) without affecting OmpA or OmpC. Importantly, these results demonstrate that artificial sRNAs discovered by our screening strategy can specifically knockdown endogenous genes much like the natural sRNAs.

Screening of Artificial sRNAs Targeting *fliC* 5′ UTR. The successful discovery of functional artificial sRNAs that target an endogenous gene (*ompF*) with efficiency and specificity similar to those of a natural sRNA (MicF) prompted us to ask if it is possible to design artificial sRNAs that target other genes for which there are no known natural sRNA regulators. Consequently, we focused on flagellin (FliC), which constitutes the filament structure of the bacterial flagella, as our next target. We anticipated that *E. coli* cells whose FliC expression was downregulated by artificial sRNAs would display an easily detectable phenotype (reduced cell motility) as previously demonstrated by *fliC* deletion mutants.[22] Moreover, to our knowledge, no sRNAs are known to target *fliC* in any bacteria.

The reporter plasmid pKP33-FliC::GFPuv was constructed by cloning the 5′ UTR (70 nt) and the first 13 codons of *fliC* mRNA in frame with *gfpuv* (*fliC::gfpuv*) as described above for *ompF* (Supplementary Figure S7). The reporter plasmid was cotransformed with the artificial sRNA libraries based on all four scaffolds (DsrA, GcvB, MicF, and Spot42), resulting in 1.2×10^5 colonies. We isolated 11 clones that showed reduced fluorescence based on visual screening on agar plates. In contrast to our previous screen against *ompF*, however, the repression efficiency of the artificial sRNAs was moderate, with the best artificial sRNA (aFliC-12) yielding 4-fold repression in overnight cultures (Supplementary Figure S4). Interestingly, sequencing of these artificial sRNAs revealed that they all contained the Spot42 scaffold except one clone that retained the DsrA scaffold (Table 2). The apparent bias for the Spot42 scaffold in both of our screens is intriguing. However, more rigorous investigations are necessary to determine the roles of various sRNA scaffolds (including those that were not used in the present study) in the context of artificial sRNAs.

Because of the relatively moderate activities of the anti-*fliC* sRNAs discovered in our initial screen, we designed a secondary screening step to optimize the sRNA activity. First, we found that the removal of the 5′ leader sequence from the vector-derived *lac* operator sequence in the reporter plasmid resulted in significant increase in the cellular fluorescence (pKP33-ΔlacO-FliC::GFPuv) (Supplementary Figure S7). The elevated baseline fluorescence of the reporter strain allowed significantly improved discrimination of

Table 2. Sequences of Artificial anti-*fliC* sRNAs

library[a]	aFliC sRNA	scaffold	randomized region[b]
1	3	Spot42	AUAAUUAAGACCCUCGGAA
1	4	Spot42	AUAAUAAGACCCUCGGAAU
1	7	Spot42	AGGCAUAAUUUGG
1	9	DsrA	CUAAUGUGCC
1	10	Spot42	UCUACAGUAUGCGACUGUUGGUGGGUG
1	12	Spot42	ACGGUAUAGUUAUCCUAAG
1	13	Spot42	CUGCAAUGGACGUCUAUUAACCCCUCGGUU
1	17	Spot42	AGGUACUUGUUAACCGGGAU
1	21	Spot42	GAAUGAAACACUGUGCCAUCG
1	24	Spot42	AGGAAUGAGCCAUGAUUGUU
1	27	Spot42	AACUCACAUUGAUAACUUUGACUUUCUUG
2	12.1	Spot42	UGUACGUCGUUAUCCUAAG
2	12.3	Spot42	AGACUAUCGUUAUCCUAAG
2	12.16	Spot42	CAUCGUUCGUUAUCCUAAG

[a] aFliC sRNAs isolated from the first (1) or the secondary (2) library as described in the manuscript. [b] Sequence corresponding to the degenerate bases depicted in Figure 1b and Supplementary Figure S1. Note that all sRNAs start with 5′-ACUCGAG derived from the vector with the exception of aFliC-9 in which the entire sequence was deleted.

