3362–3375 *Nucleic Acids Research, 2010, Vol. 38, No. 10* *Published online 3 February 2010*
doi:10.1093/nar/gkq025

A novel antisense RNA regulates at transcriptional level the virulence gene *icsA* of *Shigella flexneri*

Mara Giangrossi[1], Gianni Prosseda[2], Chi Nhan Tran[1], Anna Brandi[1], Bianca Colonna[2,*] and Maurizio Falconi[1,*]

[1]Laboratory of Molecular Genetics, Department of Biology M.C.A., University of Camerino, 62032 Camerino (MC) and [2]Istituto Pasteur-Fondazione Cenci Bolognetti, Department of Cell and Developmental Biology, Sapienza University of Rome, 00185 Roma, Italy

Received September 26, 2009; Revised December 30, 2009; Accepted January 11, 2010

ABSTRACT

The virulence gene *icsA* of *Shigella flexneri* encodes an invasion protein crucial for host colonization by pathogenic bacteria. Within the intergenic region *virA-icsA,* we have discovered a new gene that encodes a non-translated antisense RNA (named RnaG), transcribed in *cis* on the complementary strand of *icsA. In vitro* transcription assays show that RnaG promotes premature termination of transcription of *icsA* mRNA. Transcriptional inhibition is also observed *in vivo* by monitoring the expression profile in *Shigella* by real-time polymerase chain reaction and when RnaG is provided *in trans.* Chemical and enzymatic probing of the leader region of *icsA* mRNA either free or bound to RnaG indicate that upon hetero-duplex formation an intrinsic terminator, leading to transcription block, is generated on the nascent *icsA* mRNA. Mutations in the hairpin structure of the proposed terminator impair the RnaG mediated-regulation of *icsA* transcription. This study represents the first evidence of transcriptional attenuation mechanism caused by a small RNA in Gram-negative bacteria. We also present data on the secondary structure of the antisense region of RnaG. In addition, alternatively silencing *icsA* and RnaG promoters, we find that transcription from the strong RnaG promoter reduces the activity of the weak convergent *icsA* promoter through the transcriptional interference regulation.

INTRODUCTION

Until recently, bacterial gene expression was believed to be regulated essentially by repressor or activator proteins acting mainly at the level of transcription initiation. At present, ~140 bacterial small (50–500 nt) regulatory RNAs, that do not encode proteins (ncRNAs), have been identified by means of systematic computer analysis, microarrays and cloning-based screening (1). Recently, a database collecting the small RNA (sRNA) genes has become available (2). sRNAs have been characterized not only in bacteria but also in phages, plasmids and in eukaryotic cells (microRNAs and interfering RNAs), suggesting that this level of regulation is widespread among living organisms (3). Although the function of many sRNAs remains to be elucidated, current studies indicate that they act by three general mechanisms. Few sRNAs are integral parts of the RNA–protein complexes as the 4.5S RNA component of the signal recognition particle and the RNase P RNA. A second class of regulatory sRNAs is represented by the so-called 'molecular decoys', which mimic the structure of other nucleic acids. Their target is usually a protein rather than another RNA (examples are the 6S RNA and the CsrB/CsrC RNAs). The non-coding RNAs of the third class are the best known and act by basepairing (RNA as antisense) with a second RNA, usually a messenger RNA, to change its behavior. Regulatory sRNAs can be encoded either *in cis* by the opposite strand of the target mRNA or by a free-standing gene located far from the target gene. Different RNA–RNA mediated mechanisms, entailing changes in processing and degradation of the target message and alterations in the efficiency of transcription and translation, have been described (3–7). It is becoming

*To whom correspondence should be addressed. Tel: +39 737 403265; Fax +39 737 40390; Email: maurizio.falconi@unicam.it
Correspondence may also be addressed to Bianca Colonna. Tel: +39 06-49917580; Fax +39 06-49917594; Email: bianca.colonna@uniromal.it

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

increasingly evident that sRNAs, besides gene regulation in general, can also play a key role in controlling the expression of virulence genes or can affect adaptive stress-responses, which are important for bacteria to survive into the host. Indeed, recent studies have identified sRNAs in human pathogens as *Listeria monocytogenes* (8,9), *Staphylococcus aureus* (10), *Pseudomonas aeruginosa* (11), *Salmonella* (12–14) and *Vibrio cholerae* (15–18). Interestingly, in *Escherichia coli*, the majority of sRNAs are present in pathogenic strains, suggesting that they might control virulence.

Shigella flexneri is a Gram-negative pathogenic bacterium that causes human bacillary dysentery. The *icsA* gene (named also *virG)* of *Shigella*, located on the 230-kb virulence plasmid pINV, (19), encodes a protein required for the invasion of intestine epithelial cells and intercellular spread of pathogens (20). IcsA, one of the biggest proteins (1102 amino acid residues) in bacteria, is an outer membrane protein, which induces host actin polymerization at one pole of the cell, resulting in actin-tail formation which propels the bacterium from one cell to another (21). As most plasmid virulence genes, the expression of *icsA* is modulated by temperature and repressed by the nucleoid protein H-NS (22–26). In contrast to other genes of the invasivity regulon, *icsA* is not submitted to the VirF–VirB regulatory cascade but seems to be activated only by the AraC-like transcriptional effector VirF (27,28).

In this study, we have identified and characterized the first regulatory RNA encoded by the virulence plasmid of *S. flexneri*. This small RNA (named RnaG), acting as antisense, is 450 nt long and is transcribed from the complementary strand of the target *icsA* mRNA. We show that two mechanisms contribute to the RnaG-mediated regulation of *icsA*: (i) *icsA* and RnaG promoters, sponsoring convergent transcription, are subjected to transcriptional interference (TI) regulation defined as the direct negative impact of a transcriptional process on a second transcriptional process *in cis* (29); and (ii) RnaG is capable to interact with the *icsA* mRNA and to cause premature termination of transcription of the target gene possibly by a transcriptional attenuation mechanism. Chemical and enzymatic RNA probing experiments show that the RNA–RNA interaction induces a change of the secondary structure of the leader region of the *icsA* transcript, promoting the formation of an alternative spatial conformation resembling a Rho-independent terminator. Mutations, which destabilize the conserved stem structure of this intrinsic terminator, significantly affect the ability of the RnaG to inhibit *icsA* transcription *in vitro*. The secondary structure of the antisense region (~120 nt) of RnaG was also determined by chemical probing. To our knowledge, this is the first strong evidence of such a mechanism in a human life-threatening bacterial pathogen.

MATERIALS AND METHODS

Bacterial strains and general procedures

Bacterial strains used in this work were *Shigella flexneri* M90T (20) and *E. coli* K-12 DH10b [*mcrA* Δ(*mrr-hsdRMS-mcrBC*) Δ*lacX74 deoR recA1*] (30), P90C [*ara D*(*lac-pro*) *thi*] (31) and HMG11 [*araD139, Δ*(*ara, leu*)*7697, Δ*(*lacZ*) *M15, galU, galK, strA*] (32). DNA isolation, agarose gel electrophoresis, polymerase chain reaction (PCR), restriction digestions, cloning and other DNA manipulation methods were performed according to standard procedures (30). Plasmid pKK232-8 was from Pharmacia. Beta-galactosidase assays were performed as described by Miller (33) on sodium dodecyl sulfate (SDS)–chloroform-permeabilized cells grown in LB broth (pH 7). Antibiotics and chemicals were used at the following concentrations: ampicillin 50 μg/ml; cloramphenicol 25 μg/ml; trimethoprim 10 μg/ml; Xgal 20 μg/ml. Radioactivity associated with DNA or RNA was detected and quantified by Molecular Imager (Bio-Rad, model FX). Oligonucleotides used in this study are listed in Table 1.

Table 1. Desoxyoligonucleotides used in this study

Name	Sequence
AH16	5′-GTCGAATTCCATGATGAGTC-3′
GZ25	5′-GTTATATACCACGGATCCATTC-3′
RM10	5′-CATGCACTGTGTTATCAGTAAG-3′
FM10	5′-CTTACTGATAACACAGTGCAT-3′
GRNdeI	5′-GAACATATGCTGTGCTTGCTTTGA-3′
GFNdeI	5′-CAGCATATGTTCTCCCGTTGCATTG-3′
GZ24	5′-GCAGATCTTGCATTGGAGAGTC-3′
G + 120H	5′-CCCAAGCTTGTGGTTGATAAACCCC-3′
G-383E	5′-CGGAATTCTCCTGATGTATTTTCC-3′
G + 1H	5′-CCCAAGCTTGTTGCATTGATATATAACAC-3′
G + 370E	5′-GAGGAATTCCATCAACAGGTGTTAATAA-3′
AC8	5′-TAGATTTCGGCAAGCGTACAG-3′
G-112	5′-TTCTTTCTGTACGCTTGC-3′
G-14E	5′-GGTGAATTCCTCCCG TTGCATTGATATA-3′
G-220	5′-CACAACTTTCCTACCAGACTCTCC-3′
G + 187	5′-CCTGTGAACATTGGGTCAT-3′
G + 59	5′-TTTGCATATATCGTCCCTTTATTCCG-3′
G + 50	5′-GGTTGAGGCTTTGTTTAATATG-3′
G + 110	5′-TTGATAAACCCCTGAAAAGA-3′
G-100	5′-GAAAGAACTGAAAAGTTGCGG-3′
ACC9	5′-CCTTTACCGTAGGTAATTCTC-3′
G-290	5′-ATTCTTCTCTTTAAGAAAACGATAC-3′
naF	5′-TGAAGCCGCACGTTATGAAG-3′
naR	5′-TCAACGTAATCGCCCAGGTT-3′
iaF	5′-TGATGGACTTTCTCCCTTGGG-3′
iaR	5′-TACCACGCATCCATTCCATCT-3′
rgF	5′-GCTTTCCCCCTCTTTTTTTCA-3′
rgR	5′-CAACGGGAGAAATTTACCTGTGTC-3′
G501	5′-ATATCGTCCCTTTATATGTGATAAAACGGAATCT-3′
G502	5′-AGATTCCGTTTTATCACATATAAAGGGACGATAT-3′

DNA manipulations

pGT1127 has been constructed by cloning into pGEMT-Easy a 866-bp fragment obtained upon PCR amplification with oligo pair AH16 and GZ25 using as template pMYSH6601, a pBR322-derived vector containing the *virA-icsA* genetic region (4.472 kb) of the *S. flexneri* 2a virulence plasmid pMYSH6000 (34). In order to study the intrinsic activity of P_{icsA} and P_{RnaG} promoters without the effect induced by the convergent promoter activity, we inactivated either P_{RnaG} or P_{icsA}, thus giving rise to $P_{icsA/RnaGmut}$ and $P_{icsAmut/RnaG}$. The $P_{icsA/RnaGmut}$ (silenced RnaG promoter) was obtained by

the three-round PCR approach using the wild-type *icsA-virA* regulatory region of pMYSH6601 as template. Oligos AH16 and RM10 were used as primers in the first reaction round (PCR-1), while oligos FM10 and GZ25 served as primers in PCR-2. Equimolar amounts of fragments from PCR-1 and PCR-2 were used as templates in PCR-3, primed by the oligos pair AH16 and GZ25. The 866-bp fragment thus obtained was cloned into pGEMT-Easy, giving rise to pGT1129. Sequencing confirmed the presence of P_{RnaG} modified in the –10 region (5′-TGTG TT-3′) (Figure 1B).

