TITLE PAGE

Artificial *trans***-encoded small noncoding RNAs specifically silence the selected gene expression in bacteria**

Shuai Man# , Rubin Cheng# , Cuicui Miao, Qianhong Gong, Yuchao Gu, Xinzhi Lu, Feng Han, and Wengong Yu*

Key Laboratory of Marine Drugs, Chinese Ministry of Education; Key Laboratory of Glycoscience & Glycotechnology of Shandong Province; School of Medicine and Pharmacy, Ocean University of China, 5 Yushan Road, Qingdao 266003, China.

These authors contributed equally to this work.

*To whom correspondence should be addressed. Tel: +86-532-82031680. Fax: +86-532-82033054. E-mail: yuwg66@ouc.edu.cn

Keywords: atsRNA, bacteria, gene silencing, trans-encoded sRNA

Figure S1. The common structural characteristics of well-known bacterial *trans*-encoded sRNAs.

Figure S2. Effect of atsRNA on growth rate of the host bacteria *E. coli*.

Figure S3. Effect of atsRNA on non-target gene.

Figure S4. The base-pairing schemes for atsRNA (including GY series and CY series) and target mRNA.

Figure S5. Predicted secondary structure of atsRNA targeting EGFP and *uidA* gene.

Figure S6. The base-pairing schemes for atsRNAs (including CY4, L-CY4, CY6 and L-CY6) and their target *uidA* mRNA.

Figure S7. The amount and relative interference efficiency of CY4, CY6 and their corresponding atsRNA mutants.

Figure S8. The base-pairing schemes for atsRNA (including CY9, CY9-L1, CY9-L2, CY9-L3 and CY9-L4) and their target *uidA* mRNA.

Figure S9. The amount and relative interference efficiency of CY9 and CY9 mutants.

Figure S10. The stability and relative interference efficiency of atsRNAs CY1 and N-CY1.

Figure S11. The stability and relative interference efficiency of GY2, CY9 and their corresponding mRNA basepairing region null atsRNA mutants.

Figure S12. The stability and relative interference efficiency of GY2, CY9 and their corresponding Hfq binding site null atsRNA mutants.

Figure S13. Effect of atsRNAs on target gene in hfq mutant strain after the transformation of a multicopy plasmid containing hfq gene.

Figure S14. Affinity of Hfq for atsRNA revealed by gel mobility shift assays.

Figure S15. The interaction of atsRNA and target mRNA in vitro with or without the aid of Hfq revealed by gel mobility shift assays.

Figure S16. Quantitative real-time PCR analysis of uidA mRNA in HAT103 (temperature sensitive rne-1 mutant strain) containing rne gene when it was cultured at 42° C.

Figure S17. Predicted secondary structure of atsRNAs targeting against identified essential genes.

Figure S18. Ouantitative real-time PCR analysis of target mRNA in MC4100.

Figure S19. Effect of atsRNAs on *hla* gene in *S. aureus*.

Figure S20. Effect of atsRNAs on SEA gene in *S. aureus*.

Figure S21. Effect of atsRNA on growth rate of the host bacteria *S. aureus*.

Supplementary Figures Legends

Figure S1. The common structural characteristics of well-known bacterial *trans*-encoded sRNAs. The sRNAs contained roughly three parts: mRNA basepairing region (red), Hfq binding site (blue) and Rho-independent terminator (pink). At certain instances, two parts of the sRNAs are intertwined together. Overlapped area between mRNA basepairing region and Hfq binding site is in green, and that between Hfq binding site and Rho-independent terminator is in gray. The corresponding target mRNAs for the highlighted mRNA basepairing regions of natural *trans*-encoded sRNAs are as follows: *sodB* mRNA for RyhB, *flhA* mRNA for OxyS, *rpoS* mRNA for DsrA, *galK* mRNA for Spot42, *ompF* mRNA for MicF, *hapR* mRNA for Qrr1, respectively.

Figure S2. Effect of atsRNA on growth rate of the host bacteria. Growth curve of host bacteria MC4100 after the expression of atsRNAs targeting against EGFP gene (A) or *uidA* gene (B), respectively. Bacterial cells containing empty vector pRI were used as the control. Cells were grown in LB media with appropriate antibiotics at 30°C. Turbidity of the cultures was monitored at 600 nm.

