1818–1827 *Nucleic Acids Research, 2012, Vol. 40, No. 4* *Published online 3 November 2011*
doi:10.1093/nar/gkr850

An unstructured 5′-coding region of the *prfA* mRNA is required for efficient translation

Edmund Loh[1,2,3], Faranak Memarpour[1], Karolis Vaitkevicius[1,2,3], Birgitte H. Kallipolitis[4], Jörgen Johansson[1,2,3,*] and Berit Sondén[1,*]

[1]Department of Molecular Biology, [2]Laboratory for Molecular Infection Medicine Sweden (MIMS), [3]Umeå Centre for Microbial Research (UCMR), Umeå University, 90187 Umeå, Sweden and [4]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark

Received December 16, 2010; Revised and Accepted September 22, 2011

ABSTRACT

Expression of virulence factors in the human bacterial pathogen *Listeria monocytogenes* is almost exclusively regulated by the transcriptional activator PrfA. The translation of *prfA* is controlled by a thermosensor located in the 5′-untranslated RNA (UTR), and is high at 37°C and low at temperatures <30°C. In order to develop a thermoregulated translational expression system, the 5′-UTR and different lengths of the *prfA*-coding sequences were placed in front of *lacZ*. When expressed in *Escherichia coli,* the β-galactosidase expression was directly correlated to the length of the *prfA*-coding mRNA lying in front of *lacZ*. A similar effect was detected with *gfp* as a reporter gene in both *L. monocytogenes* and *E. coli*, emphasizing the requirement of the *prfA*-coding RNA for maximal expression. *In vitro* transcription/translation and mutational analysis suggests a role for the first 20 codons of the native *prfA*-mRNA for maximal expression. By toe-print and RNA-probing analysis, a flexible hairpin-loop located immediately downstream of the start-codon was shown to be important for ribosomal binding. The present work determines the importance of an unstructured part of the 5′-coding region of the *prfA*-mRNA for efficient translation.

INTRODUCTION

The human pathogen *Listeria monocytogenes* causes perinatal infections, meningo-encephalitis, meningitis, septicaemia and gastroenteritis. *Listeria monocytogenes* has turned out to be a very important model for the study of host–pathogen interactions and bacterial adaptation to mammalian hosts (1,2). Analysis of *L. monocytogenes* infections have provided considerable knowledge into how bacteria invade cells, escape the phagosome, move intracellularly and disseminate into deeper tissues. A majority of the proteins involved in the different infection steps are encoded on a 9-kb pathogenicity island, and the expression of these factors is dependent on the transcriptional activator PrfA. Expression of the virulence genes is maximal at 37°C, whereas it is very low at 30°C (3). At low temperatures, the 5′-UTR of *prfA* adopts a structure obstructing the binding of the ribosome to the ribosome-binding site. An increase in temperature induces a conformational change in the RNA structure, allowing binding of the ribosome and initiation of translation. Also, a riboswitch whose transcription terminates when binding *S*-adenosylmethionine, was recently identified as a regulator of PrfA expression, acting by an RNA:RNA antisense mechanism (4). The terminated riboswitch (SreA) binds to the 5′-end of the *prfA* thermosensor and represses *prfA* translation at least in part by destabilizing the *prfA* transcript. In contrast to the repressive effect of the thermosensor, the 5′-UTRs lying in front of *inlA*, *hly* and *actA* are each required for maximal expression of their gene-products (5–7). These proteins are essential for adhesion to cell, lysis of phagosome and actin-based motility, respectively. It has been speculated that the 5′-UTR in these cases are important to stabilize the transcript (7). However, no mechanism has yet been shown to explain the function of these 5′-UTRs.

In this article, we examined the role of the *prfA*-coding region for expression. We show that the first 20 codons of the *prfA* mRNA is required to be maintained in a flexible

*To whom correspondence should be addressed. Tel: + 46 90 7852535; Fax: +46 90 772630; Email: jorgen.johansson@molbiol.umu.se
Correspondence may also be addressed to Berit Sondén. Tel: +46 90 7852525; Fax: +46 90 772630; Email: berit.sonden@molbiol.umu.se
Present address:
Edmund Loh, Department of Medicine, Section of Microbiology, Flowers Building, Imperial College London, South Kensington, SW7 2AZ London, and The Sir William Dunn School of Pathology, University of Oxford South Parks Road, OX1 3RE. Oxford, UK.

manner, to allow ribosome binding and translation initiation.

MATERIALS AND METHODS

Oligonucleotides, strains, plasmids and growth conditions

The strains and plasmids used in this study are listed in Tables 1 and 2. The oligonucleotides used in this study are listed in Table 3. *Listeria monocytogenes* strains were grown in BHI broth (Fluka), and *E. coli* strains were grown in Luria-Bertani (LB) broth or on LB-agar. For RNA isolation, *L. monocytogenes* and *E. coli* overnight cultures were diluted 100-fold and grown to the indicated optical density (0.4) in the presence of antibiotics at the following concentrations: carbenicillin, $100\,\mu g\,ml^{-1}$; chloramphenicol, $7\,\mu g\,ml^{-1}$; and kanamycin, $50\,\mu g\,ml^{-1}$. All strains were grown at 37°C or 30°C with aeration.

Plasmid constructions

DNA fragments covering the *prfA* operon region and different length of PrfA coding sequence (1aa, 4aa, 9aa, 20aa) were PCR amplified using the Pfu enzyme and plasmid plis35 as template (PCR primers are listed in Table 3). The DNA fragments were digested with *Sca*I and cloned into *Sca*I digested pJEM12 (9) to generate plasmid pBSN73, pBSN74, pBSN75, pBSN76. To construct PrfA-GFP fusions, pBSN73 and pBSN76 were digested with *ScaI* and *Kpn*I and inserted into *Hin*dIII (blunt-end treated) and *Kpn*I cut pEGFP-N2 (Clontech) to generate plasmid pBSN77 and pBSN78. A PrfA-GFP fragment was excised from pBSN77 and pBSN78 using *Xho*I/*Not*I, filled in and inserted into *Sma*I digested pMK4 to generate plasmid pBSN79 and pBSN80, respectively. Construction of pBSN83 (*prfA*$_{20Mut}$-*lacZ*): Primer pairs pJEM15-1 and prfA20mutD as well as LacZ-D and prfA20mut-U were used in a PCR-reaction with pBSN76 as template, creating fragments A and B. These fragments were purified and used together in a new PCR-mix together with primers pJEM15-1 and LacZ-D, creating fragment C. This fragment was purified, inserted into pJET1.2 and sequenced. The resulting construct was digested with *BamH*I and *Sca*I and the short fragment was purified and inserted into *Sca*I/*BamH*I digested pBSN76 (lacking the short fragment). All constructs were sequenced for accuracy.