The $P_{icsAmut/RnaG}$ (silenced *icsA* promoter) was obtained by introducing a NdeI restriction site into the –10 consensus box. To this end, the two amplicons obtained from the *icsA-virA* region of pMYSH6601 with the oligos pairs AH16 and GRNdeI or GFNdeI and GZ25 were digested with NdeI and ligated to each other. The resulting 866-bp fragment was cloned into pGEMT-Easy to obtain pGT1083; sequence analysis confirmed the presence of a modified –10 region (5′-CATATG-3′) into the P_{icsA} promoter (Figure 1B).

Plasmid pKG673 has been constructed by cloning into the BamHI site of pKK232-8 a 673-bp DNA fragment (from positions –262 to +419) amplified by PCR using the primer pair GZ24 and GZ25 and pMYSH6601 DNA as template.

Plasmids pTZRnaG, pTZRnaG120 and pTZicsA370 were obtained by cloning, into the HindIII/EcoRI restriction sites of pTZ19R, under the control of the T7 promoter, the RnaG gene (from positions –383 and –14 to +120) and the leader region of *icsA* (from positions +1 to +370), respectively. These DNA fragments were obtained by PCR using pMYSH6601 as DNA template and the following primer pairs G+120H/G-383E, G+120H/G-14E and G+1H/G+370E. EcoRI linearized plasmids, pTZRnaG, pTZRnaG120 and pTZicsA370 were used as templates in *in vitro* transcription reactions with T7 RNA polymerase as described by Brandi *et al.* (35) to synthesize RnaG, RnaG120 and *icsA* transcripts.

Substitution of four bases on *icsA* sequence (from positions +81 to +84), to produce the *icsA81/4* mutant, was carried out by the QuikChange Site-Directed Mutagenesis Kit (Stratagene) using pGT1129 DNA as template and the mutagenic oligos G501 and G502. The resulting plasmid pGT1129M was used in *in vitro* transcription assay. The mutation was confirmed by DNA sequencing.

Construction of transcriptional fusions

Plasmids carrying fusions with the *lacZ* reporter gene were constructed by cloning different PCR-generated fragments into the multicloning site of the *lacZYA* transcriptional fusion vector pRS415 (33). Plasmid pULS1127, containing the P_{icsA}–*lacZ* fusion, was obtained by cloning a 866-bp EcoRI–BamHI fragment generated using AH16 and GZ25 as forward and reverse primers and pMYSH6601 as template. Plasmid pULS1287, containing the P_{RnaG}–*lacZ* fusion, was generated by cloning a 580-bp BamHI-BglII fragment obtained using GZ25 and AC8 as forward and reverse primers, respectively, and pMYSH6601 as template. Plasmid pULS1129, carrying the $P_{icsA/RnaGmut}$-*lacZ* fusion, and plasmid pULS1288, carrying the $P_{icsAmut/RnaG}$-*lacZ* fusion, were obtained with the oligo pairs used for the construction of the corresponding wt fusions but using plasmids pGT1129 and, respectively, pGT1083 as templates. The aforementioned *lacZ* fusions were then transferred by homologous recombination to the *lac* transducing phage λRS45, and then integrated (31) into the chromosome of *E. coli* P90C at the λ attachment site, thus generating strains ULS1127, ULS1129, ULS1287 and UL1288 (Figure 2B). Monolysogens were selected by means of a PCR test using primers corresponding to sequences flanking the bacterial and prophage attachment sites (36).

Primer extension analysis

After an initial denaturation of RNA samples at 80°C for 5 min, primer extension analysis was carried out at 42°C for 45 min in reaction mixtures (10 µl) containing the supplied buffer, 0.1 mM dNTP mix, 3 U of AMV reverse transcriptase (Roche) and 3 pmol of [γ^{-32}P]-oligo as primer. The reaction products were analyzed on 7% PAGE-urea gel in parallel with the dideoxy chain termination sequencing reaction using the same primer (30).

Mapping the 3′-end of the RnaG

Poly(A) tails were added to the 3′-termini of RNA using the *E. coli* enzyme poly(A) polymerase (GE Healthcare). The reaction (50 µl) was carried out at 37°C for 45 min in 40 mM Tris–HCl pH 7.7, 10 mM $MgCl_2$, 250 mM NaCl, 1 mM DTT, 50 µg/ml BSA, 2.5 mM $MnCl_2$, 250 µM ATP, 5 U of poly(A) polymerase and 15 µg of total RNA extracted from *E. coli* strain HMG9/pMYSH6601. The reaction was stopped at 75°C for 10 min in the presence of EDTA (f.c. 20 mM), RNA was precipitated and then dissolved in 15 µl of water. Aliquots of poly(A)-tailed RNA were used to synthesize a cDNA copy. RNA was denaturated at 65°C for 5 min and subsequently the reaction (10 µl) was carried out at 42°C for 45 min in 50 mM Tris–HCl pH 8.5, 8 mM $MgCl_2$, 30 mM KCl, 1 mM DTT, 150 µM dNTPs, 60 pmol anchored oligo($dT_{19}N)_{20}$ and 15 U of AMV reverse transcriptase (Roche). The resulting cDNAs were used as templates in PCR reaction performed with 130 pmol of the anchored oligo(dT_{14}VN) and 30 pmol of the *icsA* specific oligo G-112 as primers. The amplicon obtained after 40 PCR cycles was subjected to sequencing using the oligo G-220 as primer.

In vitro transcription assay

In vitro transcription from supercoiled pGT1127, pGT1129 and pGT1083 as DNA templates was carried out at the indicated temperature for 45 min. Each reaction mixture (40 µl) contained 40 mM Tris–HCl, pH 7.5, 150 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.01% Triton X-100, 0.5 mM each of NTPs, 2 U of ribonuclease inhibitor and 0.2 U of *E. coli* RNA polymerase (USB). The reaction was stopped on ice and RNA was precipitated with ethanol in the presence of 1-µg

tRNA as carrier. Concerning transcription assays shown in Figure 7, RnaG was added to the mixture before RNA polymerase (samples B) or at the step of RNA precipitation (samples E). This protocol ensures that the same amount of RnaG is present in the couple B/E during the following elongation step by the reverse transcriptase. The transcripts were subjected to primer extension as described above. Alternatively, *in vitro* transcription was performed incorporating [α^{-32}P]-UTP in the *de novo* synthesized RNA. DNA template (10–20 ng) were transcribed in a volume of 15 µl in the same conditions indicated above in the presence of 200 µM CTP, 10 µM UTP, 1 mM ATP, 200 µM GTP and 0.2 µCi/µl [α^{-32}P]-UTP (3000 Ci/mmol). After an incubation of 30 min at 37°C, the reaction was stopped by adding an equal volume of a solution containing 50% formamide and 10 mM EDTA and heated at 90°C for 2 min.

Chemical and enzymatic RNA probing

Chemical modification of RNA was performed as described previously (37) using the single-strand specific reagents DMS (A and C specific) and CMCT (U and G specific).The RNA was incubated for 5 min in 20 µl of buffer A (50 mM Na-cacodylate, pH 7.5, 5 mM $MgCl_2$, 100 mM KCl) or 20 min in 20 µl of buffer B (50 mM Na-borate, pH 8, 10 mM $MgCl_2$, 50 mM KCl) at 32°C in the presence of the indicated concentrations of DMS or CMCT, respectively. Control samples were treated identically with the exception that no modifying reagents were added. The modified RNA was subjected to primer extension as described above.

To perform enzymatic probing, purified *icsA*370 mRNA was dephosphorylated using calf intestinal alkaline phosphatase (Amersham) and labeled with T4 polynucleotide kinase (USB) and [γ-^{32}P]-ATP (30).The 5′-end-labeled RNA was additionally extracted from gel (PAGE). After denaturation at 90°C for 1 min and renaturation for 10 min at 32°C, RNA was treated with RNases T1 or T2 (5 min) at the indicated concentrations in 10 µl buffer A (20 mM HEPES–KOH, pH 7.5, 10 mM $MgCl_2$, 50 mM KCl) in the presence of 1 µg of carrier tRNA. The reaction products were analyzed on 10% PAGE-urea gel in parallel with ΔT1 and OH– ladders (38).

Real-time quantitative PCR

Total RNA was extracted as described by Von Gabain *et al*. (39) and cDNA synthesis was performed using the High Capacity cDNA Reverse Transcription Kit from Applied Biosystems. The 20-µl reaction mix contained 20 µg total RNA from *Shigella* M90T. Real-time quantitative PCR was performed on a 7300 Real-Time PCR System (Applied Biosystems) in a 30-µl reaction mix containing 5 µl cDNA and Power SYBR®Green PCR Master Mix (Applied Biosystems). At least three wells were run for each sample. The relative amounts of *icsA* and RnaG transcripts were analyzed using the $2^{-\Delta\Delta Ct}$ method (40) and the results were indicated as an *n*-fold increase relative to the starting sample, which was chosen as reference. Primers for the *nusA* transcript, used as endogenous control, and for the above-mentioned transcripts were designed with the aid of the Primer Express® software v2.0 (Applied Biosystems) and experimentally validated for suitability to the $2^{-\Delta\Delta Ct}$ method. The following oligos were used: naF and naR for the *nusA* transcript; iaF and iaR for *icsA*; rgF and rgR for RnaG.

