Quantification of the gene silencing performances of rationally-designed synthetic small RNAs

Ilaria Massaiu^{1,2} · Lorenzo Pasotti^{1,2} · Michela Casanova^{1,2} · Nicolò Politi^{1,2} ·
Susanna Zucca^{1,2} · Maria Gabriella Cusella De Angelis² · Paolo Magni^{1,2}

Received: 16 February 2015 / Revised: 12 June 2015 / Accepted: 27 July 2015 / Published online: 7 August 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Small RNAs (sRNAs) are genetic tools for the efficient and specific tuning of target genes expression in bacteria. Inspired by naturally occurring sRNAs, recent works proposed the use of artificial sRNAs in synthetic biology for predictable repression of the desired genes. Their potential was demonstrated in several application fields, such as metabolic engineering and bacterial physiology studies. Guidelines for the rational design of novel sRNAs have been recently proposed. According to these guidelines, in this work synthetic sRNAs were designed, constructed and quantitatively characterized in *Escherichia coli*. An sRNA targeting the reporter gene RFP was tested by measuring the specific gene silencing when RFP was expressed at different transcription levels, under the control of different promoters, in different strains, and in single-gene or operon architecture. The sRNA level was tuned by using plasmids maintained at different copy numbers. Results demonstrated that RFP silencing worked as expected in an sRNA and mRNA expression-dependent fashion. A mathematical model was used to support sRNA characterization and to estimate an efficiency-related parameter that can be used to compare the performance of

the designed sRNA. Gene silencing was also successful when RFP was placed in a two-gene synthetic operon, while the non-target gene (GFP) in the operon was not considerably affected. Finally, silencing was evaluated for another designed sRNA targeting the endogenous lactate dehydrogenase gene. The quantitative study performed in this work elucidated interesting performance-related and context-dependent features of synthetic sRNAs that will strongly support predictable gene silencing in disparate basic or applied research studies.

Keywords Small RNA · Synthetic biology · Quantitative characterization · Mathematical modelling · Operon · Lactate dehydrogenase

Introduction

Naturally occurring small RNAs (sRNAs) are short non-coding RNAs, typically between 50 and 250 nucleotide long, able to control the expression of target genes in bacteria, predominantly at the post-transcriptional level (Gottesman 2005; Man et al. 2011; Jost et al. 2011; Peterman et al. 2014). Hundreds of sRNAs have been identified in different bacteria so far, especially in *Escherichia coli*. These sRNAs can be sorted in three general categories: sRNAs that have intrinsic catalytic activity or are components of ribonucleoproteins, sRNAs that affect protein activity and sRNAs that regulate gene expression by base-pairing to a target mRNA (Man et al. 2011). The third sRNAs category is the best-characterized and the most abundant in Gram-negative bacteria. The post-transcriptional control system of sRNAs belonging to the latter category is based on a *trans*-acting mechanism, in which sRNAs bind to the 5' untranslated region (5'-UTR)

Electronic supplementary material The online version of this article (doi:10.1007/s11693-015-9177-7) contains supplementary material, which is available to authorized users.

✉ Paolo Magni
paolo.magni@unipv.it

¹ Laboratory of Bioinformatics, Mathematical Modelling and Synthetic Biology, Department of Electrical, Computer and Biomedical Engineering, University of Pavia, via Ferrata 5, 27100 Pavia, Italy

² Centre for Health Technologies, University of Pavia, via Ferrata 5, 27100 Pavia, Italy

or to the translation initiation region (TIR) of single or multiple target messenger RNAs (mRNAs) through imperfect base-pairing (Aiba 2007), although sRNA binding in regions downstream of the TIR have also been reported (Pfeiffer et al. 2009). The regulation of gene expression is carried out, upon binding, by the modulation of translation or transcript stability (Peterman et al. 2014; Wassarman et al. 2001; Masse and Gottesman 2002; Masse et al. 2003). In particular, specific sRNAs are known to change ribosome accessibility, mainly repressing translation, although examples of positive regulation have also been reported (Gottesman 2004; Peterman et al. 2014). Conversely, the final effect of other sRNAs is to change the stability of the target mRNA, mainly by accelerating its degradation (Gottesman 2004; Peterman et al. 2014; Levine et al. 2007). Morita et al. (2006) analyzed the repression effects of two sRNAs, SgrS and RyhB, which experimentally showed to degrade their target transcript. By inhibiting the mRNA degradation machinery of the host strain, the authors showed that SgrS and RyhB could act as translation inhibitors without affecting transcript stability. This double final effect of the investigated sRNAs, i.e., translation repression and mRNA degradation, could be motivated by the necessity to rid the cell of translationally inactive mRNAs (Morita et al. 2006). A major class of sRNAs requires (or is strengthened by) the RNA chaperone Hfq for efficient gene silencing (Aiba 2007; Gottesman 2004; Peterman et al. 2014).