RNA isolation

Total cellular RNA was isolated from *L. monocytogenes* and *E. coli* by dissolving pelleted culture (20 ml, $A_{600} = 0.4$) in resuspension solution (10% glucose, 12.5 mM Tris [pH 7.6] and 5 mM EDTA) and fresh EDTA (0.5 M). Samples were immediately transferred to bead beater tubes with roughly 0.4 g glass beads and 500 µl of acid phenol (pH 4.5). The bacteria were disrupted using a mini bead beater (Biospec products) for 75 s. After centrifugation (5 min, 20 800*g*), RNA was recovered by addition of 1 ml of Trizol and 100 µl of chloroform/isoamylalcohol (24:1), followed by centrifugation. Samples were there-after subjected to two additional chloroform/IAA extractions. The aqueous phase was precipitated by adding isopropanol (0.7×) and incubated at −20°C for 20 min. For collection of the pellet, the RNA samples were centrifuged for 25 min. The pellet was dissolved in 200 µl of RNase-free water.

For removal of the remaining DNA, samples were treated with 20 U of DNaseI (Ambion) for 45 min at 37°C. The reaction was terminated by addition of phenol/chloroform/IAA (25:24:1 [pH 6.6]). Centrifuged samples were chloroform/IAA extracted and ethanol precipitated. The pellet was resuspended in 200 µl RNase-free water, RNA concentration was measured on a Nanodrop (Nanodrop ND-1000 Spectrophotometer), and the RNA integrity was determined on a 1.2%

Table 1. Strains used in this study

	Relevant characteristics	Reference/ source
Escherichia coli strain		
XL1 blue	*recA1*, *lac*[F′*proABlacI^{q}ZΔM15*Tn*10*]	Stratagene
XL2 blue	*recA1*, *lac*[F′*proABlacI^{q}ZΔM15*Tn*10*(Tet^r) Amy Cam^r]	Stratagene
Listeria monocytogenes strain		
EGD-e		(8)

Table 2. Plasmids used in this study

Plasmid	Relevant characteristics	Reference source
pJEM12	Shuttle vector *E. coli-Mycobacteria*, LacZ translational fusions, Kan^r	(9)
pEGFP-N2	GFP translational fusions, Kan^r	Clontech
pMK4	Shuttle vector *E. coli -Listeria*, Amp^r, Cml^r	(10)
pJET1.2	Cloning vector	Fermentas
plis35	*prfA* clone in pMK4, Amp^r, Cml^r	(11)
pBSN73	PrfA$_1$-LacZ, Kan^r	This study
pBSN74	PrfA$_4$-LacZ, Kan^r	This study
pBSN75	PrfA$_9$-LacZ, Kan^r	This study
pBSN76	PrfA$_{20}$-LacZ, Kan^r	This study
pBSN77	PrfA$_1$-GFP in pEGFP-N2, Kan^r	This study
pBSN78	PrfA$_{20}$-GFP in pEGFP-N2, Kan^r	This study
pBSN79	PrfA$_1$-GFP in pMK4, Amp^r, Cml^r	This study
pBSN80	PrfA$_{20}$-GFP in pMK4, Amp^r, Cml^r	This study
pBSN83	PrfA$_{20Mut}$-LacZ, Kan^r	This study

Table 3. Primers used in this study, restriction enzyme sites underlined, mutations are shown in bold

Primer name	Sequence	Relevant characteristics
Nde25′	5′-GGGGCATATGCATGTCTCATCCCCCAATCGT-3′	*Nde*I, $PrfA_1$
Nde3	5′-GGGGAGTACTTTATTCCTACAAAAAAGGGTTAGT-3′	*Sca*I
Nde4	5′-GGGGCATATGTTGAGCGTTCATGTCTCATCCCCCAATCGT-3′	*Nde*I, $PrfA_4$
Nde5	5′-GGGGCATATGTTTGAATTCTTCTGCTTGAGCGTTCAT-3′	*Nde*I, $PrfA_9$
Nde6	5′-GGGGCATATGTTTTGGTTTTATCCCGTTAGTTTC-3′	*Nde*I, $PrfA_{20}$
GFP-U	5′-TAAACGGCCACAAGTTCAGC-3′	
GFP-D	5′-TCCTTGAAGAAGATGGTGCG-3′	
LacZ-U	5′-ACCCAACTTAATCGCCTTGC-3′	
LacZ-D	5′-GATCGCACTCCAGCCAGC-3′	
hns-RT1	5′-CAGCTGGAGTACGGCCTTGG-3′	
hns-RT2	5′- CGTACTCTTCGTGCGCAGGC-3′	
tmRNA-U	5′-GGATTCGACGGGATTTGCG-3′	
tmRNA-D	5′-TAGCCTGATTAAGTTTTAACGC-3′	
prfA-pT7	5′-GATAGACTTCGAAATTAATACGACTCACTATAGGTGTAAAAAACATCATTTAGCGT-3′	
LacZ-D-Toe	5′-CAGGGTTTTCCCAGTCACG-3′	
prfA20mut-U	5′-GAGACATGAACGCTCAAGCAGAAG**CG**TTCAAAAAATATTTAGAAACTAACG-3′	
prfA20mut-D	5′-CGTTAGTTTCTAAATATTTTTTGAA**CG**CTTCTGCTTGAGCGTTCATGTCTC-3′	
pJEM15-1	5′-CACAAAACGGTTTACAAGCATAAAG-3′	
LacZ-New-D(toe)	5′-TATCCGGATCCGCGGGC-3′	

agarose gel. Only RNA samples showing distinct non-processed precursors to ribosomal RNA were used in the following experiments.

Northern blot

For northern blotting, 20 µg of total RNA was separated on a formaldehyde agarose gel prior to blotting as described (12). The Hybond-N membrane was subsequently hybridized with ^{32}P-ATP α-labelled DNA fragments amplified with corresponding primers. Northern blots were developed using a STORM machine (Molecular Dynamics). Primers used are listed in Table 3. To amplify a DNA fragment for detection of *prfA*, *hns* and tmRNA, we used GFP-U and GFP-D, hns-RT1 and hns-RT2 and tmRNA-U and tmRNA-D, respectively.

RNA stability

Indicated *E. coli* strains were grown at 37°C in a shaking water bath, until $A_{600} = 0.4$. Initiation of transcription was stopped by the addition of rifampicin to 250 µg ml^{-1}, and samples were collected at indicated time points for RNA isolation.

SDS–PAGE and western blotting

The different cultures were grown in BHI (*L. monocytogenes*) or LB medium (*E. coli*) to an optical density of $OD_{600} = 0.4$. Bacteria were centrifuged and re-suspended in buffer A (200 mM KCl, 50 mM Tris–HCl [pH 8.0], 1 mM EDTA and 10% glycerol). The suspension was disrupted using a bead-beater for 1.5 min at maximum speed. After 2 min on ice, the suspension was centrifuged at 15 000 rpm for 5 min, and the supernatant (cytoplasmic fraction) was removed. Protein samples were separated on a 12% polyacrylamide gel electrophoresis before being transferred onto a PVDF membrane using a semidry blotting apparatus. Development of the membrane essentially followed the protocol of the ECL+ western blotting kit (Amersham), using anti- β-galactosidase (GenWay), anti-β-lactamase (GenWay), anti-GFP (BD-living colours), or anti-GroEL (4) as primary antibodies and HRP-conjugated anti-rabbit or anti-mouse as secondary antibodies (Bio-Rad), respectively. Measurement of protein expression was carried out using a STORM machine (Molecular Dynamics).