RESULTS

Identification and molecular characterization of the RnaG

While cloning different DNA fragments carrying the promoter and part of the coding region of *icsA* gene of *S. flexneri* into the vector pKK232-8, we observed that the expression of the promoter-less reporter gene *cat* was detected independently of the correct orientation of the *icsA* promoter (P_{icsA}). This suggested the existence of at least two convergent promoters, transcribed on different DNA strands, within the *icsA* DNA sequence. In agreement with this finding, primer extension analysis reveals the presence of a promoter, named P_{RnaG}, with reverse orientation relative to *icsA* (Figure 1A). Transcripts originated with two guanines at position +120 and, albeit to a much lesser extent, at position +118 with respect to the transcriptional start site of *icsA* (+1) previously identified by Lett *et al*. (34). The consensus hexamers –10 and –35 and the transcriptional start sites of P_{icsA} and P_{RnaG} are shown in Figure 1B. In addition, we prepared a set of different constructs in which the promoters of *icsA* and RnaG were alternatively inactivated by introducing mutations in the corresponding –10 elements to produce plasmids pGT1127, pGT1129 and pGT1083, schematically shown in Figure 1C. Promoter functioning was tested by primer extension analysis, carried out on total RNA extracted from *E. coli* cells transformed with the pGT plasmids. As expected, Figure 1D reveals that base exchanges at P_{RnaG} and P_{icsA} dramatically impair RnaG and *icsA* promoter activity, respectively. It is remarkable that the level of *icsA* mRNA is significantly higher when the RnaG promoter is inactivated (pGT1129) than in the wt condition (pGT1127). Moreover, transcription of *icsA* is increased at 37 as compared to 30°C. Although *in silico* analysis of the DNA sequence indicates the existence of a short open reading frame (31 amino acid residues) in the *virA-icsA* intergenic region, no β-galactosidase activity could be detected from translational fusions of this DNA region with the reporter gene *lacZ* (data not shown). Altogether, these observations strongly suggest the presence of a non-coding antisense RNA (RnaG) encoded in *cis* on the opposite strand of *icsA* and complementary to the first 120 nt of the mRNA.

The full length of RnaG has been determined by mapping the 3′ end of the transcript. Since bacteria do not usually show poly(A)+ RNA, a tail of adenines was incorporated *in vitro* to the 3′ end of total RNA using the poly(A) polymerase. Subsequently, a cDNA copy was made using the anchored oligo($dT_{19}N$) as primer. The DNA was then subjected to PCR amplification combining the anchored oligo($dT_{14}VN$) and a specific one designed

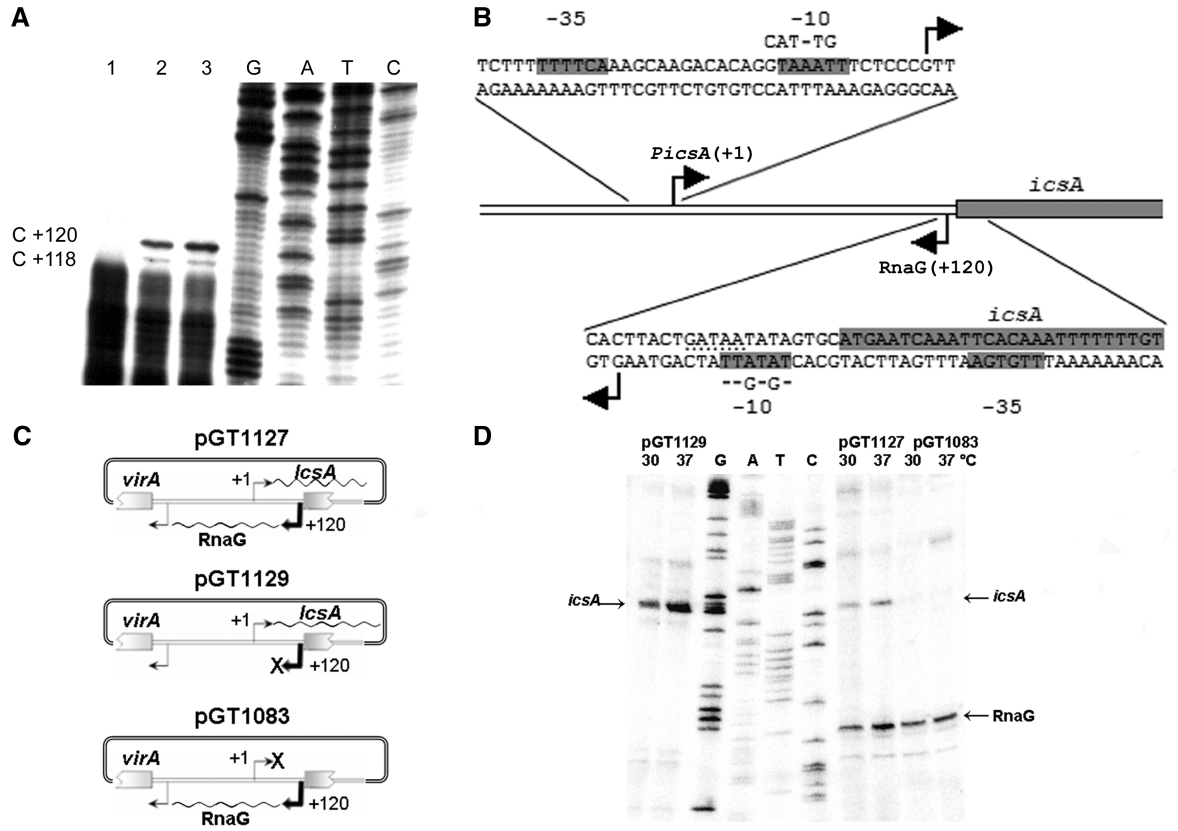

Figure 1. Identification of the RnaG promoter. (**A**) Primer extension analysis was carried out using the oligo G + 59 on 10 μg of total RNA from the *E. coli* wt strain HMG11, a MC1029 derivative, transformed with plasmid pMYSH6601 (lane 2) or pKG673 (lane 3). Lane 1 is the control sample in the absence of RNA. Lanes G, A, T and C represent the dideoxy sequencing reactions carried out with the same primer. The positions of the two transcriptional start sites are indicated. Smears below the specific signals are due to the low quality of some commercial oligonucleotides. (**B**) The relevant nucleotide sequence of the regulatory region of the *icsA* gene. The first nine coding triplets of *icsA*, the transcriptional start point, the –10 and –35 consensus sequences of the *icsA* and RnaG promoters are indicated. Mutated TATA boxes created in P_{icsA} and P_{RnaG} to construct plasmids pGT1083 and pGT1129, respectively, are reported above and below the sequence. (**C**) Schematic representation of plasmids pGT1127 ($P_{icsA/RnaG}$), pGT1129 ($P_{icsA/RnaGmut}$), pGT1083 ($P_{icsAmut/RnaG}$). (**D**) Primer extension analysis was performed on 15 μg of total RNA from the *E. coli* strain HMG11 transformed with plasmids pGT1127, pGT1129 or pGT1083 using a mixture of primers G + 110 and G + 50 to detect *icsA* mRNA and RnaG, respectively. Lanes G, A, T and C represent the sequencing reactions using the primer G + 110.

on *icsA* as primers. Finally, the amplification product was sequenced (Supplementary Figure S1AB). The nucleotide sequence indicates that transcription possibly terminates between cytosine at position –332 and adenine at position –334 so that RnaG would be ~450 nt long. Because of the presence of two A_s at the 3′-end of the transcript, that are not distinguishable from those incorporated during the poly(A) tail synthesis, we could not identify the transcription termination site with single base accuracy. Northern blot analysis confirms the estimated size of RnaG (Supplementary Figure S1C). In agreement with the experimental results, computer prediction of secondary structure of the 3′-end of RnaG shows that it might form two hairpin loops characteristic of 'Tandem/U shaped' intrinsic terminators (41) (Supplementary Figure S1D). It is important to notice that the same transcriptional termination site was also found when RnaG was transcribed *in vitro*, using the wt plasmid pGT1127 as template (data not shown), suggesting that no other factors contribute to RNA polymerase pausing and leading to termination of transcription.

The *icsA* and RnaG promoters are subjected to TI regulation

The *icsA* and RnaG promoters are convergent and 120 bp apart. This arrangement can possibly give rise to TI, which refers to a direct negative influence of one transcriptional process on a second transcriptional process occurring in *cis*. Usually the strong promoter (aggressive) reduces the activity of the weaker convergent promoter (sensitive) (29). To assess whether TI plays a role in the regulation of *icsA* and eventually to establish which promoter is the sensitive or the aggressive one, we used plasmids pGT1127, pGT1129 and pGT1083 (Figure 1C). These constructs were transcribed *in vitro* and the *icsA* and RnaG transcripts were detected by primer extension. As shown in Figure 2A, when the wt construct pGT1127 is used as template only the RnaG transcript is visible and no transcription originates from P_{icsA}. The *icsA* mRNA becomes evident when pGT1129, lacking an active P_{RnaG}, is provided as template. This observation suggests that transcription from the strong

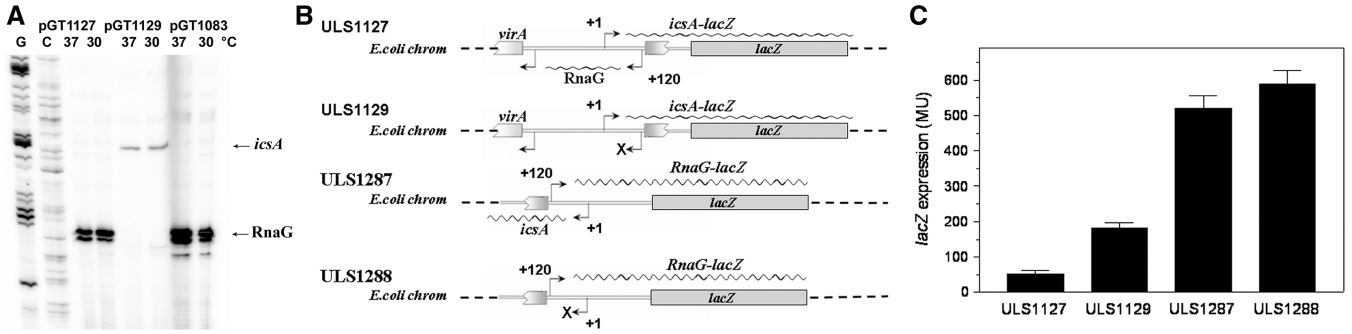

Figure 2. Transcriptional Interference regulates *icsA* and RnaG promoter activity. (**A**) About 200 ng of supercoiled plasmids pGT1127 ($P_{icsA/RnaG}$), pGT1129 ($P_{icsA/RnaGmut}$) and pGT1083 ($P_{icsAmut/RnaG}$), schematically represented in Figure 1C, were used as templates in *in vitro* transcription assays carried out at 30 and 37°C. RnaG and *icsA* transcripts were detected by primer extension using a mixture of oligos G+50 and G+110. Lanes marked G and C are the dideoxy sequencing reactions performed on pGT1127 with the oligo G+110 as primer. (**B**) Schematic organization of the *E. coli* chromosome (strain P90C) carrying the different ULS *lacZ* fusions. (**C**) β-Galactosidase activity of ULS transcriptional fusions was determined on strains grown in LB broth (A_{600} = 0.4) at 37°C. Values represent the average of four independent experiments and the standard deviation is indicated.