Figure S3. Effect of atsRNA on non-target genes. (A) Effect of CY series atsRNAs on EGFP gene in MC4100. The fluorescence was measured after the expression of atsRNA was induced with 1 mM IPTG for 30 min. (B) Effect of GY series atsRNAs on *uidA* gene in MC4100. The activity of beta-glucuronidase was assayed after the expression of atsRNA was induced with 1 mM IPTG for 30 min. Bacterial cells containing the empty vector pRI were used as the control. Average values were from triplicate experiments with std values indicated.

Figure S4. The base-pairing schemes for atsRNA and target mRNA. (A) Hybrids between GY series atsRNAs and EGFP mRNA. (B) Hybrids between CY series atsRNAs and *uidA* mRNA. For each hybrid, the lower part shows the region of atsRNA that pairs with target mRNA (upper). The base-pairing sequences between atsRNAs and target mRNAs were predicted by LALIGN and the thermodynamic pairing energies were calculated by DINAMelt. The Shine–Dalgarno (SD) sequences (indicated with arrows) and translation start codons (indicated with asterisks) are shown.

Figure S5. The predicted secondary structure of atsRNA targeting against EGFP gene and *uidA* gene. The GY series atsRNAs were designed to target against EGFP gene (A) and the CY series atsRNAs were designed to target against *uidA* gene (B). The sequences of predicted pairing region (indicated with arrows) and Hfq binding site (indicated with asterisks) of atsRNA are marked. The detailed sequences of these atsRNAs were listed in Table S1. The secondary structure of atsRNA is predicted at 37°C by MFOLD program.

Figure S6. The base-pairing schemes for atsRNAs (including CY4, L-CY4, CY6 and L-CY6) and their target *uidA* mRNA. For each hybrid, the lower part shows the region of atsRNA that pairs with target mRNA (upper). The base-pairing sequences between atsRNAs and target mRNAs were predicted by LALIGN and the thermodynamic pairing energies were calculated by DINAMelt. The Shine–Dalgarno (SD) sequences (indicated with arrows) and translation start codens (indicated with asterisks) are shown.

Figure S7. The amount and relative interference efficiency of CY4, CY6 and their corresponding atsRNA mutants. (A) The amounts of atsRNAs analyzed by Northern blot. Total RNAs were purified from cells after the expression of atsRNAs were induced with 1 mM IPTG for 30 min and analyzed using probes specific for atsRNAs and 5S rRNA respectively. (B) The relative interference efficiency of atsRNAs. We obtained the relative interference efficiency of each atsRNAs by normalizing the interference efficiency to relative RNA amounts.

Figure S8. The base-pairing schemes for atsRNA (including CY9, CY9-L1, CY9-L2, CY9-L3 and CY9-L4) and their target *uidA* mRNA. For each hybrid, the lower part shows the region of atsRNA that pairs with target mRNA (upper). The base-pairing sequences between atsRNAs and target mRNAs were predicted by LALIGN and the thermodynamic pairing energies were calculated by DINAMelt. The Shine–Dalgarno (SD) sequences (indicated with arrows) and translation start codens (indicated with asterisks) are shown.

Figure S9. The amount and relative interference efficiency of CY9 and CY9 mutants. (A) The amounts of atsRNAs analyzed by Northern blot. Total RNAs were purified from cells after the expression of atsRNAs were induced with 1 mM IPTG for 30 min and analyzed using probes specific for CY9 and 5S rRNA respectively. (B) The relative interference efficiency of atsRNAs. We obtained the relative interference efficiency of each atsRNAs by normalizing the interference efficiency to relative RNA amounts.

Figure S10. The stability and relative interference efficiency of atsRNAs CY1 and N-CY1. (A) The stability of CY1 and N-CY1 analyzed by Northern blot. The expression of atsRNAs was induced with 1 mM IPTG for 30 min, and then rifampicin was added. The incubation was continued further and cells were harvested at the indicated time points for RNA preparation. Total RNA was analyzed using probes specific for CY1, N-CY1 and 5S rRNA respectively. The relative RNA amounts were calculated according to the RNA amounts at time 0. (B) The relative interference efficiency of atsRNAs. We obtained the relative interference efficiency of each atsRNAs by normalizing the interference efficiency to relative RNA amounts.

Figure S11. The stability and relative interference efficiency of GY2, CY9 and their corresponding mRNA basepairing region null atsRNA mutants. (A) The stability of atsRNAs analyzed by Northern blot. The expression of atsRNAs was induced with 1 mM IPTG for 30 min, and then rifampicin was added. The incubation was continued further and cells were harvested at the indicated time points for RNA preparation. Total RNA was analyzed using probes specific for atsRNAs and 5S rRNA respectively. (B) The relative interference efficiency of atsRNAs. The relative RNA amounts were calculated according to the RNA amounts at time 0. We obtained the relative interference efficiency of each atsRNAs by normalizing the interference efficiency to relative RNA amounts.