In nature, sRNAs are involved in the regulation of disparate functions in bacteria, such as stress response, outer membrane protein biogenesis, quorum sensing, virulence, iron and sugar metabolism (Gottesman 2004; Man et al. 2011). Due to their importance, several studies have been recently carried out to discover sRNAs, identify their targets and characterize their regulation mechanisms, also with the help of mathematical models. Among all the research investigations, high-throughput analyses have been performed to search for novel sRNAs or specific targets of known sRNAs by microarray studies (Gottesman 2004; Wassarman et al. 2001; Masse et al. 2005) and, more recently, by the sort-seq approach (Peterman et al. 2014). Reporter gene fusions with the initial part of a target gene (including the 5'-UTR) have been adopted to quantitatively study the contribution of sRNAs and their specificity (Urban and Vogel 2007). Mathematical models, generally studied at the steady-state, were developed to quantify the effects of parameter changes, such as mRNA/sRNA levels and their half life (Levine et al. 2007; Shimoni et al. 2007). Such kinetic models were able to capture the observed behaviour of artificially-constructed systems where sRNA and mRNA levels were tuned (Levine et al. 2007; Shimoni et al. 2007). Moreover, the proposed models are sufficiently general to describe the contribution of sRNAs affecting translation and/or transcript degradation (Jost

et al. 2011). One of these models was recently refined to investigate the effects of ribosome binding site (RBS) strength on the RyhB, DsrB and OmrA sRNA efficiency towards their target gene, whose expression was studied via gene fusion and quantitative PCR (Lavi-Itzkovitz et al. 2014). Comparison between model simulations and experimental data demonstrated that, in the context of the investigated sRNAs, increasing translation rate can lead to increased repression (Lavi-Itzkovitz et al. 2014).

Inspired by the features of natural control systems, sRNAs can also play an important role in the design of synthetic biological systems. In metabolic engineering studies, metabolic fluxes towards the target bioproduct can be optimized by the simultaneous expression of heterologous genes, over-expression and down-regulation of endogenous genes. In this case, sRNAs can be used to down-regulate the expression of the target genes involved in the desired pathway. The use of an sRNA-based approach has several advantages over the common gene knockout method: (1) sRNA expression plasmids, which actuate the silencing of target genes, can be incorporated in the host strain by simple bacterial transformation and it makes sRNA systems highly portable to different hosts; (2) several combinations of sRNAs can be simultaneously tested by co-transforming different expression plasmids or assembling multiple cassettes in the same vector; (3) a scalable, sRNA sequence-dependent, repression efficiency can be obtained; (4) sRNAs can be used to down-regulate essential genes, since sRNAs can be placed under the control of inducible promoters (Man et al. 2011).

It is worth noting that other post-transcriptional control systems, such as antisense RNAs (asRNAs), also have many of the advantages described above (Tummala et al. 2003; Nakashima et al. 2006). However, asRNA-based systems are generally characterized by a lower efficiency than sRNAs (Man et al. 2011), although efforts have been recently carried out to improve their activity (Nakashima and Tamura 2009).

Another recently proposed method for programmable silencing of gene expression in bacteria is CRISPR interference (CRISPRi). It uses an engineered clustered regularly interspaced palindromic repeats (CRISPR) pathway, where a customizable single guide RNA (sgRNA) forms a complex with a catalytically inactive Cas9 protein (dCas9) that can bind DNA (Larson et al. 2013; Qi et al. 2013). Differently from sRNAs, which act at post-transcriptional level, CRISPRi relies on transcriptional regulation by steric block of promoter binding or transcription elongation (Larson et al. 2013). Activation of gene expression in bacteria has also been reported via this method, upon dCas9 protein engineering (Bikard et al. 2013). The regulation mechanism of CRISPRi is highly promising for genome-wide control of gene expression and it is

complementary to post-transcriptional regulation elements, like sRNAs and asRNAs. CRISPRi is characterized by high expression modulation efficiency, it has been shown to work in many species, and guidelines for sgRNA rational design have been proposed (Larson et al. 2013; Qi et al. 2013); compared to post-transcriptional element, an intrinsic drawback of CRISPRi is that the selective repression of an individual gene in polycistronic transcript cannot be easily achieved (Choudhary et al. 2015).