RnaG promoter (aggressive) dramatically inhibits transcription from the weaker *icsA* promoter (sensitive). The interference, calculated as ratio of RnaG (pGT1083) and *icsA* (pGT1129) promoter activity, is ~12-fold. Moreover, TI is reciprocal since also transcription of P_{icsA} negatively affects, albeit at low extent, that of the P_{RnaG}. In fact, the level of the RnaG transcript is ~1.5-fold higher in pGT1083 carrying an impaired *icsA* promoter, than in the wt construct pGT1127.

TI between the P_{icsA} and P_{RnaG} has been further analyzed *in vivo* by constructing a series of transcriptional fusions carrying the *lacZ* reporter gene under the control of either *icsA* and RnaG promoters (Figure 2B). P_{icsA} and P_{RnaG} transcription was assayed using a λ-based single-copy chromosomal *lacZ* operon fusion system. As shown in Figure 2C, inactivation of P_{RnaG} causes a ~3.5-fold increase of the expression of the ULS1129 fusion as compared to that of the wt construct ULS1127, thus confirming that RnaG transcription negatively interferes with P_{icsA} transcription. On the other hand inactivation of P_{icsA} gives rise only to a slightly enhanced β-gal level from P_{RnaG} (compare ULS1288 and ULS1287). By measuring the levels of β-gal expression of ULS1288 and ULS1129 fusions, which do not account for TI effects, it appears evident that P_{RnaG} is endowed with ~3-fold stronger transcriptional activity than P_{icsA}. Altogether, these results clearly indicate that TI contributes to the coordinate regulation of this genetic system.

RnaG downregulates *icsA* by transcription attenuation

To investigate on the possible effect of RnaG on *icsA* transcription, we cloned the DNA sequence coding for this antisense RNA into pTZ19R. The recombinant plasmid (pTZRnaG) was used to synthesize RnaG by *in vitro* run-off transcription with T7 RNA polymerase. After purification, RnaG was added to an *in vitro* transcription assay programmed with a 331-bp DNA fragment carrying the *icsA* promoter and an impaired P_{RnaG}. As seen in Figure 3A, the band representing the full-length run-off transcript of *icsA* (F) suddenly disappears even in the presence of low amounts of RnaG (200 fmol). Concomitantly, a shorter product of ~100 nt (T) is formed and its level progressively increases, becoming the only transcript detectable at higher RnaG concentrations. This result suggests that the antisense RNA may promote transcription termination of the target gene by a transcriptional attenuation mechanism. Such hypothesis is strengthened by the lack of the truncated RNA molecule (T) when the *icsA81/4* mRNA, carrying mutations in the potential intrinsic terminator (see below), is transcribed (Figure 3B). Under the same experimental conditions, the promoter activity of a control gene (*hns*) is not affected by RnaG (Figure 3C).

The role played by RnaG in the modulation of *icsA* transcription has been further investigated *in vivo* by introducing extra copies of RnaG into ULS1129, a λR45 mono lysogen strain carrying a P_{icsA}–*lacZ* fusion with an inactivated RnaG promoter (Figure 2B). As shown in Figure 4A, synthesis of RnaG from pGT1127 or pGT1083 plasmids impairs the full expression of the chromosomal P_{icsA}, giving rise to a ~3-fold reduction of β-gal level, thus confirming that, when provided *in trans,* RnaG is able to inhibit *icsA* transcription. This result depends on the presence of the RnaG molecule, since no repression is observed transforming the cells with the pGT1129 which carries a silenced P_{RnaG}. By means of a real-time quantitative PCR assay (QPCR), we also compared the level of *icsA* mRNA in *S. flexneri* M90T strain carrying either pGT1083 or the pGEMT cloning vector. The presence of RnaG encoding plasmid (pGT1083) gives rise to a 40% reduction of *icsA* mRNA (data not shown) in agreement with the result obtained in ULS1129 background. Moreover, to support the control effect of RnaG on *icsA* transcription, we decided to monitor the relative expression of *icsA* and RnaG in *S. flexneri* M90T strain throughout the growth curve by a QPCR assay. As shown in Figure 4B, during the exponential growth, RnaG expression remains fairly constant, while *icsA* expression shows a progressive increase. Interestingly, approaching the stationary phase ($OD_{600\,nm}$ = 1) RnaG expression

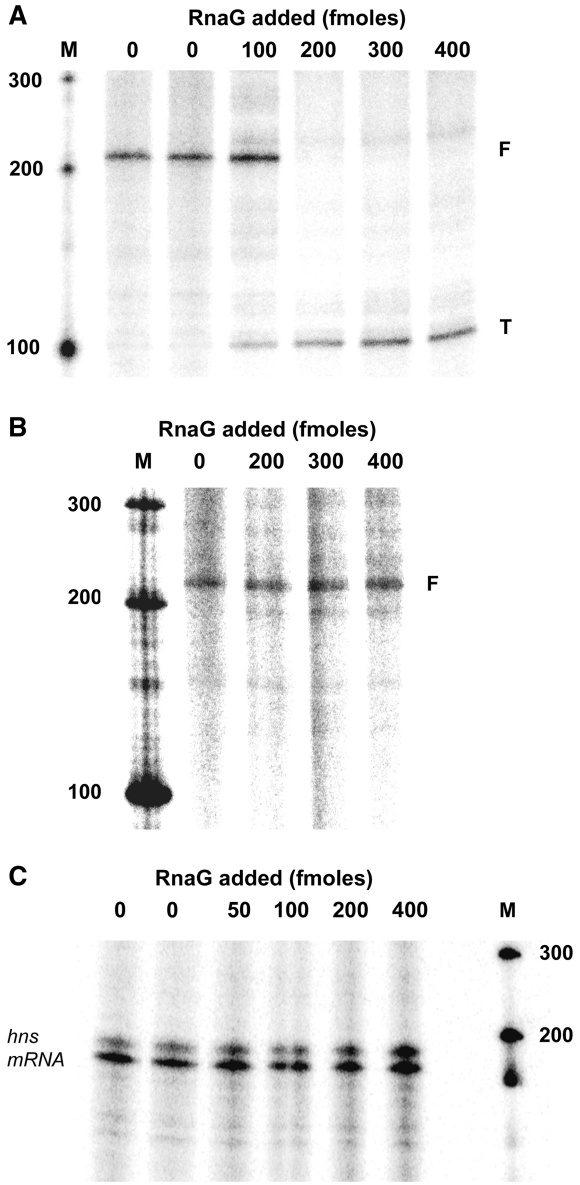

Figure 3. The RnaG downregulates *icsA* transcription. Transcription was investigated *in vitro* as function of increasing amounts of purified RnaG using as template a 331-bp DNA fragment (from position –117 to +214), corresponding to the wt *icsA* promoter (**A**), the same DNA fragment carrying the mutation 81/4 (see text and Figure 7A) (**B**) and a 480-bp DNA fragment (from position –314 to +166) corresponding to the *E. coli hns* gene as control (57) (**C**). The wt and mutated *icsA* DNA fragments were generated by PCR amplification using the primer pair G-100/ACC9 and plasmids pGT1129 and pGT1129M as templates, respectively. F and T indicate full-length run-off and terminated *icsA* transcripts, respectively. Transcripts were labeled by incorporating [α^{-32}P]-UTP and lanes marked with M are different RNA ladders.

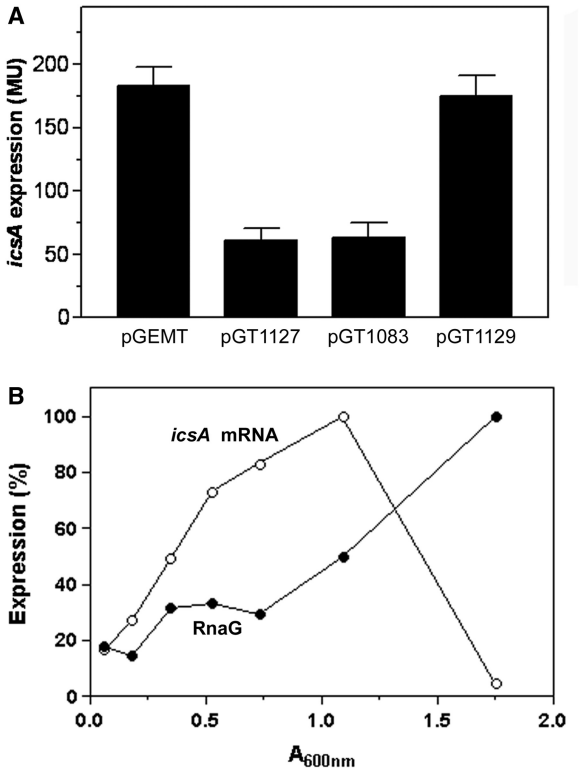

Figure 4. RnaG negatively affects *icsA* expression. (**A**) Expression of *icsA-lacZ* fusion USL1129 was monitored in cells transformed with pGT1083, pGT1127, pGT1129 or pGEMT vectors. (**B**) The *in vivo* level of *icsA* and RnaG transcripts was monitored during the growth of *Shigella* by real-time PCR. The expression trend of either template, normalized to the corresponding maximum (100%), is shown.

shows a sharp increase immediately followed by an abrupt decrease of *icsA* expression, thus supporting the hypothesis that RnaG hampers the full expression of *icsA* gene.