Figure S12. The stability and relative interference efficiency of GY2, CY9 and their corresponding Hfq binding site null atsRNA mutants. (A) The stability of atsRNAs analyzed by Northern blot. The expression of atsRNAs was induced with 1 mM IPTG for 30 min, and then rifampicin was added. The incubation was continued further and cells were harvested at the indicated time points for RNA preparation. Total RNA was analyzed using probes specific for atsRNAs and 5S rRNA respectively. (B) The relative interference efficiency of atsRNAs. The relative RNA amounts were calculated according to the RNA amounts at time 0. We obtained the relative interference efficiency of each atsRNAs by normalizing the interference efficiency to relative RNA amounts.

Figure S13. Effect of atsRNAs on target *uidA* gene in *hfq* mutant strain GSO81 after

the transformation of a multicopy plasmid containing *hfq* gene. The *hfq* gene, including its promoter and transcription terminator, was cloned into pACYC177 vector which was compatible with the atsRNA expression vector. The activity of beta-glucuronidase was assayed after the expression of atsRNA was induced with 1mM IPTG for 30 min. Bacteria cells containing empty vector pRI were used as the control. Average values were from triplicate experiments with std values indicated.

Figure S14. Affinity of Hfq for atsRNA revealed by gel mobility shift assays. 5 nM labelled atsRNAs GY2 (A), Hfq-null-GY2 (B), CY9 (C) and Hfq-null-CY9 (D) were incubated in the absence (lane 1), in the presence of 5nM (lane 2), 10nM (lane 3), 20nM (lane 4), of Hfq protein, respectively. The free RNA and RNA-protein binding complex were indicated with F and B respectively.

Figure S15. The interaction of atsRNA and target mRNA in *vitro* with or without the aid of Hfq revealed by gel mobility shift assays. The target mRNA was synthesized in vitro using T7 RNA polymerase and PCR fragments as the templates. The length of the synthesized target mRNA was 146 nt for EGFP (from $+1$ to $+146$) and 137 nt for *uidA* (from +1 to +137), respectively. The primer sequences used for transcription in *vitro* were as follows: forward primer of EGFP gene, 5'-CTAATACGACTCACTATAGGGAGTCTAGAAATAATTTTGTTTAAC-3',

reverse primer of EGFP gene, 5'-GATGGGCACCACCCCGGTGAAC-3'; forward primer of *uidA* gene, 5'-CTAATACGACTCACTATAGGGAGCTTAATGAGGAGTCCCTTATGT-3',

reverse primer of *uidA* gene, 5'- TGTAACGCGCTTTCCCACCAACG-3'. 5 nM labelled atsRNAs GY2 (A) and CY9 (B) were incubated alone (lane 1), or with 0.5 nM unlabelled target mRNA (lane 2), 10 nM Hfq alone (lane 3), 0.5 nM unlabelled target mRNA and 10 nM Hfq together (lane 4). The free atsRNAs, complexes corresponding to atsRNA-target mRNA, atsRNA-Hfq and atsRNA-target mRNA-Hfq are indicated on the right.

Figure S16. Quantitative real-time PCR analysis of *uidA* mRNA in HAT103 (temperature sensitive *rne*-1 mutant strain) containing *rne* gene when it was cultured at 42°C. The *rne* gene, including its promoter and transcription terminator, was cloned into pACYC177 vector which was compatible with the atsRNA expression vector. The mRNA level was detected after the expression of atsRNAs CY4 and CY9 were induced with 1 mM IPTG for 30 min. Bacterial cells containing empty vector pRI were used as the control. The value was normalized to the level of 16S rRNA in each sample.

Figure S17. The predicted secondary structure of atsRNAs targeting against identified essential genes. M series atsRNAs were designed to target against *murA* gene, T series atsRNAs were designed to target against *trmA* gene and Y series atsRNAs were designed to target against *ygjD* gene. The sequences of predicted pairing region (indicated with arrows) and Hfq binding site (indicated with asterisks) of atsRNA are marked. The detailed sequences of these atsRNAs were listed in Table S5. The secondary structure of atsRNA is predicted at 37°C by MFOLD program.