Many bacterial genes are organized in operon architecture. The clustering of genes in operons is an important context in *E. coli* and other prokaryotic organisms, allowing to coordinately express proteins that are involved in common processes, while greatly facilitating the ability to efficiently respond to environmental changes. Because transcription and translation are physically coupled in prokaryotes, operons provide a highly efficient method of regulating the transfer of genetic information from DNA to protein (Keene and Tenenbaum 2002). Genes in polycistronic transcripts can be naturally targeted for repression by sRNAs. Intuitively, the silencing mechanism is important when targeting operon genes, since transcript degradation leads to the down-regulation of all the operon genes, while translation repression can specifically target individual genes within an operon. Specific *E. coli* operons were analyzed in different works (Moller et al. 2002; Masse and Gottesman 2002; Masse et al. 2005; Desnoyers et al. 2009). The *sdhCDAB* operon, involved in succinate metabolism, is targeted by the *RyhB* sRNA, which binds the transcript between the first and the second gene of the operon, resulting in mRNA degradation (Masse and Gottesman 2002; Masse et al. 2005). Conversely, the *galETKM* operon, involved in galactose metabolism, is targeted by the *Spot 42* sRNA, which binds the mRNA upstream of the third gene of the operon (*galK*), but it does not result in transcript degradation. Only the *galK* gene is down-regulated by *Spot 42*, which acts as a translator inhibitor (Moller et al. 2002). Other, more complex, operon regulation mechanisms have recently been reported. For example, the *iscRSUA* transcript, involved in the Fe-S clusters biosynthesis, is targeted for degradation by the *RyhB* sRNA, but the first operon gene (*iscR*) is not down-regulated, thanks to a strong repetitive extragenic palindromic secondary structure (between *iscR* and *iscS*) which may protect the gene against ribonucleases degradation (Desnoyers et al. 2009).

Motivated by the attractive features of sRNA-based control systems, after the discovery of natural sRNAs in bacteria many efforts were carried out to design synthetic sRNAs that can repress the desired target genes. Inspired by the natural architecture of the discovered bacterial sRNAs, the synthetic sRNAs are composed of two functional parts: a target-binding sequence and a scaffold

sequence (Man et al. 2011; Sharma et al. 2012; Na et al. 2013). The first part is a sequence complementary to the TIR of its target mRNA, which specifically binds to the target and actuates the gene silencing. The scaffold sequence recruits the RNA chaperone Hfq, a highly abundant protein that facilitates the binding of the sRNA to the target mRNA at a much faster rate than that of the binding of ribosomes (Na et al. 2013).

However, owing to a lack of full understanding of the sRNA silencing mechanism in prokaryotes, the first studies on synthetic sRNA design mainly focused on random screening methods (Man et al. 2011; Sharma et al. 2012). Man et al. (2011) developed a semirational strategy for sRNA design based on the sequence of well-known trans-encoded *E. coli* sRNAs. The target-binding sequence was complementary to the 5'-UTR of the target mRNA and then appropriately adjusted to have a secondary structure with at least two stem loops. The Hfq-binding scaffold sequence and the transcriptional terminator were extracted from a list of well-studied endogenous sRNAs and randomly combined to the target-binding part. These candidate sRNAs were finally filtered according to their secondary structure and a shorter list of candidates was obtained. This method resulted in the design of successful sRNAs, although out of the 16 initially selected candidates, only two repressed the target gene expression by 70 % or more.

Sharma et al. (2012) developed a screening strategy that can identify synthetic sRNAs capable of regulating endogenous genes. They constructed a large library of artificial sRNAs by fusing a randomized antisense domain to a scaffold sequence from four natural sRNAs that interact with the Hfq protein. In order to select the sRNA actually targeting the mRNA of interest, expression plasmids including the random sRNA library were co-transformed with an expression plasmid including a fluorescent reporter gene (GFP) fused with the 5' leader sequence of the mRNA of interest. Fluorescence detection by visual inspection of transformation plates identified the colonies containing the desired sRNAs. The described method enabled the obtainment of sRNAs that repressed the *ompF* target gene by 45- to 145-fold, but the approach required the screening of a large number of clones ($>10^5$) and was characterized by a low probability to find a clone where fluorescence was repressed (0.03 %).