In vivo protein stability experiment

To determine the intracellular stability of PrfA-LacZ, we used a technique described by (13). Protein stability was monitored after the protein synthesis had been inhibited by the addition of spectinomycin (100 µg ml^{-1}) to bacterial cultures grown to $OD_{600} = 0.4$ in LB medium supplemented with kanamycin, 50 µg ml^{-1} at 37°C. Samples to be analysed by western blotting were removed at indicated times.

In vitro transcription/translation

One microgram of pT7p*prfA$_1$-gfp* and pT7p*prfA$_{20}$-gfp* plasmids (T7 driven *prfA$_1$-gfp* and *prfA$_{20}$-gfp*, amplified using primers prfA-pT7 and GFP-D were inserted into pGEM-T) were *in vitro* transcribed in an S30 T7 high yield *in vitro* Transcription/Translation Kit (Promega) according to the manufacturer's instructions. In brief, the mixtures were incubated at 25°C for 5 min before transfer to 37°C for an additional 5 min. Samples were acetone-precipitated, re-suspended in sample buffer, and separated on a 12% polyacrylamide gel before being transferred onto a PVDF membrane using a wet blotting apparatus (Biorad). Development of the membrane essentially followed the protocol of the ECL+ western blotting kit (Amersham), using anti-GFP (BD-living colours) and anti- β-lactamase (GenWay) as primary antibodies and HRP-conjugated anti-mouse and anti-rabbit as secondary antibodies (Bio-Rad).

Fluorescent imaging on agar plate

Bacterial strains were streaked onto a LB-plate containing carbenicillin (100 µg/ml) and were grown overnight. Fluorescence imaging was performed with an IVIS Spectrum imaging system (Xenogen). A GFP filter (excitation wavelength 445–490 nm and emission 515–575 nm) was used for acquiring fluorescence imaging. Identical illumination settings, such as exposure time (1 s) and field of views (15×15 cm), were used for acquiring all images. Fluorescence emission was normalized to photons per second per centimeter squared per steradian ($p\,s^{-1}\,cm^{-2}\,sr^{-1}$). Images were acquired and analysed using Living Image 3.0 software (Xenogen).

β-galactosidase assay

For the β-galactosidase assay, samples were taken at $OD_{600} = 0.5$. The β-galactosidase reactions were assayed essentially as described by (14), with the exception that we used chloroform and 0.002% SDS to disrupt the bacteria. All data represent the average from assays performed in duplicate in three independent experiments, and the means + standard deviations are shown in the plotted graphs.

Toe-print assay

Templates for *in vitro* transcription of *prfA*$_{20}$-*lacZ* and *prfA*$_{20Mut}$-*lacZ* were constructed by PCR using the primers (prfA-pT7 and LacZ-D-Toe) listed in Table 3. The templates contain a 5′-end T7 promoter. *In vitro* transcription was performed using the RiboMAX™ Large Scale RNA production sytems-SP6 and T7 kit as described by the manufacturer (Promega). *In vitro* transcribed RNA was ethanol precipitated, resuspended in formamide loading dye and separated on an 8% denaturing polyacrylamide gel. The RNA was visualized by UV shadowing, excised from the gel and transferred to 300 µl 2 M NH_4Acetate. After overnight incubation at 14°C, the RNA was phenol extracted followed by ethanol precipitation. Quantification was performed on a NanoDrop 2000. *In vitro* transcribed RNA was 5′-end-labelled using the KinaseMax kit as described by the manufacturer (Ambion).

Toeprinting experiments were performed in 10 µl reactions with 0.1 pmol of *prfA*$_{20}$-*lacZ* or *prfA*$_{20Mut}$-*lacZ*. The RNA was pre-incubated for 20 min and subsequently mixed with 0.6 pmol of 5′-end-labelled LacZ-D-Toe probe in a buffer containing 60 mM NH_4Cl, 10 mM Tris–acetate [pH 7.5], 10 mM DTT, 1 µl RNAguard and 100 µM dNTP. The mixture was incubated 2 min at 94°C and then placed on ice for 5 min and 37°C for 5 min. Three different concentrations of 30S ribosomes (0.4, 1.0 and 1.5 pmol) (*E. coli* MRE600) were added followed by 10 min incubation. The mixture was supplemented with 10 µM uncharged $tRNA_f^{Met}$ (Sigma) followed by 15 min incubation after which, 2 U of AMV reverse transcriptase was added. The reaction was stopped after 30 min by the addition of 10 µl formamide loading dye. In parallel, sequencing reactions were prepared using *prfA*$_{20}$-*lacZ* and *prfA*$_{20Mut}$-*lacZ* DNA as templates. The resulting DNA was separated on an 8% denaturing polyacrylamide sequencing gel and the resulting toe-print was measured with a STORM machine using the signal obtained from the sample without $tRNA_f^{Met}$ as background controls.

T1 ribonuclease structure mapping

The *prfA* UTR region was amplified by PCR from pBSN76 and pBSN83 plasmids with primers prfA-PT7 and LacZ-New-D(toe) using Phusion DNA polymerase (Finnzymes). An amount of 4 µg of gel-purified products were used as template for *in vitro* transcription with a RiboMAX RNA Production System T7 (Promega) in a total volume of 100 µl according to manufacturer's instructions. After DNase treatment the reaction products were chloroform extracted, ethanol precipitated, resuspended in DEPC treated water and purified by size exclusion using NucAway spin columns (Ambion). The synthesized RNA was separated by electrophoresis on a denaturing 8 M urea, 6% AA/bisAA (29:1) gel in a TBE buffer. The bands were detected by UV shadowing, excised, and eluted overnight at 4°C into 500 µl of 500 mM ammonium acetate, 1 mM EDTA pH 6.5 in presence of 100 µl acid phenol/chloroform (Ambion). Eluted RNA was chloroform extracted, ethanol precipitated and dissolved in water. An amount of 10 pmol of purified RNA were dephosphorylated with FastAP alkaline phosphatase (Fermentas) and 5′-^{32}P-labelled using a T4 polynucleotide kinase (Fermentas) as described by manufacturer. Following chloroform extraction the unincorporated label was removed by size exclusion using ProbeQuant spin columns (GE). The labelled RNA was gel-purified as described above. For structural probing ~0.1 pmol of labelled RNA and 1 µg total yeast RNA (Ambion) were used per reaction. Before structural probing the RNA was denatured by incubating at 95°C for 1 min and cooling on ice for 5 min. Following denaturation the RNA was diluted in 1× Structure Buffer (Ambion) and renatured at 37°C for 20 min. An amount of 2 µl of appropriately diluted RNAse T1 (Ambion) were added into 10 µl aliquots containing ~0.1 pmol labelled RNA and 1 µg total yeast RNA and continued to incubate at 37°C for 5 min. To generate RNAse T1 sequencing ladder 2 µl of RNA (~0.2 pmol labelled RNA and 2 µg total yeast RNA) were mixed with 9 µl of 1× Sequencing Buffer (Ambion), and incubated with 1 µl of 0.4 U µl^{-1} RNase T1 at 50°C for 5 min. Alkaline hydrolysis ladder was prepared by mixing 2 µl of RNA (~0.2 pmol labelled RNA and 2 µg total yeast RNA) with 10 µl 1× Alkaline Hydrolysis Buffer (Ambion) and incubation at 95°C for 15 min. Reactions were stopped by addition of 12 µl Gel Loading Buffer II (Ambion) and immediate freezing the tubes in dry ice. Prior to electrophoresis RNA samples for structure probing were incubated for 1 min at 95°C and kept on ice. The RNA was separated on a denaturing 8 M urea, 6% AA/bisAA (19:1) gel in TBE buffer.