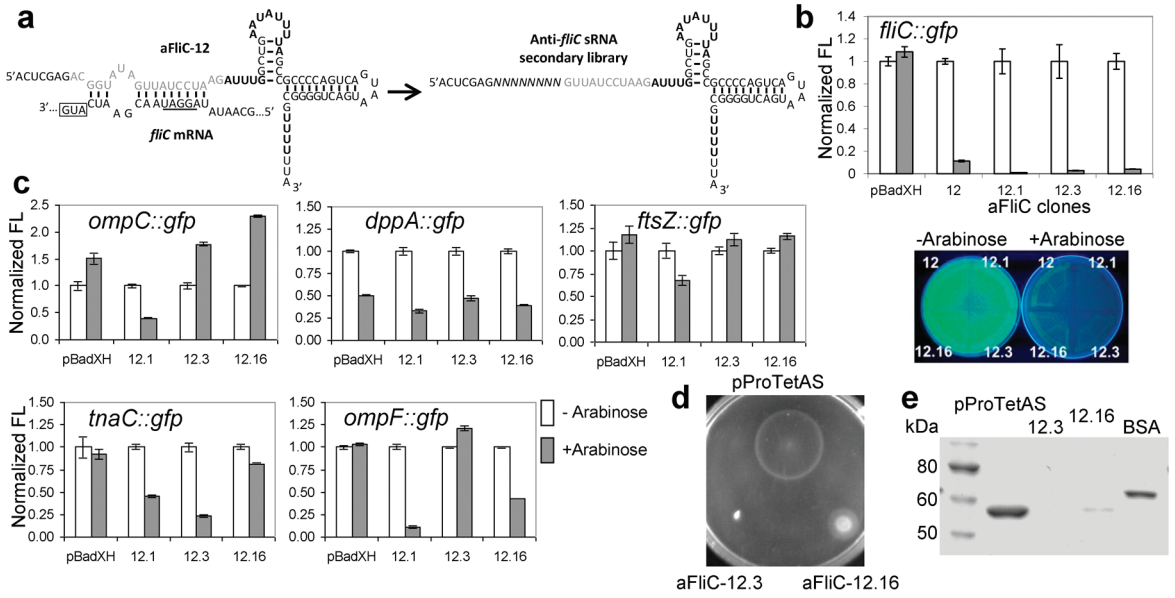


Figure 3. Artificial sRNAs targeting *fliC*. (a) Schematic illustration of the secondary library design based on aFliC-12. The core 9-base antisense sequence in aFliC-12 was conserved, and the adjacent 8 bases in the artificial sRNA were randomized. The model depicting aFliC-12 targeting the *fliC* mRNA was derived using IntaRNA. (b) Reporter gene assay (*fliC::gfp*) of the anti-*fliC* artificial sRNAs in liquid culture (graph) and on LB agar plates (photograph) visualized over UV (365 nm) illumination. (c) Effects of the sRNAs on other untargeted mRNA leader sequences. (d) Swimming motility assay using *E. coli* MG1655 cells expressing aFliC-12.3 and aFliC-12.16. Cells transformed with the empty vector (pProTetAS) were used as a negative control. The artificial sRNAs dramatically reduce cell motility. (e) Detection of the extracellular flagellin protein (FliC) isolated from MG1655 cells expressing the artificial sRNAs. While strong FliC expression was observed without sRNA expression (pProTetAS), FliC was barely detectable in the cells expressing the artificial sRNAs. Bovine serum albumin (BSA) was loaded as a control. (b, c) The data (in the graphs) are averages of triplicate overnight cultures, and the error bars indicate SD. Cellular fluorescence in the absence of arabinose was normalized to 1.0.

the sRNA activities during visual colony screening (i.e., increased dynamic range). Second, we designed a new focused artificial sRNA library based on the best artificial sRNA clone from the initial screen (aFliC-12). As shown in Figure 3a, aFliC-12 was predicted to hybridize with and near the SD sequence in the *fliC* 5′ UTR. A new focused library was constructed by conserving the core 9-base antisense sequence of aFliC-12 while randomizing the adjacent eight bases on the 5′ end. A total of 7.5×10^4 unique clones were generated for the secondary library. Not surprisingly, a majority of the colonies showed reduced fluorescence. Therefore, we sampled and sequenced three clones (aFliC-12.1, aFliC-12.3, and aFliC-12.16) for more detailed characterization.