Guidelines for the rational design of customized sRNA were recently proposed by Na et al. (2013). The authors used reporter genes to test different features of sRNA expression systems, by investigating the repression capability as a function of different scaffold sequences, hybridization energy, binding position of sRNA within the transcript and target-binding sequence length. From their investigations, they selected the *MicC* (Urban and Vogel 2007; Chen et al. 2004) sequence as the best scaffold

among four candidates, 24 nucleotides as the optimal length of target-binding sequence and an hybridization energy lower than -20 kcal/mol. In their work, the proposed guidelines were successfully applied to metabolic engineering, demonstrating that complex pathways can be optimized via large libraries of designed sRNAs and such strategy can be easily adopted to search the best producer among a collection of candidate *E. coli* strains.

The aim of our work is to quantitatively evaluate the performance of synthetic sRNAs designed with guidelines proposed by Na and colleagues. We designed the sRFP silencer, which represses the expression of reporter target gene RFP, encoding the Red Fluorescent Protein, to evaluate its performance on different ad-hoc constructed model systems, in two *E. coli* strains, as a function of sRNA and mRNA levels, also with the help of mathematical modelling. Since the operon context has never been quantitatively tested before using designed sRNAs (Yoo et al. 2013), in this work we studied the down-regulation of a target gene in a synthetic operon. Finally, we present data on the silencing of an endogenous gene, *ldhA*, which has a crucial role in the fermentation pathway of *E. coli* and in metabolic engineering studies, by means of another ad-hoc designed sRNA. The quantitative study performed in this work elucidated interesting performance-related and context-dependent features of synthetic sRNAs that have never been investigated before. The obtained results and data will strongly support predictable gene silencing in disparate basic or applied research studies via novel designed sRNAs.

Materials and methods

Strains and plasmids

The *E. coli* TOP10 (Invitrogen) strain was used for cloning. The TOP10 and W (ATCC 9637; Archer et al. 2011) strains were used for quantitative experiments.

We designed the sRFP, sLDH, sACK, sFRD and sPFL synthetic sRNAs, targeting the RFP, *ldhA*, *ackA*, *frdA* and *pflB* genes, respectively, according to the guidelines proposed by Na et al. (2013). The target-binding sequence was designed as the reverse complement of the first 24 bp of the coding sequence included in the target mRNA. The hybridization energy of the target-binding sequence was computed via the UNAFold web server (<http://mfold.rna.albany.edu/>; Markham and Zuker 2005), to verify that it was lower than -20 kcal/mol. The MicC scaffold sequence was included downstream of the target-binding sequence to obtain the final sRNA. This sRNA sequence is placed between the strong promoter P_R upstream and the T1/TE

transcriptional terminator downstream. The genomic sequences of the DH10B (NC_010473.1), closely related to TOP10 and with the same genotype, and the W (NC_017664.1) strains in the NCBI database were used to retrieve the *ldhA*, *ackA*, *frdA* and *pflB* gene sequences. All the genes had identical nucleotides in the initial 24 bp of their coding sequences between the two strains. The sequence of the RFP gene was retrieved from the BBa_E1010 entry in the MIT Registry of Standard Biological Parts (Registry) (<http://partsregistry.org>; Knight 2003).

The sRNA expression cassettes were de-novo synthesized by the GenScript gene synthesis service (Piscataway, NJ, USA). They were designed with the standard BioBrick^{TM} prefix upstream and suffix downstream (Knight 2003) to facilitate their transfer in different plasmid vectors. These cassettes were delivered in the pUC57-Simple shipping vector and they were subsequently transferred, upon EcoRI/PstI digestion, both into the pSB3K3 and into the pSB1A2 BioBrick^{TM} vectors (Shetty et al. 2008).

All the other parts were either physically retrieved from the Registry DNA Distribution 2008–2011 or assembled in this study from existing BioBrick^{TM} parts, by using the BioBrick^{TM} Standard Assembly procedure (Knight 2003).