Convergent transcription from face-to-face promoters, as in the case of this genetic system, produces two transcripts that partially basepair, possibly originating a RNA duplex. The potential interaction between *icsA* and RnaG transcripts was monitored by primer extension analysis performed, in the presence or in the absence of purified RnaG, on bulk RNA extracted from an *E. coli* HMG11 strain harboring the pGT1129, which synthesizes only the *icsA* mRNA. We find that, in the absence of RnaG, the 5′ terminus matches with the previously identified transcriptional start point of *icsA* (34). By contrast, upon adding RnaG, the canonical start point of *icsA* is apparently moved from position +1 to position +108 (Figure 5A). Assuming that the 5′ leader region of the *icsA* mRNA interacting with the RnaG yields a peculiar secondary structure, which prevents the elongation of the cDNA by the enzyme reverse transcriptase, we investigated on the secondary structure of the *icsA* transcript by RNA probing. Initially, the purified 5′ terminus (~370 nt) of the *icsA* mRNA was treated with the two single-strand specific reagents dimethyl sulfate (DMS modifies unpaired adenines and cytosines) and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-p-toluene sulfonate (CMCT modifies unpaired uridines and guanines) (42). Data obtained by chemical probing (Figure 5B) were superimposed on a computer prediction generated by the MFOLD program (43). A model of the structural organization of the first 145 nt of free *icsA* mRNA is presented in Figure 5E. This RNA, particularly its 5′-end, forms a pronounced secondary structure which is composed of a very long hairpin motif (AH1) and a second helix containing both an apical and an internal loop (AH2). Conversely, the 3′-end of the *icsA*

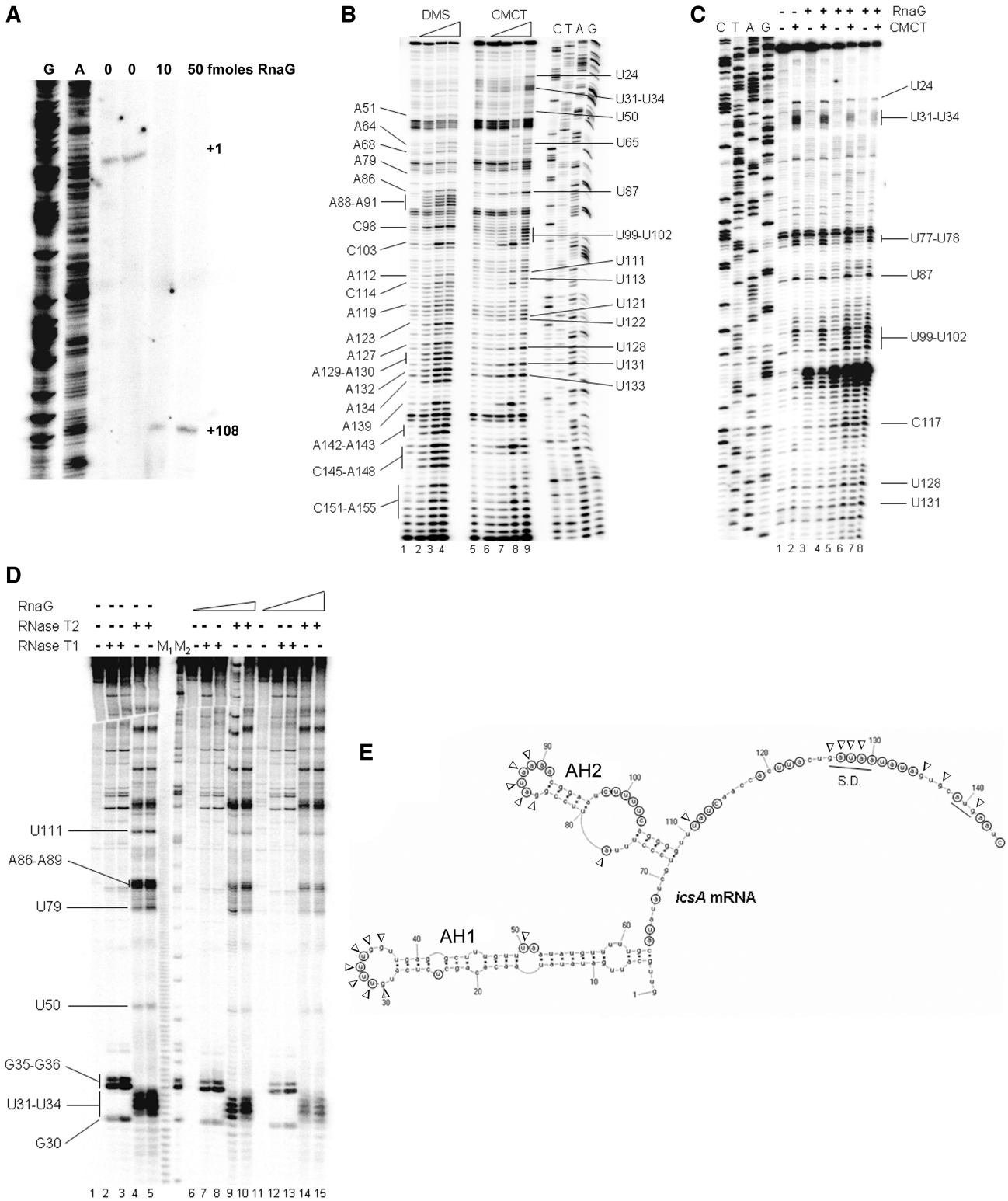

Figure 5. RNA probing of the *icsA* mRNA leader region either alone or in combination with RnaG. (**A**) Total RNA (15 µg) was primer-extended in the absence or in the presence of the indicated amounts of purified RnaG using the oligo ACC9. (**B**) *In vitro* transcribed *icsA* RNA (2 pmol) was treated with increasing amounts of DMS (lanes 1, 0%; lane 2, 0.3%; lanes 3, 0.6%; lane 4, 1.2%) and CMCT (lanes 5, 0 mg/ml; lane 6, 0.25 mg/ml; lane 7, 0.5 mg/ml; lane 8, 1.0 mg/ml). (**C**) The *icsA* mRNA (2 pmol) was treated with CMCT (3 mg/ml) either without (lanes 1 and 2) or with different amounts of RnaG: 1 pmol (lanes 3 and 4), 2 pmol (lanes 5 and 6) and 4 pmol (lanes 7 and 8). Relatively to both (B) and (C), modified nucleotides were detected by primer extension using the oligo G + 187 and accessible sites were evaluated comparing samples incubated in the absence (–) and in the presence of either DMS or CMCT. Lanes C, T, A and G correspond to DNA sequencing ladders made with the same primer. (**D**) About 300 fmol of [^{32}P]-end-labeled *icsA* mRNA were digested with RNases T1 (2.5×10^{-2} U, lanes 2, 7 and 12; 5×10^{-2} U, lanes 3, 8 and 13) or T2 (1.1×10^{-3} U, lanes 4, 9 and 14; 2.2×10^{-3} U, lanes 5, 10 and 15) in the absence (lanes 1–5) and in the presence of 290 (lanes 6–10) and 580 fmol (lanes 11–15) of RnaG as indicated. M_1 and M_2 represent the OH– and ΔT1 ladders, respectively. (**E**) Schematic representation of the secondary structure of *icsA* mRNA (nucleotides from +1 to +145). The AH1 and AH2 motifs, nucleotides reactive to DMS or CMCT (circles), cleavage points by T1 and T2 RNases (triangles), the initiation triplet and the S.D. sequence are indicated.

mRNA, downstream residue 110, is highly accessible to chemical modification, indicating that this region of the molecule is single stranded. Next, we probed the structure of the *icsA* mRNA together with the RnaG. As seen in Figure 5C and Supplementary Figure S3, in the presence of RnaG, an extended region (~80 nt) of the 5′ terminus of the *icsA* mRNA is no longer exposed to CMCT and DMS modification, particularly the unpaired nucleotides forming the apical loop (U31–U34, A28), the uridine 24 and the adenine 51 of AH1. This strongly suggests that an RnaG–*icsA* mRNA duplex is formed. Differently, the non-folded intervening sequence (A68, C70, U72) between AH1 and AH2, the apical loop (A86–A91) and the bulged structure of AH2 (U77–A79 and A96–U102) remain still accessible to the modifying agents even adding a 2-fold excess of RnaG (Figures 5C and S3, lanes 7 and 8), indicating that nucleotides 80–120, although complementary to RnaG, are less important for RNA pairing. The *icsA* mRNA structure obtained by chemicals was confirmed by enzymatic RNA probing with T1 and T2 ribonucleases (Figure 5D). This technique was also used to analyzed the structure of *icsA* mRNA in the presence of the RnaG. Depending on the amount of RnaG added, cleavages produced by these two single-stranded RNA specific enzymes at level of motifs AH1 (G30–G36, U50) and AH2 (U79, A86–A89), on *icsA* transcript, are clearly reduced or even hardly detectable, while no difference is observed with or without RnaG downstream U111. RNA probing experiments combined with MFOLD predictions and the *in vitro* synthesis of a incompletely extended *icsA* mRNA (Figure 3A) suggest that the progressive complementation between the nascent transcript and the RnaG hinders proper folding of helices AH1 and AH2 (at least in part), thereby inducing structural changes in the leader region of sense RNA. Such an alternative organization of the *icsA* mRNA has the potential to form a stem–loop structure, characteristic of a Rho-independent terminator (Figure 7A), which is likely responsible to promote premature termination of the nascent *icsA* transcript by an antisense RNA-mediated transcriptional attenuation mechanism (44). The proposed model is shown in Figure 6.

To verify the validity of our model, we mutagenized the pGT1129 creating the pGT1129M plasmid which carries a four-base substitution at positions 81–84 on *icsA* and encodes the mutated *icsA81/4* mRNA (Figure 7A). This mutation was designed to destabilize the stem structure of the intrinsic terminator possibly formed, during the *icsA* mRNA tramscription, upon interaction with the antisense RnaG. As previously shown in Figure 3B, transcription of the *icsA81/4* mRNA was not able to originate the truncated product in the presence of RnaG. Thus, we further analyzed the effect of this mutation by means of a technically different *in vitro* transcription assay programmed with supercoiled plasmid DNA. The *icsA* transcript was monitored by primer extension using a [^{32}P]-labeled oligo recognizing only untruncated transcripts. Since we found that the binding of RnaG to *icsA* mRNA can interfere with elongation by reverse transcritpase (Figure 5A), RnaG was added either at beginning (B) or at the end (E) of

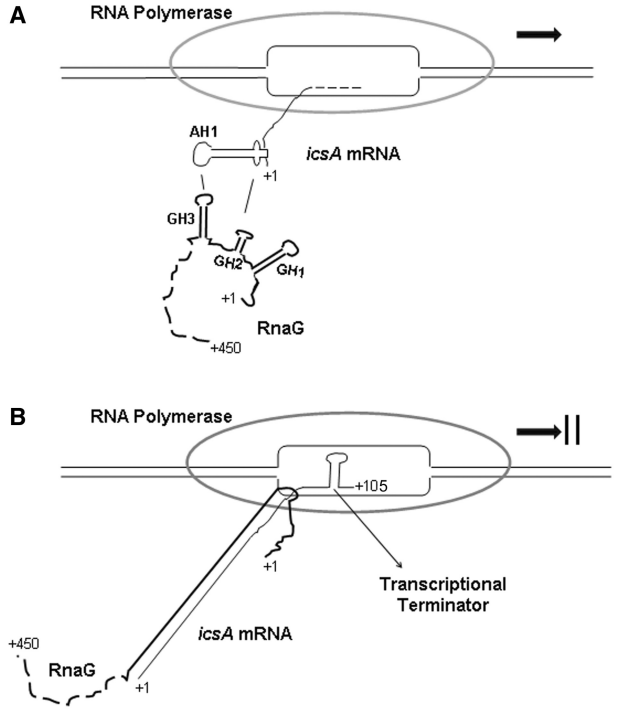

Figure 6. Model showing the RnaG-mediated transcriptional attenuation mechanism. (**A**) The interactions of GH3 and GH2 with AH1 possibly provide the initial nucleation points leading to duplex formation between RnaG and *icsA* mRNA. The GH1 hairpin is not represented contacting AH2 because at this stage the latter motif is not completely transcribed. (**B**) RNA–RNA pairing causes modifications of the secondary structure of the *icsA* mRNA resulting in the formation of a potential intrinsic terminator between nucleotides +78 and +105 (see also Figures 7A and 8).

the *in vitro* transcription reaction. We devised this protocol to properly discriminate the negative effects, caused by RnaG, on transcription from those on the following elongation step. Therefore, this procedure has an internal control. As seen in Figure 7B, RnaG mostly represses transcription of the wt *icsA* (lanes B) without significantly affecting the primer extension detection (lanes E). Transcription inhibition is also observed using a shorter form of RnaG, named RnaG120, containing only the first 120 nt, suggesting that the antisense region plays a key role in controlling *icsA* synthesis (Supplementary Figure S2). The study of deletion mutants of RnaG is actually in progress. On the contrary, under the same experimental conditions, RnaG is not able to cause transcription termination (except at the highest amount tested), when the mutated *icsA81/4* mRNA is synthesized (Figure 7C). The different response of *icsA* and *icsA81/4* mRNAs to the RnaG-mediated control is clearly evidenced when transcription is expressed as the ratio B/E (Figure 7D). These results strengthen the proposed model based on transcriptional attenuation mechanism. Moreover, we analyzed the secondary structure of RnaG and actually the antisense region (nucleotides 1–120) has been elucidated. As seen in Figure 8, the 5′-end of RnaG is characterized by three stem–loop motifs indicated as GH1–GH3 that might play a role in the initial pairing with the *icsA* mRNA.