Characterization of Artificial sRNAs Targeting *fliC* 5′ UTR. The anti-*fliC* artificial sRNAs suppressed the *fliC::gfpuv* reporter gene expression with respectable efficiency. Although aFliC-12

discovered in the initial screen yielded 9-fold repression when assayed with pKP33-ΔlacO-FliC::GFPuv, those isolated from the secondary library exhibited repression of 23- to 85-fold in overnight cultures (Figure 3b). Qualitatively similar results were obtained in log-phase cultures (Supplementary Figure S5). These results suggest that although the initial screens may not yield highly active artificial sRNAs, secondary screens based on the initial screening outcomes may lead to improved sRNA activity.

The anti-*fliC* sRNAs isolated from the secondary library (aFliC-12.1, aFliC-12.3, and aFliC-12.16) were subsequently analyzed for target specificity (Figure 3c). While aFliC-12.1 exhibited the strongest anti-*fliC* activity in the reporter gene expression assay (Figure 3b), it showed significant nonspecific activities against the other genes tested, in particular against *ompF*, which was downregulated 9-fold (Figure 3c). On the other hand, aFliC-12.3 showed no significant activities against the panel of genes compared to the empty vector (pBadXH) except for the modest downregulation of *tnaC* (4-fold) (Figure 3c). Similarly, aFliC-12.16 displayed no significant activities except for the weak repression of *ompF* (2-fold) (Figure 3c).

Finally, *E. coli* cells expressing the anti-*fliC* sRNAs were assayed for cell motility. We found that cell motility of TOP10 cells was significantly compromised compared to the strains with fewer genomic manipulations. Therefore, we used MG1655 that exhibits high cell motility. The two anti-*fliC* artificial sRNAs (aFliC-12.3 and aFliC-12.16) were subcloned into a high-copy plasmid (pPROTetAS) under the control of a strong constitutive promoter (P_{LtetO}) (Supplementary Figure S6c). As shown in Figure 3d, cells expressing the artificial sRNAs (aFliC-12.3 and aFliC-12.16) exhibited significantly reduced motility, consistent with other results. Furthermore, extracellular flagellin protein was isolated and analyzed by SDS-PAGE.[33,34] As expected, we observed significant reduction of FliC (Figure 3e) consistent with the cell motility assay.

Mutational Analysis of aFliC-12.3. To gain a better understanding of the role of the antisense domain of the artificial sRNAs, we designed and evaluated several mutants of aFliC-12.3. First, IntaRNA was used to construct a model of aFliC-12.3 targeting the *fliC* mRNA leader sequence. The model predicts a perfect 11-bp hybridization at and near the SD sequence of the *fliC* mRNA (Figure 4a). Two-base mutations M1 and M2, both predicted to pair with the SD sequence, resulted in almost complete loss of activity (Figure 4b). Although less dramatic, two-base mutation at the end of the putative antisense sequence (M3) also resulted in a significant loss of target gene repression (Figure 4b). These results support the antisense mechanism model of the artificial sRNAs.

Furthermore, the first 21 bases of the Spot42 scaffold, which contains two of the three U-rich regions known to contact Hfq,[5] were removed to evaluate the role of the scaffold (M4) (Figure 4a). The remaining U-rich region immediately downstream of the 3′ stem-loop of the scaffold was retained as it also functions as a part of the transcription terminator. Although M4 was partially active, its repression efficiency (6.1-fold repression) was measurably compromised compared to that of aFliC-12.3 (19.6-fold repression) (Figure 4b). Similarly, substitution of the Spot42 scaffold with the MicF, DsrA, or GcvB scaffold reduced the artificial sRNA efficacy (4.0-, 8.8-, and 3.8-fold repression, respectively) (Figure 4b). Together, these results suggest that the sRNA scaffolds play a significant role in artificial sRNA function.