All the strains were grown in 5 ml of L-broth (LB; Sambrook et al. 1989) at 37 °C, 220 rpm. When required, ampicillin (100 mg/l), kanamycin (20 mg/l) and chloramphenicol (12.5 mg/l) were added to cultures to maintain plasmids. Long-term glycerol stocks, stored at -80 °C, were prepared for all the recombinant strains by mixing 750 μ l of bacterial culture and 250 μ l of sterile 80 % glycerol. Plasmids were extracted from overnight cultures through the NucleoSpin Plasmid kit (Macherey–Nagel). DNA was digested as appropriate, with the EcoRI/XbaI/SpeI/PstI enzymes, and the fragments of interest were extracted from 1 % agarose gel by NucleoSpin Extract II kit (Macherey–Nagel) before proceeding with ligation. DNA-modifying enzymes were purchased from Roche Diagnostics and used according to manufacturer's instructions.

We constructed model systems to quantitatively evaluate the performance and the specificity of the synthetic sRNAs (see Fig. 1). In particular, these systems include synthetic circuits expressing RFP and/or GFP in single gene or operon, which are driven by promoters and RBSs with different strength. All these constructs were placed in the pSB4C5 low-copy plasmid and they were co-transformed with an sRNA expression cassette, placed in the pSB3K3, pSB1A2 or pUC57-Simple plasmid. This expression systems design allows to study genes transcribed/translated at different levels (through promoter/RBS/inducer concentration changes) in combination with sRNAs expressed at different levels (through plasmid copy

Fig. 1 Synthetic circuits used in this study. The BioBrick^{TM} codes are reported in brackets and can be found in the Registry of Standard Biological Parts with their nucleotide sequences. **a** J101-R (BBa_J107029): single-gene cassette for the expression of RFP driven by the constitutive BBa_J23101 promoter, with the BBa_B0034 RBS upstream of the RFP coding sequence. **b** J101-R32 (BBa_K516132): single-gene cassette for the expression of RFP driven by the constitutive BBa_J23101 promoter, with the BBa_B0032 RBS upstream of the RFP coding sequence. **c** Plux-R (BBa_J107032): single-gene cassette for the expression of RFP driven by the inducible P_{lux} promoter, with the BBa_B0034 RBS upstream of the RFP coding sequence. **d** Plux-G (BBa_T9002): single-gene cassette for the expression of GFP driven by the P_{lux} promoter, with the BBa_B0032 RBS upstream of the GFP coding sequence. **e** Plux-RG (BBa_J107042): RFP–GFP operon driven by the P_{lux} promoter,

with the BBa_B0034 and BBa_B0032 RBSs upstream of RFP and GFP, respectively. **f** Plux-GR (BBa_J107043): GFP–RFP operon driven by the P_{lux} promoter, with the BBa_B0032 and BBa_B0034 RBSs upstream of GFP and RFP, respectively. **g** Plux-RG30 (BBa_J107044): RFP–GFP operon driven by the P_{lux} promoter, with the BBa_B0034 and BBa_B0030 RBSs upstream of RFP and GFP, respectively. **h** Plux-G30R (BBa_J107045): GFP–RFP operon driven by the P_{lux} promoter, with the BBa_B0030 and BBa_B0034 RBSs upstream of GFP and RFP, respectively. All the described constructs are present in the pSB4C5 low-copy vector. Curved arrows represent promoters, ovals represent RBSs, straight arrows represent genes and hexagons represent transcriptional terminators. RBS34, RBS32 and RBS30 are the BBa_B0034, BBa_B0032 and BBa_B0030 BioBrick^{TM} RBSs

number changes). A similar experimental design, including a two-plasmid expression system for reporter gene and silencer, respectively, has been used by Levine et al. (2007) and Lavi-Itzkovitz et al. (2014) to characterize the effects of transcription, translation and RNA degradation parameters change. We use copy number change to tune the sRNA level in order to reproduce the same expression system design proposed by Na et al. (2013), which proved to be successful as sRNA production cassette.