In silico RNA folding

RNA sequences of different constructs/mRNAs were analysed using the RNAfold web server of the Vienna

RNA package (http://rna.tbi.univie.ac.at/egi-bin/RNAfold.cgi). For each sequence, the minimum free energy in kcal mol^{-1} was predicted (15).

RESULTS

The expression of PrfA is directly correlated to the amount of *prfA*-coding sequence

In order to develop a thermo-inducible translation system for Mycobacterial species, the *prfA*-thermosensor from *L. monocytogenes* was chosen as a scaffold due to its temperature-sensing properties (3). It has previously been shown that only six codons (18 bases) of the *prfA*-coding mRNA was sufficient for proper thermosensing, when fused in front of *gfp* (3). To investigate if a difference in the length of the *prfA*-coding region affected thermosensing, DNA encoding; 1 codon (3 bases), 4 codons (12 bases), 9 codons (27 bases) or 20 codons (60 bases), respectively, was inserted in front of *lacZ* and the constructs were introduced into *Escherichia coli* (Supplementary Figure S1 and 'Materials and Methods' section). These constructs all harboured the identical native *prfA* promoters and were inserted in the identical *Sca*I cloning site in the Multiple Cloning site (MCS) of the vector generating translation fusions (*i.e.* all constructs contained the same length of the MCS). To test whether the *prfA-lacZ* fusions still were thermoregulated, β-galactosidase expression was measured at 30°C and 37°C. Except for the one codon construct, thermosensing was still retained in the different *prfA-lacZ* fusions, with 2- to 4-fold higher expression at 37°C compared to 30°C (Supplementary Figure S2). More strikingly though, was the correlation between the β-galactosidase expression and the length of the *prfA*-coding sequence (Figure 1A). The β-galactosidase expression was ~15-fold higher when 20 codons of *prfA* were inserted in front of *lacZ,* (creating *prfA*$_{20}$-*lacZ*), compared to one codon (*prfA*$_1$-*lacZ*). The constructs carrying either four (*prfA*$_4$-*lacZ*) or nine codons (*prfA*$_9$-*lacZ*) had a β-galactosidase expression lying in between *prfA*$_{20}$-*lacZ* and *prfA*$_1$-*lacZ* (Figure 1A). Importantly, the β-galactosidase activity was directly correlated with protein expression as determined by western blotting (Figure 1B). An equal amount of plasmids could be extracted from each strain grown to mid-log phase, demonstrating that the difference in β-galactosidase expression among the constructs were not due to variations in plasmid maintenance and stability (data not shown).

The increased *prfA*-expression is not due to the reporter mRNA

It could be hypothesized that the *lacZ*-gene would, by some mechanism, cause the differences in β-galactosidase expression observed. To test this, the *prfA*-UTR with either 1 or 20 codons was inserted in front of *gfp* in the identical site of the vector before introduction into *E. coli.* We reasoned that if the difference in expression between the *prfA*$_1$ and the *prfA*$_{20}$ was still detected with *gfp* as a reporter mRNA, it would furthermore demonstrate the

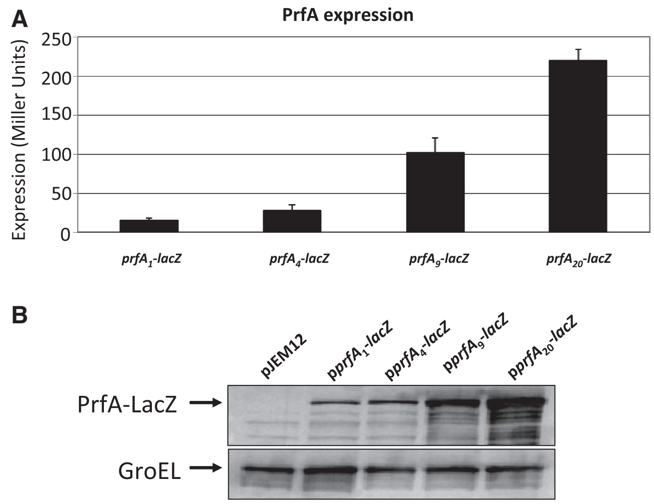

Figure 1. (**A**) Measurement of *prfA*-expression with *prfA-lacZ* translation fusions. *Escherichia coli* carrying the indicated plasmids were grown in LB medium until OD_{600} = 0.5. β-galactosidase activity was monitored as described in 'Material and Methods' section. (**B**) Western blot analysis of PrfA-LacZ expression. Total protein was isolated from *E. coli* carrying the indicated plasmids and subjected to western blot analysis. Membranes were probed with antibodies recognizing β-galactosidase or GroEL (loading control).

importance of the 20 first codons for efficient PrfA expression and rule out effects caused by the reporter genes. A large difference in fluorescence was detected between strains expressing the PrfA$_1$-GFP or the PrfA$_{20}$-GFP on bacterial agar-plates (Figure 2A). By western blotting, we determined that the difference in the level of PrfA$_1$-GFP and PrfA$_{20}$-GFP was similar to the difference detected between the short and the long *prfA-lacZ* constructs (compare Figures 1B and 2B). Altogether, these results suggest that the 20 first codons of the *prfA*-coding mRNA are important for the expression of PrfA, in a mechanism independent of the reporter mRNAs (*lacZ* or *gfp*).

The *prfA*-coding sequence does not affect the stability of the *prfA-gfp* transcripts

One possible explanation of the above results would be that *prfA-gfp* mRNAs carrying 20 codons was more expressed or more stable than *prfA-gfp* mRNAs with only one codon. To test this, we isolated RNA from cultures (OD_{600} = 0.4) of the strains carrying 1 or 20 codons of *prfA* in front of *gfp* (*prfA*$_1$-*gfp* and *prfA*$_{20}$-*gfp,* respectively). Differences in the length of *prfA* did not affect the steady-state levels of the *prfA-gfp* transcripts (data not shown). These data, however, did not rule out the possibility that the different lengths of *prfA* might affect stability of the transcripts and hence PrfA expression. We therefore performed a transcript stability experiment where rifampicin was added to cultures carrying 1 or 20 codons of *prfA* upstream of *gfp*. Samples were taken at 0, 5, 10 and 15 min after addition of rifampicin. Northern blot results revealed that the *prfA*$_1$-*gfp* and the *prfA*$_{20}$-*gfp* transcripts were equally stable with a half-life of ~8 min (Figure 3). As a control, the decay of the *hns* transcript was followed, showing a half-life of ∼3–5 min (Figure 3). The results show that the variation observed in PrfA expression of the different constructs is not due to an altered transcription or transcript stability.