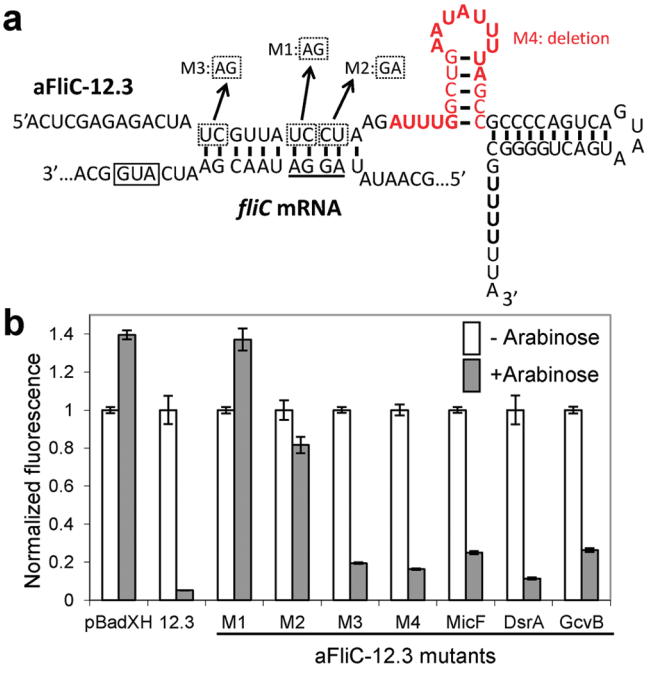


Figure 4. Mutational studies of aFliC-12.3. (a) Model of aFliC-12.3 hybridized with the *fliC* mRNA as predicted by IntaRNA. Two-base mutations (M1−M3) prepared in this study are indicated by the dotted boxes. Partial deletion of the scaffold (M4) is shown in red. The SD sequence is underlined, and the start codon is boxed in the *fliC* mRNA. (b) Reporter gene assay (*fliC::gfp*) of the aFliC-12.3 mutants. MicF, DsrA, and GcvB indicate scaffold substituted analogues of aFliC-12.3. The data are averages of triplicate overnight cultures. and the error bars indicate SD. Cellular fluorescence in the absence of arabinose was normalized to 1.0.

Summary and Conclusion. In this study, we described a practical strategy for developing bacterial artificial sRNAs that posttranscriptionally knockdown endogenous gene expression by targeting the 5′ UTR and the translation initiation region of endogenous mRNAs. The key elements of the strategy include the use of natural sRNA scaffolds that are known to recruit the cellular proteins involved in gene silencing such as Hfq[5,28] and high-throughput screening based on the GFP reporter system.[35] The strategy was inspired by the previously observed modular architectures of the natural sRNAs.[7,8] The successful discovery of numerous functional sRNAs that target *ompF* and *fliC* from a relatively small library of partially randomized artificial sRNAs strongly indicate that our strategy may be generally applicable to other genes in *E. coli* and possibly in other bacteria. It should be noted that numerous parameters such as stabilities and secondary structures of the RNA species are expected to affect the overall performance of the artificial sRNAs. One advantage of the described screening strategy is that it is possible to identify potent sRNAs without knowing the underlying mechanisms. However, it will be of significant interest to investigate how these artificial sRNAs function and how their mechanisms depend on the sRNA scaffolds and/or target mRNAs.

Artificial sRNAs enable dynamic and tunable regulation of bacterial genes that may be useful for various applications in synthetic biology. Ultimately, accumulation and analysis of additional artificial sRNAs targeting diverse genes may lead to robust design principles facilitating rational design.

METHODS

Bacterial Strains and Growth Conditions. *E. coli* TOP10 cells (Invitrogen, F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC), φ80*lacZ*ΔM15, Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*)7697, *gal*U, *gal*K, *rps*L (StrR), *end*A1, *nup*G) were used in all experiments, unless specified otherwise. Motility assays and flagellin analysis were performed using *E. coli* K-12 MG1655 cells (F- λ- *ilvG- rfb-50 rph-1*) obtained from Coli Genetic Stock Center (CGSC). The cells were grown at 37 °C (unless specified otherwise) from single-colony isolates or diluted from overnight cultures in LB media supplemented with L-arabinose (Calbiochem) as necessary. For plasmid maintenance, ampicillin (Fisher Scientific) and kanamycin (Sigma) were used at concentrations of 100 and 50 μg/mL, respectively.