The J101-R (Fig. 1a) and J101-R32 (Fig. 1b) constructs have the same constitutive promoter (BBa_J23101) upstream of RFP, but different RBSs (the BBa_B0034 RBS is stronger than BBa_B0032 when placed upstream of RFP; Pasotti et al. 2012). The Plux-R (Fig. 1c) circuit contains RFP driven by the P_{lux} inducible promoter. P_{lux} in the induced state is about eightfold stronger than BBa_J23101 (Zucca et al. 2012). These circuits allow to characterize RFP silencing as a function of mRNA level (J101-R and Plux-R constructs) and RBS strength (J101-R and J101-R32 constructs).

We studied the unspecific silencing by comparing the output of the Plux-R and Plux-G (Fig. 1d) circuits, where the latter includes an inducible expression cassette for GFP, which is not targeted by the designed sRNAs. We also studied unspecific silencing by characterizing the output of

the above constructs in presence of sLDH, sACK, sFRD or sPFL, which do not target RFP and GFP.

The RFP–GFP and GFP–RFP operons (circuits Plux-RG, Plux-GR, Plux-RG30 and Plux-G30R, see Fig. 1e–h) allow the study of the specific and unspecific gene silencing in polycistronic mRNA by red and green fluorescence quantification. The RFP and GFP genes of these circuits are in different positions and the GFP gene is placed under two different RBSs (the BBa_B0030 RBS is stronger than BBa_B0032 when placed upstream of GFP).

Transformation was carried out in TOP10 and W by heat shock at 42 °C.

Both the ampicillin-resistant high-copy plasmids pSB1A2 and pUC57-Simple have a ColE1-based replication origin, but it has a single nucleotide mismatch (according to their sequence in the Registry and in the provided GenScript document, respectively), which could contribute to a different copy number. Finally, the pSB3K3 plasmid has a p15A replication origin and the pSB4C5 plasmid has a pSC101 origin (Shetty et al. 2008).

Fluorescence assays

Recombinant strains were grown in 2-ml tubes at 37 °C, 220 rpm for 16–20 h in 0.5 ml of M9 supplemented

medium (11.28 g/l M9 salts—M6030, Sigma Aldrich, 2 mM MgSO₄, 0.1 mM CaCl₂, 2 g/l casamino acids, 1 mM thiamine hydrochloride and 4 ml/l glycerol; Sambrook et al. 1989) with antibiotics as required, inoculated with a single colony from a streaked selective LB-agar plate (at least 3 independent biological replicates were carried out for each recombinant strain). The grown cultures were 100-fold diluted in 200 µl of M9 in a 96-well microplate. When required, 2 µl of the *N*-3-oxohexanoyl-L-homoserine lactone (HSL) inducer (K3007, Sigma Aldrich) were added to reach the desired final concentration. Unless differently indicated, 100 nM of HSL were used to induce the *P_{lux}* promoter. The microplate was incubated at 37 °C in the Infinte F200 reader (Tecan) and the following kinetic cycle, programmed via the i-control software (Tecan), was carried out: linear shaking 15 s (amplitude 3 mm), wait 5 s, absorbance measurement (600 nm), fluorescence measurement (excitation 485 nm, emission 540 nm for GFP; excitation 535 nm, emission 620 nm for RFP, gain of 50 or 80), sampling time 5 min (Pasotti et al. 2012; Zucca et al. 2013).

Data analysis for fluorescence assays

Matlab R2010a (MathWorks) and Microsoft Excel were used to analyze the absorbance and fluorescence time series to obtain doubling time and average RFP or GFP synthesis rate per cell (S_{cell} ; Kelly et al. 2009; Politi et al. 2014). S_{cell} is expressed in arbitrary units (AU), proportional to the average per-cell protein synthesis rate. In each experiment, the absorbance of M9 without bacteria (background absorbance) and the fluorescence of the TOP10 and W strains without reporter genes (background fluorescence) were measured. The background absorbance time series was subtracted from the absorbance of each culture of interest to obtain a time series (OD_{600}) proportional to bacterial cell density (see Figure S1A–B; Kelly et al. 2009; Politi et al. 2014). Similarly, the RFP fluorescence background (which is not characterized by a relevant OD_{600} -dependent autofluorescence, see Figure S1B–C) time series was subtracted from the raw RFP fluorescence of each culture to yield a time series proportional to the total RFP proteins in the microplate well. Since GFP shows a relevant OD_{600} -dependent autofluorescence (see Figure S1D), a different background subtraction procedure was carried out: a standard curve was obtained by fitting GFP background fluorescence against OD_{600} via linear regression for each of the two strains (see Figure S1E; Mutalik et al. 2013); the fitted standard curve was used to subtract GFP background fluorescence from the raw GFP fluorescence of each culture at the same OD_{600} , yielding a time series proportional to the total GFP proteins in the microplate well (Mutalik et al. 2013). Raw and background-subtracted