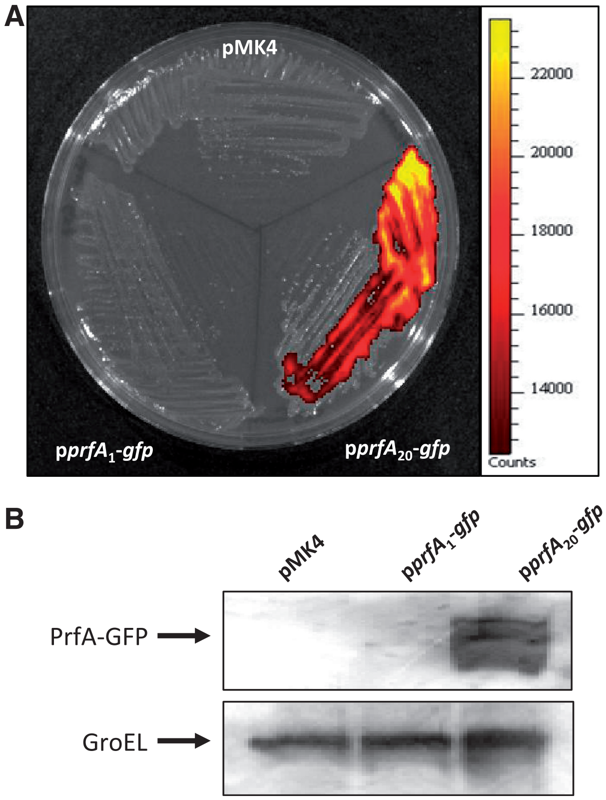

Figure 2. (**A**) Fluorescence measurement of PrfA-GFP fusions. *Escherichia coli* strains carrying the indicated plasmids were streaked on LB-agar plate and grown at 37°C for 24 h. Fluorescence was measured using an IVIS-Spectrum imaging system. Colour scale represents level of fluorescence intensity ranging from high (yellow) to low (dark-red). (**B**) Western blot analysis of PrfA-GFP expression. *E. coli* strains carrying the indicated plasmids were grown to an OD_{600} = 0.4 at 37°C. Total protein was isolated and subjected to western blot analysis. Membrane was probed with antibodies recognizing GFP or GroEL (loading control).

The first 20 codons of the *prfA*-mRNA are important for PrfA expression in *L. monocytogenes*

To test whether the 20 first codons of the *prfA*-mRNA would be required for efficient PrfA-expression in its natural strain background and rule out *E. coli* specific artefacts, the *prfA*$_1$-*gfp* and the *prfA*$_{20}$-*gfp* constructs were introduced into *L. monocytogenes* and the protein expression measured by western blotting. As seen in Figure 4, the amount of PrfA$_{20}$-GFP was higher than the amount of PrfA$_1$-GFP when expressed in *L. monocytogenes*, similar to the difference detected in *E. coli* (compare Figures 2B and 4).

Similar protein stabilities are detected in the PrfA$_1$ and PrfA$_{20}$ codon constructs

In order to examine if the varied PrfA expression in the different constructs was due to a difference in protein stability, strains harbouring *prfA*$_1$-*lacZ* or *prfA*$_{20}$-*lacZ*, were grown to mid-log phase before translation was inhibited by the addition of spectinomycin. After 48 h of spectinomycin treatment, no proteolytic degradation could be observed for either PrfA$_1$-LacZ or PrfA$_{20}$-LacZ (Supplementary Figure S3). This suggests that an altered protein stability cannot explain the reduced amount of PrfA$_1$-LacZ, at least during the time-period of our β-galactosidase experiments.

Thus far, our results indicate that the first 20 codons of the *prfA*-mRNA are necessary for efficient PrfA expression. The mechanism controlling PrfA expression is not governed by altered expression/stability of the *prfA* messenger, nor does it affect the stability of the PrfA protein. Moreover, the mechanism is functioning in both *E. coli* and *L. monocytogenes*.

The 20 first codons of *prfA* are required for efficient translation *in vitro*

To investigate whether the 20 first codons of *prfA* are important for translation, an *in vitro* transcription/translation assay was used (4). An equal amount of the *prfA*$_1$-*gfp* and the *prfA*$_{20}$-*gfp* plasmid constructs were transcribed and translated in a continuous manner. From the reactions, *in vitro* synthesized protein was extracted and the levels measured by western blotting. The level of

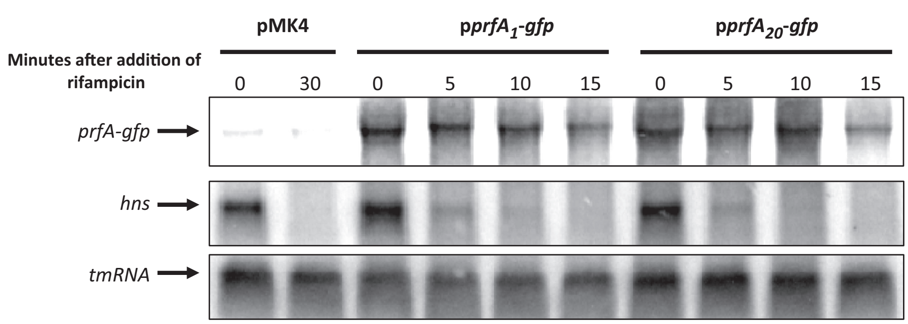

Figure 3. Northern blot analysis examining transcript stability. *Escherichia coli* strains carrying the indicated plasmids were grown until an OD_{600} = 0.4 when rifampicin was added to block further transcription. Samples were isolated at indicated time-points prior to RNA isolation. RNA samples (20 µg) were separated on agarose:formaldehyde gel and subjected to a northern blot analysis. The membrane was hybridized with *gfp*, *hns* and tmRNA (loading control) probes, respectively.

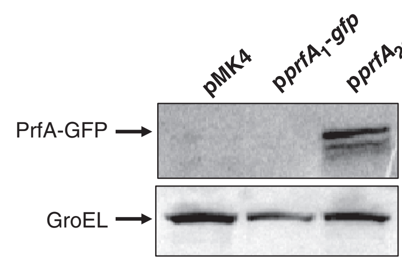

Figure 4. Western blot analysis of PrfA-GFP expression in *L. monocytogenes*. *Listeria monocytogenes* strains carrying the indicated plasmids were grown to an $OD_{600} = 0.4$ at 37°C. Total protein was isolated and subjected to western blot analysis. Membrane was probed with antibodies recognizing GFP or GroEL (loading control).

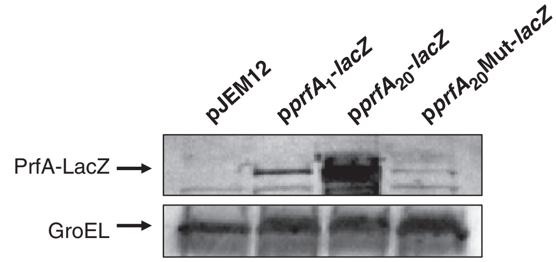

Figure 6. Western blot analysis of PrfA-LacZ expression. Total protein was isolated from *E. coli* carrying the indicated plasmids and subjected to western blot analysis. Membranes were probed with antibodies recognizing β-galactosidase or GroEL (loading control).