Figure S18. Quantitative real-time PCR analysis of target mRNA in MC4100. Quantitative real-time PCR analysis of *murA* mRNA (A), *trmA* mRNA (B) and *ygjD* mRNA (C). IPTG was added to a final concentration of 0.1 mM to induce the expression of atsRNA when OD_{600} reached 0.5. The mRNA level was detected after

the expression of atsRNAs was induced for 30 min. Bacterial cells containing empty vector pRI were used as the control. The value was normalized to the level of 16S rRNA in each sample.

Figure S19. Effect of atsRNAs on *hla* gene in *S. aureus*. (A) The effect of atsRNAs on α-hemolysin activity. (B) Northern blot analysis of atsRNA transcripts targeting *hla* gene. Bacterial cells containing the empty vector pSB2035 were used as the control. The 5S rRNA was used as the loading control.

Figure S20. Effect of atsRNAs on SEA gene in *S. aureus*. (A) The effect of atsRNAs on *S. aureus* enterotoxin activity. (B) Upper: comparison of *S. aureus* enterotoxin A levels. Lower: Northern blot analysis of atsRNA transcripts targeting SEA gene. Bacterial cells containing the empty vector pSB2035 were used as the control. The 5S rRNA was used as the loading control.

Figure S21. Effect of atsRNA on growth rate of the host bacteria *S. aureus*. Growth curve of host bacteria *S. aureus* RN4220 after the expression of atsRNAs targeting against *hla* gene (A) or SEA gene (B), respectively. Bacterial cells containing empty vector pSB2035 were used as the control. Cells were grown in LB media with appropriate antibiotics at 37°C. Turbidity of the cultures was monitored at 600 nm.

- RyhB UAUUACUUAGCCAGCCGGGUGCUGGCUUUUU
- OxyS UCGAAU AACU AAAGCCAACGUGAACUUUUGCGGAUCUCCAGGAUCCGC
- DsrA AUCCCGACCCCCUCAGGGUCGGGAUUU
- GGCUGAAUAUUUUAGCCGCCCCAGUCAGUAAUGACUGGGCCGUUUUUUA Spot42
- MicF UUUCGACCGGAUGCUUCGC AUCCGGUUUUUUUU
- **AACAGCCAAUAGACUCAUUCUAUUGGUCAUUUUUUU** Qm1

Supplementary Figure S2

Supplementary Figure S3

Supplementary Figure S6

A

 \boldsymbol{B}

Supplementary Figure S8

Supplementary Figure S10 $\mathbf A$

Supplementary Figure S11

 $A-GY2$

 $C-CY9$

B-(Hfq-null-GY2)

D-(Hfq-null-CY9)

Supplementary Figure 16

Supplementary Figure S18

Supplementary Figure S19 \overline{A}

 \bf{B}

5S

Supplementary Figure S20 \mathbf{A}

Supplementary Figure S21 \overline{A}

Supplementary Tables

 Table S1 Sequence of atsRNA targeting against EGFP gene

The sequences of GY series atsRNAs (targeted against EGFP gene) were listed above. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue). The three elements of atsRNA were color labeled to make it easy to discern the three modules of atsRNA. The sources of Hfq binding site and Rho-independent terminator from natural *trans*-encoded sRNAs were also listed in this table. In some cases, the sequences of Hfq binding site and Rho-independent terminator need to be adjusted so that the secondary structure of atsRNA meets our designing principle.

Table S2 Sequence of atsRNA targeting against uidA gene

The sequences of GY series atsRNAs CY series atsRNAs (targeted against *uidA* gene) were listed above. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue). The three elements of atsRNA were color labeled to make it easy to discern the three modules of atsRNA. The sources of Hfq binding site and Rho-independent terminator from natural *trans*-encoded sRNAs were also listed in this table. In some cases, the sequences of Hfq binding site and Rho-independent terminator need to be adjusted so that the secondary structure of atsRNA meets our designing principle.

Table S3 Sequence of atsRNA and their corresponding atsRNA mutants

The sequences of atsRNAs (including CY4, CY6, CY9 and CY1) and their corresponding atsRNA mutants (including L-CY4, L-CY6, CY9-L1, CY9-L2, CY9-L3, CY9-L4 and N-CY1) were listed above. Compared with the corresponding atsRNAs, L-CY4 and L-CY6 lost the stem loop of mRNA basepairing region in secondary structure. CY9-L1, CY9-L2, CY9-L3 and CY9-L4 were CY9 mutants with different number of unpaired nucleotides in the loop structure of mRNA basepairing region. The Hfq binding site and Rho-independent terminator remained constant in these mutants. N-CY1 was constructed by adding a 5' stem-loop structure without pairing sequence with target mRNA. The mutant N-CY1 had the same mRNA basepairing region, Hfq binding site and Rho-independent terminator with that of CY1. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue).