absorbance and fluorescence data are shown in Figure S1F–I (for RFP-expressing cultures) and in Figure S1K–N (for GFP-expressing cultures). The slope of the $\ln(OD_{600})$ time series in the OD_{600} range 0.05–0.18 (exponential growth phase) was computed, via linear regression, to calculate the cell growth rate. Doubling time was computed as $\ln(2)$ divided by the slope. A signal proportional to the RFP or GFP synthesis rate per cell was computed as the numeric time derivative of RFP or GFP time series, divided by OD_{600} (see Figure S1J and O for representative data of RFP- and GFP-expressing cultures, respectively). This signal was averaged over the exponential growth phase and the obtained value was divided by the average synthesis rate per cell of a reference culture expressing RFP or GFP, to compute S_{cell} . RFP and GFP reference cultures were recombinant strains (TOP10 or W) bearing an RFP (BBa_I13507) and a GFP (BBa_E0240) expression system under the control of the constitutive BBa_J23101 promoter, in pSB4C5. The silencing capability ($Eff\%$) for a given gene in each of the above illustrated constructs was computed as reported in Eq. 1.

$$Eff\% = 100 * \left(1 - \frac{S_{cell \text{ with silencer}}}{S_{cell \text{ without silencer}}} \right) \quad (1)$$

However, in case S_{cell} with silencer was higher than without silencer, $Eff\%$ was set to zero. Assuming that pSB4C5 is present at 5 copies per cell, the per-cell copy numbers of pSB3K3, pSB1A2 and pUC57-Simple in the TOP10 and W strains were estimated as reported in Eqs. 2–4 (Lutz and Bujard 1997; Zucca et al. 2012).

$$Copy \text{ number } pSB3K3 = \left(\frac{5}{S_{cell4C5}} \right) * S_{cell3K3} \quad (2)$$

$$Copy \text{ number } pSB1A2 = \left(\frac{5}{S_{cell4C5}} \right) * S_{cell1A2} \quad (3)$$

$$Copy \text{ number } pUC57_{Simple} = \left(\frac{5}{S_{cell4C5}} \right) * S_{cellpUC} \quad (4)$$

where $S_{cell4C5}$, $S_{cell3K3}$, $S_{cell1A2}$ and $S_{cellpUC}$ are the S_{cell} values of cultures bearing the J101-R construct (Fig. 1a) in the pSB4C5, pSB3K3, pSB1A2 and pUC57-Simple vectors, respectively, assuming that no metabolic overload affects cells at the highest copy numbers (Zucca et al. 2012).

Lactate dehydrogenase assay

The activity of lactate dehydrogenase (LdhA) was determined by a specific enzymatic assay. LdhA catalyzes the conversion of pyruvate and NADH to lactate and NAD⁺, respectively. The decrease of NADH concentration is measured by absorbance (340 nm) in order to compute the reaction rate, which is proportional to the enzyme concentration in the sample (Gay et al. 1968).

2 ml of LB with 100 mM phosphate buffer and 40 g/l of glucose were inoculated with 5 μ l of recombinant bacteria from a glycerol stock and incubated at 37 °C, 220 rpm for 16–20 h. The grown cultures were 100-fold diluted in 9 ml of the same medium and incubated as before for 4 h. One ml of sample was taken, centrifuged at 13,000 rpm for 1 min and the supernatant was removed. The bacterial pellet was resuspended with 1 ml of Tris–HCl 100 mM pH 7.3, the vial was centrifuged and the supernatant discarded as before. The CellLytic B (Sigma Aldrich) lysis buffer, supplemented with protease inhibitor cocktail, was used to resuspend the pellet, and the vial was incubated at room temperature for 10 min under slow shaking conditions. Cell debris were separated by centrifugation (13,000 rpm, 5 min). 20 μ l of supernatant, which includes the intracellular content, were transferred into the well of a microplate and mixed with 180 μ l of a solution containing Tris–HCl pH 7.3 100 mM, NADH 0.4 mM and sodium pyruvate 10 mM. The 96-well microplate was incubated at 25 °C in the Infinite F200 reader and absorbance (340 nm) was read every 5 min.