Plasmid Construction. *E. coli* MG1655 genomic DNA was used for cloning the wild-type sRNAs and the target mRNA leader sequences. The parent sRNA expression vector pBadXH ("empty vector") was derived from pBAD/HisB (Invitrogen), which includes P_{araBAD} promoter and *araC* for arabinose-inducible gene expression on the pBR322 backbone with an ampicillin-selectable marker (Supplementary Figure S6a). Genes encoding the small RNAs were cloned into the XhoI and *Hin*dIII restriction sites in pBadXH. The reporter plasmids were derived from pK184, which contains a p15a origin of replication and a kanamycin-selectable marker.[36] The parent vector pKP33-GFPuv contains a synthetic constitutive promoter P_{33} (BBa_J23110, Registry of Standard Biological Parts, http://partsregistry.org/Part:BBa_J23110), *Eco*RI and *Bam*HI restriction sites for cloning an mRNA leader sequence (using primers listed in Supplementary Table S1), and *gfpuv* fluorescent protein reporter gene (Supplementary Figure S6b). The artificial sRNA expression plasmids used in flagellin analysis and swimming motility assays in MG1655 were derived from pPROTet.E333 (Clontech) with a high-copy origin of replication (ColE1). The plasmid pPROTetE.333 was modified to contain an ampicillin-selectable marker and appropriate restriction sites to yield pPROTetAS ("empty vector"), which was used to clone the anti-*fliC* artificial sRNAs (Supplementary Figure S6c). Detailed sequence information of the plasmids and primers described in this work is given in Supporting Information (Figures S6, S7 and Table S1).

Library Construction. Natural sRNA expression vectors (pBAD-DsrA/GcvB/MicF/Spot42) were used as templates in PCR reactions to substitute the antisense domain with 20 or 30 degenerate bases. A common reverse primer Sartrev that contains 20 degenerate bases was used with a sRNA-specific forward primer with none or 10 degenerate bases (Supplementary Table S1) to amplify the whole plasmid, self-ligated, and transformed into TOP10 cells to construct the artificial sRNA libraries. Libraries based on the four scaffolds were constructed separately, and purified plasmid libraries were mixed in equal amounts to yield the combined library. The complexities of the individual scaffold libraries ranged from 1×10^6 to 5×10^6, and the combined library (all four scaffolds) theoretically contained 1.2×10^7 unique clones. Note that the typical number of clones screened was significantly smaller ($\sim 10^5$).

Library Screening. TOP10 competent cells harboring a reporter plasmid were transformed with an appropriate mixture of the artificial sRNA library plasmids and plated on LB agar plates supplemented with 0.1% arabinose, ampicillin, and kanamycin. Cells were plated so that approximately 3,000−4,000 colonies grew on each plate (100 mm diameter). The plates were first incubated at 37 °C overnight and subsequently cooled at 4 °C to enhance GFPuv fluorescence. The plates were examined visually over a UV transilluminator (UVP) at 365 nm for colonies with diminished fluorescence. Promising colonies were restreaked on fresh plates to confirm the artificial sRNA activity and to ensure clonal isolation. Furthermore, plasmids encoding active artificial sRNA clones were isolated and retransformed to eliminate the possibility of mutations in the host genome or the reporter plasmid.

Fluorescent Reporter Gene Assays for Artificial sRNA Activities. TOP10 cells transformed with an artificial sRNA expression plasmid and a reporter plasmid (mRNA leader-*gfpuv* fusion) were used in the reporter gene assays for the artificial sRNA activities. The cells were cultured overnight from single colonies or further diluted 100-fold (v/v) into fresh LB (1 mL) and grown for 5 h or overnight at 37 °C. Next, the cells (200 μL) were harvested by centrifugation, washed once with phosphate-buffered saline (PBS), and transferred to a 96-well plate in PBS (200 μL). Optical density at 600 nm (OD_{600}) and GFPuv fluorescence (excitation wavelength 395 nm; emission wavelength 509 nm) of the cells were measured using Safire[2] microplate reader (Tecan). Fluorescence data from TOP10 cells transformed with pBadXH were used to subtract the background cellular fluorescence. The background-corrected fluorescence data were further normalized by OD_{600}. Finally, for each sRNA clone, fluorescence of the sRNA-induced cells (+arabinose) was normalized by that of the uninduced cells (= 1.0, −arabinose).

Expression Levels of Outer Membrane Porins. The cells containing an appropriate artificial sRNA expression plasmid were grown in LB medium (25 mL) with or without 0.1% arabinose to OD_{600} ~0.6 at 24 °C shaken at 275 rpm. The cells were harvested by centrifugation at 5000*g* for 30 min at 4 °C, washed once with LB medium, and stored at −80 °C until use. Components of the cell envelope were isolated as previously described.[31,37,38] The isolated proteins were separated by SDS-PAGE in 12% polyacrylamide gels containing 4 M urea and stained with SimplyBlue SafeStain (Invitrogen).