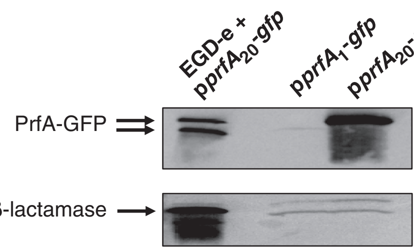

Figure 5. *In vitro* transcription/translation analysis. Indicated plasmids were subjected to *in vitro* transcription/translation analysis as indicated in 'Materials and Methods' section. First lane shows protein extract from EGDe harbouring *prfA*$_{20}$-*gfp*. Samples were removed for western analysis and the membranes were probed with antibodies recognizing GFP or β-lactamase (loading control).

PrfA$_{20}$-GFP was higher than the level of PrfA$_{1}$-GFP construct in a range similar to *prfA*$_{20}$-*gfp* and *prfA*$_{1}$-*gfp* in *E. coli* and *L. monocytogenes* (Compare Figures 5 with Figures 2B and 4). As a control, expression of β-lactamase (encoded on the same plasmid as *prfA-gfp*) was analysed from the same extracts. Expression of β-lactamase did not alter between the samples, showing that the reduced expression of PrfA$_{1}$-GFP compared with PrfA$_{20}$-GFP was not due to a general expression-defect of the plasmid.

A mutation stabilizing the *prfA*$_{20}$-*lacZ* secondary RNA-structure dramatically decreases PrfA expression

Previous reports have shown a correlation between the stability of the mRNA secondary structure and translation. This has been indicated for regions just downstream of the startcodon (within the coding RNA) (16). Also, strong mRNA secondary structures can inhibit an initial interaction between the ribosome and the mRNA at ribosome standby sites located on the mRNA (17,18). Therefore, the mRNA secondary structure stability for the entire 5′-UTR + 60 extra bases downstream of A in AUG were predicted for the different constructs (Supplementary Figure S4). The results indicated that the thermosensor remained relatively intact in all constructs, which is in agreement with our β-galactosidase expression results at different temperatures (Supplementary Figure S2). Instead, the predicted RNA-structure differed downstream of the startcodon, with the shorter constructs (1 or 4 codons) being more stable than the longer constructs [9 or 20 codons (Supplementary Figure 4)]. If the *in silico* predictions of RNA secondary structure stabilities were correct, an increased stability of the *prfA* RNA secondary structure should decrease PrfA expression. To test this, an AA to CG base-substitution mutation at position 137–138 (downstream of the startcodon) was constructed (see 'Materials and Methods' section and Supplementary Figure S4). When measuring β-galactosidase protein expression, it was observed that the *prfA*$_{20\mathrm{Mut}}$-*lacZ* mutant construct displayed a level dramatically reduced compared to the wild-type *prfA*$_{20}$-*lacZ* (Figure 6). To avoid possible effects of the reporter RNA in the *prfA*$_{1}$, *prfA*$_{4}$ and *prfA*$_{9}$ constructs, the subsequent experiments were carried out using only the *prfA*$_{20}$ and the *prfA*$_{20\mathrm{Mut}}$ constructs.

To examine *in vitro* if the structure within the coding region was affected between the *prfA*$_{20}$-*lacZ* and the *prfA*$_{20\mathrm{Mut}}$-*lacZ* transcripts, an RNaseT1 structural probing assay was undertaken. During these conditions, RNaseT1 recognizes unpaired guanine bases. From the data (Figure 7), it was evident that the overall structure of the thermosensor did not differ between the *prfA*$_{20}$-*lacZ* and the *prfA*$_{20\mathrm{Mut}}$-*lacZ* transcripts, in agreement with results from a previous RNA-probing experiment (3). However importantly, a dramatic difference was observed at the hairpin loop predicted to be downstream of the AUG start codon. For the *prfA*$_{20}$-*lacZ* transcript, the guanine bases at positions 128, 131 and 134 were equally unpaired, indicating an unstructured hairpin loop. In contrast, at the equivalent region of the *prfA*$_{20\mathrm{Mut}}$-*lacZ* transcript, only the guanine base at position 128 was unpaired, suggesting a rigid inflexible hairpin with a short loop.

A weak secondary structure within the *prfA* coding RNA increases ribosome binding

A plausible explanation of the higher expression in the *prfA*$_{20}$-*lacZ* construct compared to the *prfA*$_{20\mathrm{Mut}}$-*lacZ* construct would be that the ribosome binds more strongly to the SD-region of *prfA*$_{20}$-*lacZ* than *prfA*$_{20\mathrm{Mut}}$-*lacZ*. To test this, toe-print experiments were conducted, by analysing the capability of ribosome binding to *in vitro* synthesized *prfA*$_{20}$-*lacZ* or *prfA*$_{20\mathrm{Mut}}$-*lacZ* RNA samples (Supplementary Figure S5). Our data indicate that the