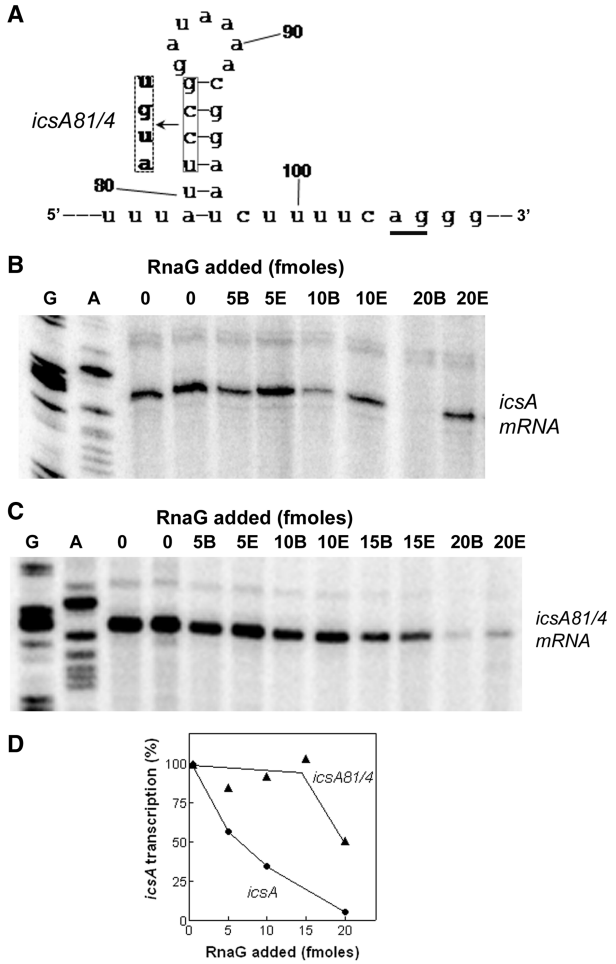

Figure 7. The mutation *icsA81/4* abolishes the RnaG-mediated transcriptional termination. (**A**) The hairpin followed by a polyU sequence typical of Rho-independent transcription terminators and base exchanges to create the *icsA*81/4 mutated mRNA are shown. The underlined nucleotides represent the putative 3′ terminus (positions +104 and +105) of the abortive *icsA* mRNA predicted according to the TransTerm algorithm (58). I*n vitro* transcription was carried out at 37°C on 300 ng of supercoiled plasmid DNA using pGT1129 (**B**) and pGT1129M (**C**) as templates. The *icsA* and *icsA81/4* mRNAs were detected by primer extension using the oligo G+110 (from position +97 to +116). Lanes G and A represent the sequencing reactions using the same primer. After imager quantization, the radioactivity associated with mRNA in (B) and (C) was used to calculate the ratio B/E at the same RnaG concentration. These values have been plotted as percentage (**D**) assuming 100% transcription in the absence of RnaG.

DISCUSSION

During the past decade, small non-coding RNAs have been discovered in all organisms. A major role of these ncRNAs, acting by basepairing to their target mRNAs, is primarily to regulate translation and messengers stability (3,6). In bacteria, antisense RNA were first detected on extrachromosomal elements as plasmids, transposons and bacteriophages (45); however, now it is becoming increasingly clear that most of the chromosomally encoded small RNAs (sRNAs) are part of regulatory circuits required for fast adaptive response to stress and environmental changes (46). The invasion of a host by a pathogenic microorganism, which results in reprogramming of the transcriptional activity, implies a prompt adaptation to new growth conditions related to the transition from a free living to a host-associated state. In this context, 19 novel island-encoded sRNAs have been recently identified in *Salmonella* and several of these genes are expressed under stress conditions when bacteria reside within macrophages (13). Thus, RNA–RNA interactions appear decisive for virulence.

In the present study, we have identified a novel gene encoding an antisense RNA (RnaG), located within the *icsA* sequence of the *S. flexneri* virulence plasmid (pINV). The *icsA* gene (also known as *virG*) encodes a structural protein involved in the reorganization of the cytoskeleton upon penetration of bacteria into host cells and its activity is crucial for the intra- and inter-cellular spreading of the bacterial pathogen (20). After mapping the 5′- and 3′- ends of RnaG, we focused on its ability to regulate the expression of *icsA*. Recent studies have shown that many sRNAs, acting as antisense RNA, modulate the synthesis of outer membrane proteins, thereby controlling the surface composition of Gram-negative bacteria (12,47,48). IcsA is indeed an outer membrane protein and RnaG represents a novel example of this type of regulation. As opposed to other sRNAs and specifically to those which control membrane proteins expression, affecting mostly the ability of target mRNAs to be translated, we found that the RnaG is capable to downregulate *icsA* transcription through dual not mutually exclusive mechanisms.

We have shown that *icsA* and RnaG promoters are subjected to TI regulation. TI provides an additional platform for gene regulation and it has been described in phages, transposable elements, plasmids, bacteria, yeast but less frequently in higher eukaryotes. The main mechanisms, by which TI can occur, are: (i) occlusion in which the passage of elongating RNA polymerases (RNAPs) blocks the access to the promoter; (ii) collisions between elongating RNAPs, moving in opposite directions, leading the premature termination of transcription; and (iii) 'sitting duck' interference which refers to the removal of promoter-bound complexes by the passage of RNAP from the opposing promoter (29). We have found that the activity of P_{icsA} is significantly reduced when a transcription process starts from the convergent P_{RnaG}, indicating that this phenomenon plays a key role in the regulation of *icsA* (Figure 2). The extent of TI is higher *in vitro* (~12-fold) than *in vivo* (~3–4-fold), suggesting that other factors (i.e. regulatory proteins), and, not simply the promoter strength, could contribute to the regulation of this genetic system into bacterial cells. Indeed, *E. coli* strains used in this assay express the chromosomal protein H-NS acting as *icsA* repressor, but not its natural activator VirF (see below) carried by the large pINV virulence plasmid. At present, our results do not allow us to figure out which type of the three aforementioned mechanisms is likely to be the major contributor to the overall TI and this will be the aim of further investigations. To our knowledge, we believe that RnaG is the first example of a non-coding RNA, which negatively affects transcription of the target gene by TI mechanism.

In addition to TI inhibition, RnaG can also cause a direct repression of *icsA* transcription acting primarily as antisense RNA by targeting the *icsA* messenger. Such a negative effect on *icsA* promoter activity is observed both in *S. flexneri* and in *E. coli* and when the RnaG is provided *in trans*. In particular, the analysis of the expression trend of RnaG during the growth of *S. flexneri* reveals that a marked RnaG increase toward the onset of the stationary phase is followed by a rapid reduction of the *icsA* transcript (Figure 4). *In vitro* transcription assays programmed with a linear DNA fragment containing P_{icsA} display, solely in the presence of the RnaG, the appearance of an abortive product of ~100 nt in place of the full-length transcript. The truncated molecule is not formed when the mutation 81/4, abolishing the function of the transcriptional terminator, is introduced in the *icsA* mRNA (Figure 3). Additionally, transcription termination mediated by RnaG is observed also when a supercoiled plasmid, carrying *icsA*, is used as template. In fact, the shortened *icsA* mRNA can not be primer-extended with the G+110 oligo, which fails to pair with its target sequence on *icsA* mRNA, resulting in the reduction and/or disappearance of the signal as function of RnaG concentration (Figure 7). Although, RnaG binds and exerts its inhibitory effect apparently without helper proteins, Hfq and/or other chaperonins (49) might influence the *in vivo* RnaG-dependent regulation of *icsA*. The inhibition of P_{icsA} is highly specific since it is completely abolished in cells not expressing RnaG due to inactivation of promoter and transcription of the control gene *hns* is not affected *in vitro* even by adding RnaG at elevated concentrations.

Combining computational predictions and data from chemical and enzymatic RNA probing (Figures 5 and S3), we have demonstrated that binding of RnaG to the target *icsA* mRNA alters the secondary structure of the untranslated region of the sense RNA. Not all of the antisense region (120 nt) of RnaG seems required to elicit this structural change. Particularly, nucleotides downstream position 80 on *icsA* mRNA display a significant high reactivity to modifying agents even in the presence of the antisense RNA, indicating that the first 40 nt of RnaG are not stably interacting with *icsA* transcript. Indeed, we propose that as the formation of the RNA–RNA hybrid proceeds (~80 nt), the hairpins AH1 and AH2 (in part) are not properly folded and the *icsA* mRNA adopts a different conformation. This alternate structure, which originates in the course of the transcription process but only in the presence of RnaG, gives rise to an intrinsic terminator in the nascent messenger that likely leads to premature termination of *icsA* synthesis (Figure 6). The location of the terminator (between position +78 and +105) is fully consistent with the size of the truncated transcript (~100 nt) deduced from *in vitro* transcription assays and with the RnaG-mediated shortening of cDNA in the primer extension experiment. The proposed model is also strengthened by the considerably reduced capability of RnaG to stop transcription and to generate an abortive product with the *icsA*81/4 mutant, which is impaired in forming the stem structure characteristic of functional intrinsic terminators (Figures 3 and 7). It is remarkable that RNA probing experiments provide a static picture of final structure of the *icsA*–RnaG complex, generated by two molecules which are already synthesized, denatured and refolded together, conditions that facilitate their annealing through the entire complementary region. Moreover, little differences are obviously observed depending on the nature of agents used (chemicals or enzymes). For these reasons, *in vitro* transcription experiments represent the assay that closer mimics the *in vivo* condition because sense–antisense pairing likely occurs during *icsA* mRNA elongation and folding.