Table S4 Sequence of GY2, CY9 and their corresponding mutants

The atsRNA mutants B-P-null-GY2 and Hfq-null-GY2 were designed based on GY2 and the mutants B-P-null-CY9 and Hfq-null-CY9 were designed based on CY9. Compared with the corresponding atsRNAs, the mRNA basepairing regions of B-P-null-GY2 and B-P-null-CY9 were non-pairing sequences with target mRNAs and the Hfq binding sites of Hfq-null-GY2 and Hfq-null-CY9 were GC rich sequences. Furthermore, the two other modules of atsRNA remained constant during this mutation. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue).

Table S5 Sequence of atsRNA targeting against essential gene

The sequences of M series atsRNAs (targeting against *murA* gene), T series atsRNAs (targeting against *trmA* gene) and Y series atsRNAs (targeting against *ygjD* gene) were listed above. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue). The sources of Hfq binding site and Rho-independent terminator from natural *trans*-encoded sRNAs were also listed in this table. In some cases, the sequences of Hfq binding site and Rho-independent terminator need to be adjusted so that the secondary structure of atsRNA meets our designing principle.

Table S6 Sequence of atsRNA targeting against *hla* gene

The sequences of H series atsRNAs (targeted against *hla* gene) were listed above. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue). The predicted hybrid pairing sequences with target mRNAs were marked (underlined sequences). The sources of Hfq binding site and Rho-independent

terminator from natural *trans*-encoded sRNAs were also listed in this table. In some cases, the sequences of Hfq binding site and Rho-independent terminator need to be adjusted so that the secondary structure of atsRNA meets our designing principle.

 Table S7 Sequence of atsRNA targeting against SEA gene

The sequences of S series atsRNAs (targeted against SEA gene) were listed above. The atsRNAs sequences were composed of three elements: mRNA basepairing sequences (red), Hfq binding sites (green) and Rho-independent terminators (blue). The predicted hybrid pairing sequences with target mRNAs were marked (underlined sequences). The sources of Hfq binding site and Rho-independent terminator from natural *trans*-encoded sRNAs were also listed in this table. In some cases, the sequences of Hfq binding site and Rho-independent terminator need to be adjusted so that the secondary structure of atsRNA meets our designing principle.

Name	Sequence information				$dG(kcal/mol)^e$	Target	Inhibitory
							efficiency ^f
	Length ^a	mRNA base-pairing region		Hfq binding site ^d			
		Length $^{\rm b}$	Complementary sequences ^c				
$H-1$	74 nt	28 _{nt}	22nt	12nt	-27.8	hla	21.5%
$H-2$	72nt	26nt	3nt	12nt	-22.8	hla	null
$H-3$	77 nt	30 _{nt}	16nt	14nt	-25.6	hla	16.9%
$H-4$	78nt	30nt	10 _{nt}	14nt	-25.6	hla	14.4%
$H-5$	78 _{nt}	28 _{nt}	22nt	0 _{nt}	-22.5	hla	20.9%
$H-6$	84nt	28 _{nt}	15nt	0 _{nt}	-22.8	hla	16.1%
$H-7$	81nt	29nt	12nt	14nt	-19.4	hla	13.9%
$H-8$	83nt	29nt	12nt	16nt	-15.1	hla	9.3%
$S-1$	75nt	30nt	23nt	12nt	-21.3	SEA	18.3%
$S-2$	73nt	28nt	5nt	12nt	-22.8	SEA	null
$S-3$	74 nt	30nt	16nt	12nt	-20.3	SEA	15.6%
$S-4$	75nt	29nt	11nt	14 nt	-20.3	SEA	11.8%
$S-5$	75nt	30nt	23nt	0 _{nt}	-21.9	SEA	18.9%
$S-6$	76nt	30nt	23nt	14nt	-17.1	SEA	9.8%

Table S8. Summary of atsRNAs of *S. aureus* in this study

^a The full length of atsRNA. b The length of mRNA basepairing region of</sup> atsRNA. ^c The length of predicted complementary sequences between atsRNA and target mRNA. d The length of Hfq binding site of atsRNA. e The minimum free energy of atsRNA predicted at 37°C by MFOLD program. ^f The repression efficiency was calculated after the expression of each atsRNA for 30 min and bacterial cells containing empty vector pSB2035 was used as the control. The detailed sequences of these atsRNAs were listed in Supplementary Table S6 and 7.