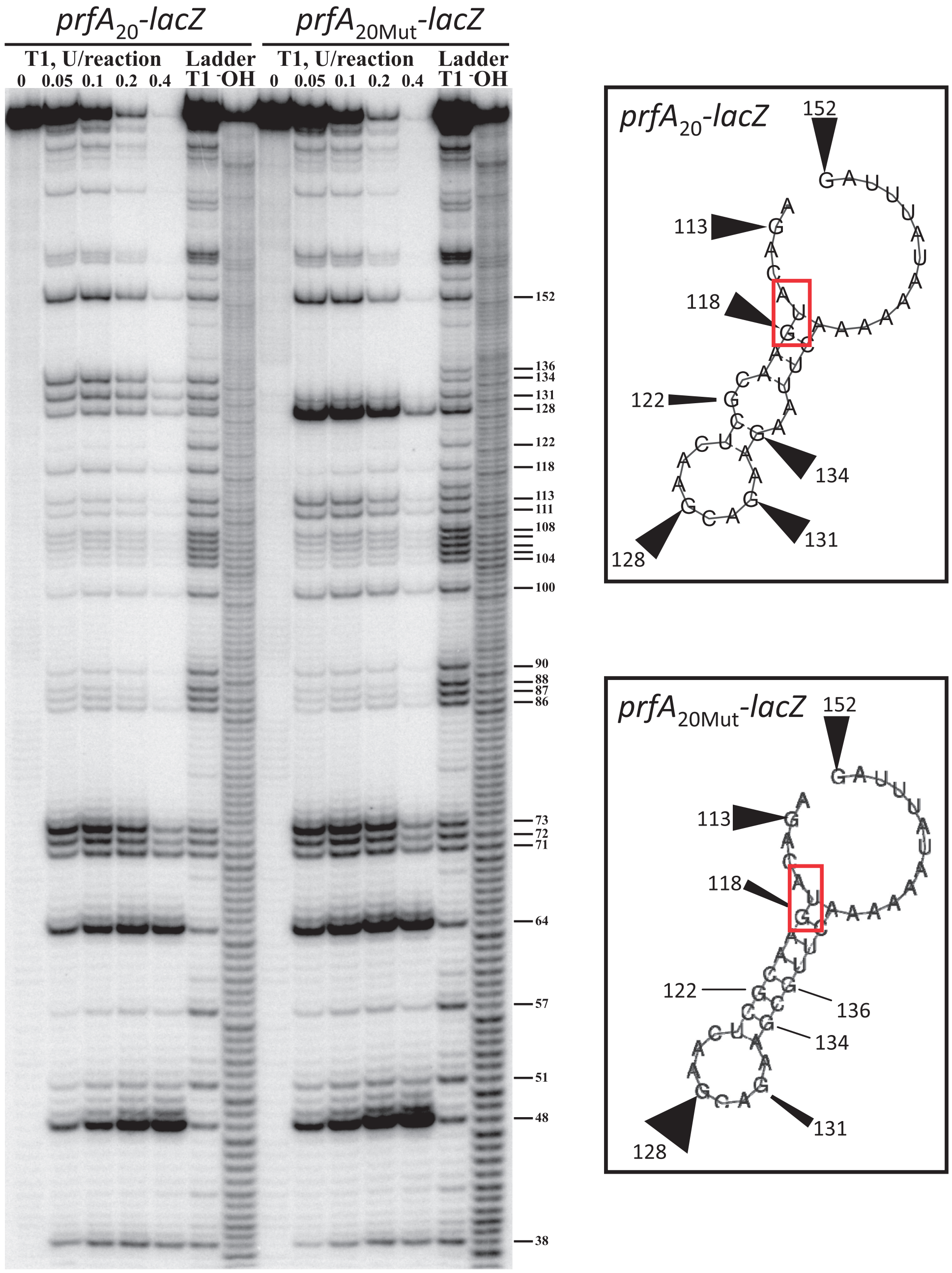

Figure 7. RNase T1 secondary structure probing of the *prfA*$_{20}$-*lacZ* and *prfA*$_{20\mathrm{Mut}}$-*lacZ* transcripts. RNase T1 probing was performed as described in 'Materials and Methods' section. To the left of the gel is shown the location of guanine residues as suggested by T1 sequencing ladder. The OH ladder shows all bases. The insets show the secondary structures of the *prfA*$_{20}$-*lacZ* and *prfA*$_{20\mathrm{Mut}}$-*lacZ*, between positions 114 and 154, respectively. Black arrowheads show the relative prevalence of free guanine residues, being low with a tiny arrowhead and high with a wide arrowhead. Red boxes highlight the AUG startcodon.

ribosome indeed binds the SD-region of the *prfA*$_{20}$-*lacZ* RNA more strongly than it binds the *prfA*$_{20\mathrm{Mut}}$-*lacZ* RNA.

DISCUSSION

In this study, we show that the 5′-end of the *prfA*-coding RNA is important for its expression. Our work shows that the ribosome requires an unstructured RNA, within the first 20 bases downstream of the AUG start codon for efficient binding and translation initiation. Stabilizing this structure severely impairs ribosomal interaction with the RNA leading to a decreased translation. Particularly, a hairpin-loop, located within the first nine codons must be in a flexible state to allow efficient ribosome binding. It has been suggested that a strong mRNA secondary structure in the 5′-part of the coding RNA affects expression negatively, by preventing binding of the ribosome (16,19). We identified an inverse correlation between the PrfA-fusion protein expression levels both *in vivo* and *in vitro* versus the stability of the predicted mRNA secondary structures and particularly an hairpin loop located downstream of the startcodon. Our results are in agreement with the ribosomal standby model (17,18,20). In the article by de Smit and van Duin (17), it was suggested that the ribosome initially binds to an unpaired region of the transcript, the standby site, instead of binding directly to the SD-region if it is occluded by a paired structure (like the thermosensor). By binding to the standby site, the ribosome can more easily compete with SD-regions trapped in secondary structures, during their time of opening (the more stable structure, the shorter the time of opening). It could therefore be hypothesized that the unstructured region downstream of AUG of *prfA* is a ribosomal standby site, where the ribosome can bind and 'wait' for the SD-region to be accessible. Once bound, the ribosome can more efficiently compete with the thermosensor structure occluding the SD-region, thereby increasing the frequency of translation initiation. If the ribosome is prevented to bind to the standby site (by mutations creating a more stable secondary structure) the binding of the ribosome to the SD will be reduced. Also, the unfolding of a structured mRNA after an initial interaction with the ribosome is important to allow the start codon to interact with the initiator tRNA (21).

Previously, the 5′-UTR of the *prfA* transcript has been shown to function as a thermosensor and has also been shown to be regulated by a *trans*-acting riboswitch (3,4). Our results suggest that the coding RNA of *prfA* does not participate in thermoregulation, but rather is important for efficient translation initiation. Vice versa, the RNA probing assay indicates that the thermosensor RNA does not interfere with the unstructured downstream region. This suggests that the binding of the ribosome to the SD is independent of the thermosensor.

Several alternative mechanisms were tested to determine if they could explain the expression difference detected between the *prfA* constructs: (i) Sprengart and colleagues (22), suggested that the presence of a downstream box (DB), located at the 5′-end of the coding-mRNAs, allows direct base-pair interactions between the DB and the 16S rRNA in the ribosome. However, no such DB-box showing complementarity against the 16S rRNA could be detected in any of our constructs (data not shown). (ii) It could be hypothesized that the shorter constructs carrying codons mainly from the MCS contain certain codons or stretches of bases that prevents maximal translation (i.e. rare codons). One method to determine codon bias is to measure the codon adaptive index [CAI, (23)]. By measuring CAI, we observed a slightly lower value for the shorter constructs compared with the longer, when the first 20 codons were analysed. However, all values were high (CAI > 0.68), arguing against an effect of rare codons (data not shown). (iii) Mechanisms involving rare/specific codons or poly-A/multiple CAs close to the SD have been suggested by others (24–28). However, no such codons/stretches of bases could explain the varied expression levels we observe among the constructs (data not shown). The nature of our constructs (using the same insertion site and differing only in the amount of codons inserted) also argues against such mechanisms, since no difference in expression should be observed between the 4, 9 or the 20 codon constructs if any of these mechanisms would apply. (iv) When examining the constructs for the most striking favoured/disfavoured codons (29), no such codons were present within the first 60 bases of our gene-fusions.

We were surprised that the MCS of commercial plasmid vectors harbour these very strong RNA secondary structures. The efficiency of translation would probably be remarkably higher if an MCS expressing a more unstructured RNA sequence would be developed.

Expression of PrfA is subject to several layers of regulation, acting at the transcriptional, translational and at the post-translational level. The reason for this multiple levels of regulation of PrfA is obviously to maintain the level and/or activity of PrfA at an optimal level at each time-point. Absence of PrfA completely attenuates the virulence capability of *L. monocytogenes* and a deregulated PrfA expression leads to increased virulence gene expression during inappropriate conditions (*i.e.* low temperature, (3)). Expression of PrfA has been shown to be controlled at many steps during initiation of translation. First, an RNA thermosensor located within the *prfA* 5′-UTR obstructs binding of the ribosome at low temperatures. Second, a *trans*-acting riboswitch has been shown to down-regulate PrfA translation by binding to the thermosensor at higher temperatures. Third, in this work, we show that maximal translation of PrfA require an unstructured 5′-region of the coding mRNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–5.