Recently, we also started to investigate on the secondary structure of free RnaG and so far the first 120 nt have been analyzed (Figure 8). Complete clarification of its structure is actually in progress. Inspection of the structural conformation of sense and antisense RNAs reveals that the unpaired bases of the apical loop (nucleotides 29–37) and of the basal bulge (nucleotides 57–61) of AH1 might basepair with loops GH3 (nucleotides 84–92) and GH2 (nucleotides 60–64), respectively. Analogously, the apical (nucleotides 85–91) and internal loops (nucleotides 98–102) of AH2 are complementary to single-stranded regions found in the topologically similar motif GH1 (nucleotides 30–36 and 18–22) of RnaG (Figures 5E and 8B). Possibly, these interactions provide the initial contact (kissing complex), which is followed by rapid helix progression leading to RNA–RNA hybrid formation. To date, this binding pathway was restricted to plasmid replication as in R1 and in ColE1 and to the regulation of insertion sequence IS10 (50). Transcription attenuation mechanism was first discovered many years ago in the replication control of the staphylococcal plasmid pT181 and of the streptococcal plasmids pIP501 and pAMβ1 (6,44). More recently, attenuation has been hypothesized also in the regulation of *repABC* genes involved in the replication of tumor-inducing plasmids of *Agrobacterium tumefaciens* (51), although the existence of a terminator, upon RNA–RNA interaction, was only predicted *in silico*. Thus, our study represents the first strong evidence of the occurrence of a transcription attenuation mechanism in genes from Gram-negative bacteria and not restricted to plasmids replication. Regulation of transcription by ncRNAs is certainly the less frequently adopted mode of action even if it constitutes a very powerful strategy to control the first step of the flow of genetic information. In this context, a subset of small RNAs associates directly with and regulates components of the transcription machinery. The best-studied example is the *E. coli* 6S RNA that inhibits transcription by competing for DNA interaction with the RNA polymerase (particularly with the 4.2 binding region of σ70), being recognized as a open promoter (52,53). Other non-coding RNAs have been found in eukaryotic cells targeting mostly the RNA polymerase (likely to the 6S RNA) and transcriptional factors (54).

RNA probing reveals that the AUG start codon and the ribosome binding site of the *icsA* messenger, located between positions +125 and +140, reside in a region that remains single stranded and thus these sites are presumably available for ribosome recognition. Moreover, the accessibility of this region to modifying agents and RNases does not significantly change also in the

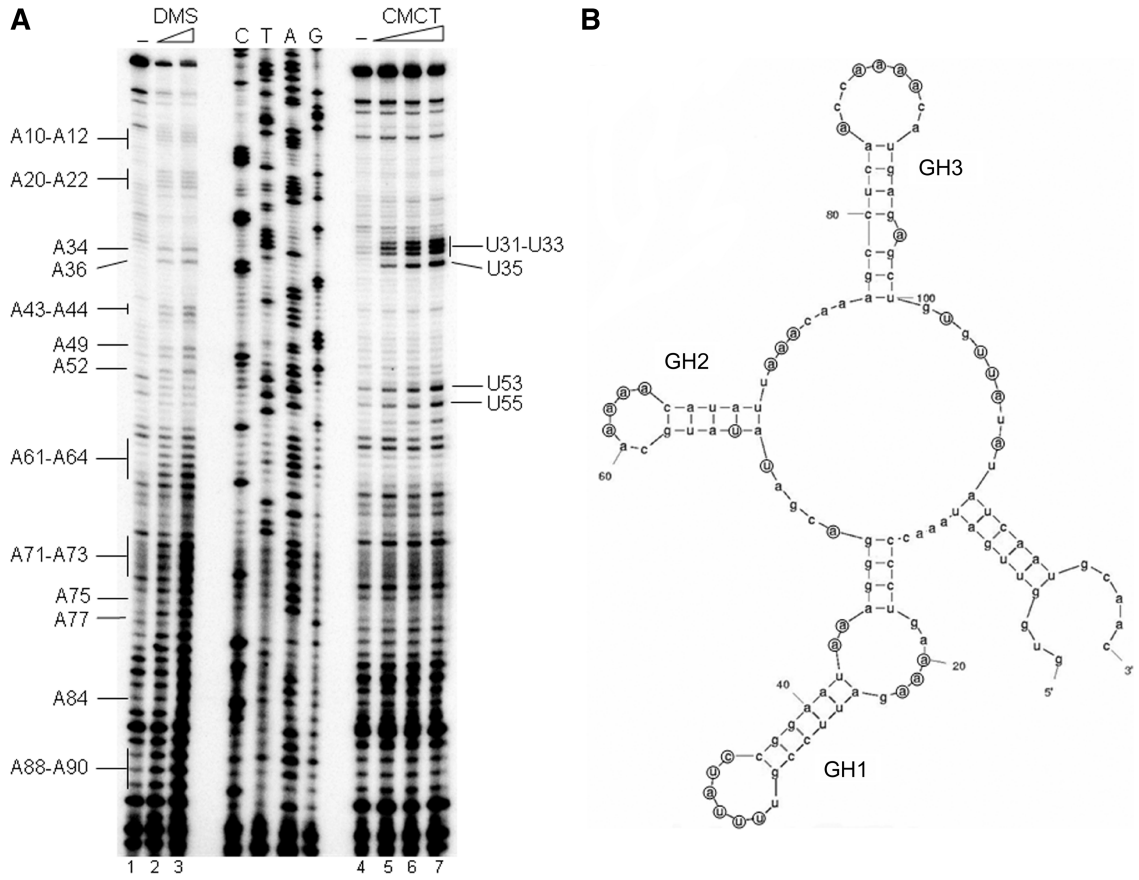

Figure 8. Secondary structure of the antisense region of RnaG. (**A**) Chemical probing of RnaG has been carried out essentially as described in Figure 5B, using the oligo G + 1H, in the presence of increasing amounts of DMS (lane 1, 0%; lane 2, 0.3%; lane 0.6%) and CMCT (lane 4, 0 mg/ml; lane 5, 1.5 mg/ml; lane 6, 3 mg/ml; lane 7, 6 mg/ml). (**B**) Schematic representation of the secondary structure of RnaG (nucleotides from +1 to +120). Numbering is according to 5′ end of antisense RNA (position +120 on *icsA* sequence) and nucleotides reactive to DMS or CMCT in (A) are circled in (B) except those between positions +97 and +108 that have been visualized in an additional experiment performed with an other primer (data not shown). The same secondary structure was obtained using RnaG120 (data not shown).

presence of RnaG once mRNA–RnaG hybrid is formed (Figure 5). Although further studies are needed to better understand the complex regulation of *icsA*, these observations do not support the hypothesis that RnaG, in addition to transcription, may also control the expression of the target gene at post-transcriptional level. Alternatively, we cannot exclude that the long duplex between sense and antisense RNAs may represent a target for RNase III resulting in RNA degradation *in vivo* (6).

RnaG is one of largest regulatory sRNAs (450 nt long) so far identified and RNA probing of the 5′-end and *in silico* prediction show a high structured organization consisting of several stem–loop motifs. Importantly, a deleted form, RnaG120, containing only the antisense region still preserves its ability to represses *icsA* transcription *in vitro* (Supplementary Figure S3) as well as the entire RnaG indicating that the first 120 nt possibly represent a functional domain. In light of its complexity, RnaG, besides *icsA,* might be implicated in contacting and modulating the expression of other virulence genes. It is worth stressing that the 3′-end of RnaG overlaps the –35 promoter element of *virA*, an important virulence gene responsible for destroying host cell microtubules (55) and having its transcriptional start site reversely oriented (position –364) with respect to that of *icsA*. Possibly, transcription from RnaG promoter could affect the activity of the tandemly transcribed *virA* promoter interfering with the binding of the RNA polymerase and/or of its specific regulatory proteins through TI regulation mechanisms. In this context, a well-characterized example of a multiple target regulator is represented by RNAIII of *S. aureus* which turns on/off the expression of genes involved in pathogenesis in response to environmental and host signals (56).

The expression of *icsA* is subjected to a complex regulatory network that besides being affected by RnaG, depends also on the coordinated action of two regulators, VirF and H-NS (27,28). VirF is required to induce the synthesis of the IcsA protein at the permissive temperature of 37°C, a condition which the pathogen meets upon invading the host. Preliminary experiments performed in our laboratory show that VirF can stimulate the *in vitro* activity of P_{icsA} while repressing that of P_{RnaG} (unpublished data). Although several aspects of the regulation of *icsA* mediated by VirF, H-NS and RnaG are

currently under investigation, it is reasonable to hypothesize that VirF positively affects *icsA* transcription both directly and by relieving the inhibitory effect caused by RnaG.

In the past few years, predictive bioinformatics searches have greatly facilitated the discovery of dozens of sRNAs in different species but identification of direct targets and understanding of their mechanistic aspects still lag behind. Elucidating the mode of action of novel non-coding sRNAs is obviously important, particularly in the light of their role in response to the host environment and during the infection process. This study contributes to provide a cleaner picture of the biological functions of antisense RNAs in human bacterial pathogens and clarifies some relevant aspects of the complex regulation of the invasion protein IcsA whose expression is critical for construction of attenuated vaccinal strains.

ACKNOWLEDGEMENTS

We are thankful to A. Giuliodori and G. Spedalieri for technical advices in the RNA probing methods and to C. Sasakawa for plasmid pMYSH6601. The critical reading of the manuscript and helpful discussion of S. Marzi, C.O. Gualerzi. M.L. Bernardini and G. Micheli are gratefully acknowledged.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Funding for open access charge: Fondi Ateneo Ricerca (FAR) to M.F. and Fondazione Cenci Bolognetti to B. C.

Conflict of interest statement. None declared.