Expression Levels of Flagellin. FliC expression levels in MG1655 were analyzed as previously described.[33,34] Briefly, the cells were grown in LB medium (1.5 mL) at 37 °C for 6 h. The cells were pelleted (5000*g*, 10 min, 4 °C) and resuspended in 0.9% NaCl (50 μL). Flagellin (FliC) was released from the cells by vigorous agitation on a vortex mixer for 5 min. The cells were removed by centrifugation, and the supernatant was analyzed on 12% SDS-PAGE.

Swimming Motility Assay. Soft agar plates (1% Bactotryptone, 0.5% NaCl, 0.25% agar) were prepared fresh on the day of cell inoculation in 10 mm plates (25 mL per plate). The plates were allowed to solidify for at least 2 h prior to inoculation. Tryptone-broth-grown cultures were used to inoculate the soft agar plates in 2 μL volume ($\sim 10^5$ cells) by stabbing the agar with a pipet tip and expelling the culture gently into the medium. After 10 min of incubation at room temperature to ensure that the cells were properly inoculated, the plates were incubated at 29 °C for 4 h and photographed.

ASSOCIATED CONTENT

Supporting Information. Supporting figures and table. This material is available free of charge via the Internet at http://pubs.acs.org.

Notes

The authors declare no competing financial interest.

AUTHOR INFORMATION

Corresponding Author

*Tel: +1-530-754-9676. Fax: +1-530-754-5739. E-mail: yoko@ucdavis.edu.

ACKNOWLEDGMENT

We thank Chris Noriega (Valley Stewart lab, UC Davis) for help with the outer membrane protein analysis. Michelle Igo and the Department of Microbiology (UC Davis) generously allowed us to use their facilities. This work was supported by New Research Initiative grant from UC Davis Academic Senate and National Science Foundation (1016357).

REFERENCES

(1) Hannon, G. J., and Rossi, J. J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature 431*, 371–378.

(2) Meister, G., and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA. *Nature 431*, 343–349.

(3) Mizuno, T., Chou, M. Y., and Inouye, M. (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc. Natl. Acad. Sci. U.S.A. 81*, 1966–1970.

(4) Gottesman, S., and Storz, G. (2010) Bacterial small RNA regulators: versatile roles and rapidly evolving variations *Cold Spring Harbor Perspect Biol.* Epub October 27, 2010, DOI: 10.1101/cshperspect.a003798.

(5) Møller, T., Franch, T., Højrup, P., Keene, D. R., Bächinger, 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.

(6) Møller, T., Franch, T., Udesen, C., Gerdes, K., and Valentin-Hansen, P. (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev. 16*, 1696–1706.

(7) Bouvier, M., Sharma, C. M., Mika, F., Nierhaus, K. H., and Vogel, J. (2008) Small RNA binding to 5′ mRNA coding region inhibits translational initiation. *Mol. Cell 32*, 827–837.

(8) Papenfort, K., Bouvier, M., Mika, F., Sharma, C. M., and Vogel, J. (2010) Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA. *Proc. Natl. Acad. Sci. U.S.A. 107*, 20435–20440.

(9) Coleman, J., Green, P. J., and Inouye, M. (1984) The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. *Cell. 37*, 429–436.

(10) Kim, J. Y. H., and Cha, H. J. (2003) Down-regulation of acetate pathway through antisense strategy in *Escherichia coli*: Improved foreign protein production. *Biotechnol. Bioeng. 83*, 841–853.

(11) Nakashima, N., and Tamura, T. (2009) Conditional gene silencing of multiple genes with antisense RNAs and generation of a mutator strain of *Escherichia coli*. *Nucleic Acids Res. 37*, e103.

(12) Nakashima, N., Tamura, T., and Good, L. (2006) Paired termini stabilize antisense RNAs and enhance conditional gene silencing in *Escherichia coli*. *Nucleic Acids Res. 34*, e138.

(13) Stefan, A., Schwarz, F., Bressanin, D., and Hochkoeppler, A. (2010) Shine-Dalgarno sequence enhances the efficiency of lacZ repression by artificial anti-lac antisense RNAs in *Escherichia coli*. *J. Biosci. Bioeng. 110*, 523–528.

(14) Stefan, A., Tabler, M., and Hochkoeppler, A. (2007) Efficient silencing of the gene coding for the epsilon subunit of DNA polymerase III in *Escherichia coli* is triggered by antisense RNAs featuring stability in vivo. *FEMS Microbiol. Lett. 270*, 277–283.