ACKNOWLEDGEMENTS

The authors thank Eva Maria Sternkopf Lillebæk for excellent technical assistance.

FUNDING

Magn. Bergvalls Stiftelse (to B.S.), Umeå University (to B.S. and J.J.), the Swedish Research Council grants (K2007-57X-20355-01-3 to B.S. and K2011-56X-15144-08-6 and 621-2009-5677 to J.J.), an ERC starting grant (no 260764 - RNAntibiotics to J.J.). and The Danish Council for Independent Research / Natural Sciences (09-063992 to B.H.K.). Funding for open access charge: Swedish Research Council grant (621-2009-5677).

Conflict of interest statement. None declared.

REFERENCES

1. Dussurget,O., Pizarro-Cerda,J. and Cossart,P. (2004) Molecular determinants of *Listeria monocytogenes* virulence. *Annu. Rev. Microbiol.*, **58**, 587–610.
2. Hamon,M., Bierne,H. and Cossart,P. (2006) *Listeria monocytogenes*: a multifaceted model. *Nat. Rev. Microbiol.*, **4**, 423–434.

3. Johansson,J., Mandin,P., Renzoni,A., Chiaruttini,C., Springer,M. and Cossart,P. (2002) An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell*, **110**, 551–561.
4. Loh,E., Dussurget,O., Gripenland,J., Vaitkevicius,K., Tiensuu,T., Mandin,P., Repoila,F., Buchrieser,C., Cossart,P. and Johansson,J. (2009) A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell*, **139**, 770–779.
5. Stritzker,J., Schoen,C. and Goebel,W. (2005) Enhanced synthesis of internalin A in aro mutants of *Listeria monocytogenes* indicates posttranscriptional control of the *inlAB* mRNA. *J. Bacteriol.*, **187**, 2836–2845.
6. Shen,A. and Higgins,D.E. (2005) The 5′ untranslated region-mediated enhancement of intracellular listeriolysin O production is required for *Listeria monocytogenes* pathogenicity. *Mol. Microbiol.*, **57**, 1460–1473.
7. Wong,K.K., Bouwer,H.G. and Freitag,N.E. (2004) Evidence implicating the 5′ untranslated region of *Listeria monocytogenes actA* in the regulation of bacterial actin-based motility. *Cell. Microbiol.*, **6**, 155–166.
8. Glaser,P., Frangeul,L., Buchrieser,C., Rusniok,C., Amend,A., Baquero,F., Berche,P., Bloecker,H., Brandt,P., Chakraborty,T. *et al.* (2001) Comparative genomics of *Listeria* species. *Science*, **294**, 849–852.
9. Timm,J., Lim,E.M. and Gicquel,B. (1994) *Escherichia coli*-Mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: the pJEM series. *J. Bacteriol.*, **176**, 6749–6753.
10. Sullivan,M.A., Yasbin,R.E. and Young,F.E. (1984) New shuttle vectors for *Bacillus subtilis* and *Escherichia coli* which allow rapid detection of inserted fragments. *Gene*, **29**, 21–26.
11. Mengaud,J., Dramsi,S., Gouin,E., Vazquez-Boland,J.A., Milon,G. and Cossart,P. (1991) Pleiotropic control of *Listeria monocytogenes* virulence factors by a gene that is autoregulated. *Mol. Microbiol.*, **5**, 2273–2283.
12. Toledo-Arana,A., Dussurget,O., Nikitas,G., Sesto,N., Guet-Revillet,H., Balestrino,D., Loh,E., Gripenland,J., Tiensuu,T., Vaitkevicius,K. *et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature*, **459**, 950–956.
13. Geuskens,V., Mhammedi-Alaoui,A., Desmet,L. and Toussaint,A. (1992) Virulence in bacteriophage Mu: a case of trans-dominant proteolysis by the *Escherichia coli* Clp serine protease. *EMBO J.*, **11**, 5121–5127.
14. Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
15. Hofacker,I.L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res.*, **31**, 3429–3431.
16. Kudla,G., Murray,A.W., Tollervey,D. and Plotkin,J.B. (2009) Coding-sequence determinants of gene expression in *Escherichia coli*. *Science*, **324**, 255–258.
17. de Smit,M.H. and van Duin,J. (2003) Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. *J. Mol. Biol.*, **331**, 737–743.
18. Studer,S.M. and Joseph,S. (2006) Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol. Cell.*, **22**, 105–115.
19. Gu,W., Zhou,T. and Wilke,C.O. (2010) A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLoS Comput. Biol.*, **6**, e1000664.
20. Unoson,C. and Wagner,E.G. (2007) Dealing with stable structures at ribosome binding sites: bacterial translation and ribosome standby. *RNA Biol.*, **4**, 113–117.
21. Marzi,S., Myasnikov,A.G., Serganov,A., Ehresmann,C., Romby,P., Yusupov,M. and Klaholz,B.P. (2007) Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell*, **130**, 1019–1031.
22. Sprengart,M.L., Fatscher,H.P. and Fuchs,E. (1990) The initiation of translation in E. coli: apparent base pairing between the 16srRNA and downstream sequences of the mRNA. *Nucleic Acids Res.*, **18**, 1719–1723.
23. Sharp,P.M. and Li,W.H. (1987) The codon Adaptation Index–a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.*, **15**, 1281–1295.
24. Brock,J.E., Paz,R.L., Cottle,P. and Janssen,G.R. (2007) Naturally occurring adenines within mRNA coding sequences affect ribosome binding and expression in *Escherichia coli*. *J. Bacteriol.*, **189**, 501–510.
25. Gonzalez de Valdivia,E.I. and Isaksson,L.A. (2004) A codon window in mRNA downstream of the initiation codon where NGG codons give strongly reduced gene expression in *Escherichia coli*. *Nucleic Acids Res.*, **32**, 5198–5205.
26. Martin-Farmer,J. and Janssen,G.R. (1999) A downstream CA repeat sequence increases translation from leadered and unleadered mRNA in *Escherichia coli*. *Mol. Microbiol.*, **31**, 1025–1038.
27. Sato,T., Terabe,M., Watanabe,H., Gojobori,T., Hori-Takemoto,C. and Miura,K. (2001) Codon and base biases after the initiation codon of the open reading frames in the *Escherichia coli* genome and their influence on the translation efficiency. *J. Biochem.*, **129**, 851–860.
28. Zamora-Romo,E., Cruz-Vera,L.R., Vivanco-Dominguez,S., Magos-Castro,M.A. and Guarneros,G. (2007) Efficient expression of gene variants that harbour AGA codons next to the initiation codon. *Nucleic Acids Res.*, **35**, 5966–5974.
29. Welch,M., Govindarajan,S., Ness,J.E., Villalobos,A., Gurney,A., Minshull,J. and Gustafsson,C. (2009) Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS One*, **4**, e7002.