REFERENCES

1. Altuvia,S. (2007) Identification of bacterial small non-coding RNAs: experimental approaches. *Curr. Opin. Microbiol.*, **10**, 257–261.
2. Huang,H.Y., Chang,H.Y., Chou,C.H., Tseng,C.P., Ho,S.Y., Yang,C.D., Ju,Y.W. and Huang,H.D. (2009) sRNA Map: genomic maps for small non-coding RNAs, their regulators and their targets in microbial genomes. *Nucleic Acids Res.*, **37**, (Database issue):D150–D154.
3. Gottesman,S. (2004) The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu. Rev. Microbiol.*, **58**, 303–328.
4. Gottesman,S. (2002) Stealth regulation: biological circuits with small RNA switches. *Genes Dev.*, **16**, 2829–2842.
5. Storz,G., Opdyke,J.A. and Zhang,A. (2004) Controlling mRNA stability and translation with small, noncoding RNAs. *Curr. Opin. Microbiol.*, **7**, 140–144.
6. Brantl,S. (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Curr. Opin. Microbiol.*, **10**, 102–109.
7. Repoila,F. and Darfeuille,F. (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol. Cell*, **101**, 117–131.
8. Christiansen,J.K., Nielsen,J.S., Ebersbach,T., Valentin-Hansen,P., Sogaard-Andersen,L. and Kallipolitis,B.H. (2006) Identification of small Hfq-binding RNAs in *Listeria monocytogenes*. *RNA*, **12**, 1383–1396.
9. Mandin,P., Repoila,F., Vergassola,M., Geissmann,T. and Cossart,P. (2007) Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. *Nucleic Acids Res.*, **35**, 962–974.
10. Pichon,C. and Felden,B. (2005) Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc. Natl Acad Sci. USA*, **102**, 14249–14254.
11. Livny,J. and Waldor,M.K. (2007) Identification of small RNAs in diverse bacterial species. *Curr. Opin. Microbiol.*, **10**, 96–101.
12. Pfeiffer,V., Sittka,A., Tomer,R., Tedin,K., Brinkmann,V. and Vogel,J. (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. *Mol. Microbiol.*, **66**, 1174–1191.
13. Padalon-Brauch,G., Hershberg,R., Elgrably-Weiss,M., Baruch,K., Rosenshine,I., Margalit,H. and Altuvia,S. (2008) Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. *Nucleic Acids Res.*, **36**, 1913–1927.
14. Vogel,J. (2009) A rough guide to the non-coding RNA world of Salmonella *Mol. Microbiol.*, **71**, 1–11.
15. Romby,P., Vandenesch,F. and Wagner,E.G.H. (2006) The role of RNAs in regulation of virulence-gene. *Curr. Opin. Microbiol.*, **9**, 1–8.
16. Toledo-Arana,A., Repoila,F. and Cossart,P. (2007) Small noncoding RNAs controlling pathogenesis. *Curr. Opin. Microbiol.*, **10**, 182–188.
17. Song,T., Mika,F., Lindmark,B., Liu,Z., Schild,S., Bishop,A., Zhu,J., Camilli,A., Johansson,J., Vogel,J. *et al.* (2008) A new *Vibrio cholerae* sRNA modulates colonization and affects release of outer membrane vesicles. *Mol. Microbiol.*, **70**, 100–111.
18. Liu,J.M., Livny,J., Lawrence,M.S., Kimball,M.D., Waldor,M.K. and Camilli,A. (2009) Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. *Nucleic Acids Res.*, **37**, e46.
19. Buchrieser,C., Glaser,P., Rusniok,C., Nedjari,H., D'Hauteville,H., Kunst,F., Sansonetti,P. and Parsot,C. (2000) The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. *Mol. Microbiol.*, **38**, 760–771.
20. Bernardini,M.L., Mounier,J., d'Hauteville,H., Coquis-Rondon,M. and Sansonetti,P.J. (1989) Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proc. Natl Acad. Sci. USA*, **10**, 3867–3871.
21. Goldberg,M.B. (2001) Actin–based motility of intracellular microbial pathogen. *Microbiol. Mol. Biol. Rev.*, **65**, 595–626.
22. Falconi,M., Colonna,B., Prosseda,G., Micheli,G. and Gualerzi,C.O. (1998) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of *virF* promoter to transcriptional repressor H-NS. *EMBO J.*, **17**, 7033–7043.
23. Prosseda,G., Falconi,M., Giangrossi,M., Gualerzi,C.O., Micheli,G. and Colonna,B. (2004) The *virF* promoter in *Shigella*: more than just a curved DNA stretch. *Mol. Microbiol.*, **51**, 523–537.
24. Stella,S., Spurio,R., Falconi,M., Pon,C.L. and Gualerzi,C.O. (2005) Nature and mechanism of the *in vivo* oligomerizzation of nucleoid protein H-NS. *EMBO J.*, **24**, 2896–2905.
25. Stella,S., Falconi,M., Lammi,M., Gualerzi,C.O. and Pon,C.L. (2006) Environmental control of the *in vivo* oligomerization of nucleoid protein H-NS. *J. Mol. Biol.*, **355**, 169–174.
26. Dorman,C.J. (2007) H-NS, the genome sentinel. *Nature Rev.*, **5**, 157–161.
27. Dorman,C.J. and Porter,M.E. (1998) The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. *Mol. Microbiol.*, **29**, 677–684.
28. Prosseda,G., Falconi,M., Nicoletti,M., Casalino,M., Micheli,G. and Colonna,B. (2002) Histone-like proteins and *Shigella* invasivity regulon. *Res. Microbiol.*, **153**, 461–468.
29. Shearwin,K.S., Callen,B.P. and Egan,J.B. (2005) Transcriptional interference – a crash course. *Trends Genet.*, **21**, 339–345.
30. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning. A Laboratory Manual*. CSHL Press, Cold Spring Harbor, NY.

31. Simons,R.W., Houman,F. and Kleckner,N. (1987) Improved single and multicopy lac-based cloning vectors for protein and operon fusions. *Gene*, **53**, 85–96.
32. Göransson,M., Sondén,B., Nilsson,P., Dagberg,B., Forsman,K., Emanuelsson,K. and Uhlin,B.E. (1990) Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature*, **12**, 682–685.
33. Miller,J.H. (1992) *A Short Course in Bacterial Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
34. Lett,M.C., Sasakawa,C., Okada,N., Sakai,T., Makino,S., Yamada,M., Komatsu,K. and Yoshikawa,M. (1989) *virG*, a plasmid-coded virulence gene of *Shigella flexneri*: identification of the VirG protein and determination of the complete coding sequence. *J. Bacteriol.*, **171**, 353–3599.
35. Brandi,A., Pietroni,P., Gualerzi,C.O. and Pon,C.L. (1996) Post-transcriptional regulation of CspA expression in *Escherichia coli*. *Mol. Microbiol.*, **19**, 231–240.
36. Powell,B.S., Rivas,M.P., Court,D.L., Nakamura,Y. and Turnbough,C.L. Jr (1994) Rapid confirmation of single copy lambda prophage integration by PCR. *Nucleic Acids Res.*, **22**, 5765–5766.
37. Holmberg,L., Melander,Y. and Nygård,O. (1994) Probing the structure of mouse Ehrlich ascites cell 5.8, 18S and 28S ribosomal RNA in situ. *Nucleic Acids Res.*, **22**, 1374–1382.
38. Donis-Keller,H., Maxam,A.M. and Gilbert,W. (1977) Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res.*, **4**, 2527–2538.
39. von Gabain,A., Belasco,J.G., Schottel,J.L., Chang,A.C. and Cohen,S.N. (1983) Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. *Proc. Natl Acad. Sci. USA*, **80**, 653–657.
40. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.
41. Unniraman,S., Prakash,R. and Nagaraja,V. (2002) Conserved economics of transcription termination in eubacteria. *Nucleic Acids Res.*, **30**, 675–684.
42. Ehresmann,C., Baudin,F., Mongel,M., Romby,P., Ebel,J.P. and Ehresmann,B. (1987) Probing the structure of RNAs in solution. *Nucleic Acids Res.*, **15**, 9109–9128.
43. Zuker,M., Mathews,D.H. and Turner,D.H. (1999) Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide in RNA biochemistry and biotechnology. In Barciszewski,J. and Clark,B.F.C. (eds), *NATO ASI Series*. Kluwer Academic Publishers, Dordrecht, the Netherlands.
44. Brantl,S. (2002) Antisense-RNA regulation and RNA interference. *Biochim. Biophys. Acta*, **1575**, 15–25.
45. Wagner,E.G., Altuvia,S. and Romby,P. (2002) Antisense RNAs in bacteria and their genetic elements. *Adv. Genet.*, **46**, 361–398.
46. Storz,G., Altuvia,S. and Wassarman,K.M. (2005) An abundance of RNA regulators. *Annu. Rev. Biochem.*, **74**, 199–217.
47. Guillier,M., Gottesman,S. and Storz,G. (2006) Modulating the outer membrane with small RNAs. *Genes Dev.*, **20**, 2338–2348.
48. Vogel,J. and Papenfort,K. (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr. Opin. Microbiol.*, **9**, 605–611.
49. Valentin-Hansen,P., Eriksen,M. and Udesen,C. (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol. Microbiol.*, **51**, 1525–1533.
50. Wagner,E.G.H. and Brantl,S. (1998) Kissing and RNA stability in antisense control of plasmid replication. *Trends Biochem. Sci.*, **23**, 451–454.
51. Chai,Y. and Winans,S.C. (2005) A small antisense RNA downregulates expression of an essential replicase protein of *Agrobacterium tumefaciens* Ti plasmid. *Mol. Microbiol.*, **56**, 1574–1585.
52. Wassarman,K.M. (2007) 6S RNA: a small RNA regulator of transcription. *Curr. Opin. Microbiol.*, **10**, 164–168.
53. Cavanagh,A.T., Klocko,A.D., Liu,X. and Wassarman,K.M. (2008) Promoter specificity for 6S RNA regulation of transcription is determined by core promoter sequences and competition for region 4.2 of sigma70. *Mol. Microbiol.*, **67**, 1242–1256.
54. Barrandon,C., Spiluttini,B. and Bensaude,O. (2008) Non-coding RNAs regulating the transcriptional machinery. *Biol Cell.*, **100**, 83–95.
55. Yoshida,S., Handa,Y., Suzuki,T., Ogawa,M., Suzuki,M., Tamai,A., Abe,A., Katayama,E. and Sasakawa,C. (2006) Microtubule-severing activity of *Shigella* is pivotal for intercellular spreading. *Science*, **10**, 985–989.
56. Boisset,S., Geissmann,T., Huntzinger,E., Fechter,P., Bendridi,N., Possedko,M., Chevalier,C., Helfer,A.C., Benito,Y., Jacquier,A. *et al.* (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes Dev.*, **21**, 1353–1366.
57. Falconi,M., Higgins,N.P., Spurio,R., Pon,C.L. and Gualerzi,C.O. (1993) Expression of gene encoding the major bacterial nucleoid protein H-NS is subject to transcriptional auto-repression. *Mol. Microbiol.*, **10**, 273–282.
58. Ermolaeva,M.D., Khalak,H.G., White,O., Smith,H.O. and Salzberg,S.L. (2000) Prediction of transcription terminators in bacterial genomes. *J. Mol. Biol.*, **301**, 27–33.