(15) Man, S., Cheng, R., Miao, C., Gong, Q., Gu, Y., Lu, X., Han, F., and Yu, W. (2011) Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res. 39*, e50.

(16) Bartel, D. P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell. 136*, 215–233.

(17) Kawamoto, H., Koide, Y., Morita, T., and Aiba, H. (2006) Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol. Microbiol. 61*, 1013–1022.

(18) Schmidt, M., Zheng, P., and Delihas, N. (1995) Secondary structures of *Escherichia coli* antisense micF RNA, the 5′-end of the target ompF mRNA, and the RNA/RNA duplex. *Biochemistry 34*, 3621–3631.

(19) Andersen, J., Delihas, N., Ikenaka, K., Green, P. J., Pines, O., Ilercil, O., and Inouye, M. (1987) The isolation and characterization of RNA coded by the micF gene in *Escherichia coli*. *Nucleic Acids Res. 15*, 2089–2101.

(20) Kuwajima, G., Asaka, J., Fujiwara, T., Fujiwara, T., Node, K., and Kondo, E. (1986) Nucleotide sequence of the hag gene encoding flagellin of *Escherichia coli*. *J. Bacteriol. 168*, 1479–1483.

(21) Macnab, R. M., and Parkinson, J. S. (1991) Genetic analysis of the bacterial flagellum. *Trends Genet. 7*, 196–200.

(22) Kuwajima, G. (1988) Construction of a minimum-size functional flagellin of *Escherichia coli*. *J. Bacteriol. 170*, 3305–3309.

(23) Sledjeski, D. D., Gupta, A., and Gottesman, S. (1996) The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J. 15*, 3993–4000.

(24) Brescia, C. C., Mikulecky, P. J., Feig, A. L., and Sledjeski, D. D. (2003) Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA 9*, 33–43.

(25) Lease, R. A., Smith, D., McDonough, K., and Belfort, M. (2004) The small noncoding DsrA RNA is an acid resistance regulator in *Escherichia coli*. *J. Bacteriol. 186*, 6179–6185.

(26) Pulvermacher, S. C., Stauffer, L. T., and Stauffer, G. V. (2009) Role of the *Escherichia coli* Hfq protein in GcvB regulation of oppA and dppA mRNAs. *Microbiology 155*, 115–123.

(27) 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.

(28) Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R., and Storz, G. (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J. 17*, 6061–6068.

(29) Smith, C., Heyne, S., Richter, A. S., Will, S., and Backofen, R. (2010) Freiburg RNA Tools: a web server integrating INTARNA, EXPARNA and LOCARNA. *Nucleic Acids Res. 38*, W373–377.

(30) Vogel, J., and Papenfort, K. (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr. Opin. Microbiol. 9*, 605–611.

(31) Chen, S., Zhang, A., Blyn, L. B., and Storz, G. (2004) MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J. Bacteriol. 186*, 6689–6697.

(32) Andersen, J., Forst, S. A., Zhao, K., Inouye, M., and Delihas, N. (1989) The function of micF RNA. micF RNA is a major factor in the thermal regulation of OmpF protein in *Escherichia coli*. *J. Biol. Chem. 264*, 17961–17970.

(33) Nishida, S., Mizushima, T., Miki, T., and Sekimizu, K. (1997) Immotile phenotype of an *Escherichia coli* mutant lacking the histone-like protein HU. *FEMS Microbiol. Lett. 150*, 297–301.

(34) Kuwajima, G., Kawagishi, I., Homma, M., Asaka, J., Kondo, E., and Macnab, R. M. (1989) Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway. *Proc. Natl. Acad. Sci. U.S.A. 86*, 4953–4957.

(35) Urban, J. H., and Vogel, J. (2007) Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res. 35*, 1018–1037.

(36) Jobling, M. G., and Holmes, R. K. (1990) Construction of vectors with the p15a replicon, kanamycin resistance, inducible lacZ alpha and pUC18 or pUC19 multiple cloning sites. *Nucleic Acids Res. 18*, 5315–5316.

(37) Filip, C., Fletcher, G., Wulff, J. L., and Earhart, C. F. (1973) Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. *J. Bacteriol. 115*, 717–722.

(38) Forst, S., Delgado, J., Ramakrishnan, G., and Inouye, M. (1988) Regulation of ompC and ompF expression in *Escherichia coli* in the absence of envZ. *J. Bacteriol. 170*, 5080–5085.