The enzymatic activity of the sample in the well was computed by linear regression of the absorbance time series. Since this activity depends from cell lysis efficiency and initial amount of cells in the sample, we computed the specific enzymatic activity by dividing the enzymatic activity of the sample by the milligrams of total proteins extracted during lysis, quantified with the Micro BCA Protein Assay Kit (Thermo Scientific). The specific activity of all the bacterial cultures analyzed is divided by the wild type activity.

Statistical tests

Statistical analysis was performed on S_{cell} values via the Kruskal–Wallis (KW) nonparametric test to evaluate the statistical significance of repression in the assayed conditions. When the KW test detects at least a significantly different S_{cell} value (p value <0.05) among groups, the least significant difference (LSD) method was used to evaluate the significantly different conditions by multiple comparisons. We implemented the test by the *kruskalwallis* Matlab function. In such multiple comparisons, we focused on the significance of silencing (specific or non-specific) of each condition compared to the recombinant strain without sRNA. For this reason, we only evaluated the contexts where S_{cell} was lower than the reference context, by one-sided test. An analogous procedure was used to analyze the statistical significance of specific LdhA enzymatic activity among the tested contexts.

Mathematical modelling

The kinetic model of Eqs. 5–9 was considered (Levine et al. 2007; Shimoni et al. 2007; Canton et al. 2008; Kelly et al. 2009) and a summary of species and parameters is reported in Table 1.

$$\frac{dm}{dt} = \alpha_m - \beta_m \cdot m - k_+ \cdot s \cdot m + k_- \cdot c \quad (5)$$

$$\frac{ds}{dt} = \alpha_s - \beta_s \cdot s - k_+ \cdot s \cdot m + k_- \cdot c \quad (6)$$

$$\frac{dc}{dt} = k_+ \cdot s \cdot m - k_- \cdot c - \beta_c \cdot c \quad (7)$$

$$\frac{di}{dt} = \rho \cdot m - (z + \mu) \cdot i \quad (8)$$

$$\frac{dr}{dt} = z \cdot i - \mu \cdot r \quad (9)$$

The equations above describe the transcription process of the target mRNA (m) and of the sRNA (s), assuming constant transcription rates (α_m and α_s , respectively) and linear degradation rates (β_m and β_s , respectively). The c state variable represents the mRNA–sRNA complex, which is formed upon m and s interaction with kinetic constant k_+ ; the complex releases m and s with kinetic constant k_- and its degradation rate is linear (β_c). The immature (i.e., non-fluorescent; Kelly et al. 2009) protein (i) synthesis process is described by a linear production term ($\rho \cdot m$, where ρ is the translation rate per mRNA unit) and a linear extinction rate ($z + \mu$, where z is the maturation rate to yield the fluorescent form, and μ is the cell growth rate which represents the protein dilution due to cell division). The last equation describes protein maturation, to yield the fluorescent form r of the reporter protein. The described model assumes that RNA degradation rate is much faster than cell growth rate, while protein degradation rate is negligible and cell division is the only responsible of the intracellular protein extinction rate.

Considering the steady-state of the system, $S_{cell} = q \cdot z \cdot \bar{i}$ (where the bar indicates the steady-state) is the experimentally observable variable (Canton et al. 2008; Pasotti et al. 2013), already defined above, where q represents the unit conversion constant between the actual protein synthesis rate per cell and the S_{cell} values (in AU) obtained in the experiments described above. S_{cell} is also proportional to the steady-state mRNA level: $S_{cell} = b \cdot \bar{m}$, where $b = \frac{q \cdot z \cdot \rho}{z + \mu}$.

The solution of the system, which is S_{cell} , can be analytically computed as (Levine et al. 2007):

$$S_{cell} = \frac{1}{2} \cdot \left(a - a_s - a_\lambda + \sqrt{(a_s + a_\lambda - a)^2 + 4 \cdot a \cdot a_\lambda} \right) \quad (10)$$

Table 1 Description of kinetic model species and parameters. Parameters refer to the ordinary differential equation model (Eqs. 5–9)

Species or parameter	Description	Units
m	mRNA per-cell level	Molecules
s	sRNA per-cell level	Molecules
c	mRNA–sRNA complex per-cell level	Molecules
i	Immature reporter protein per-cell level	Molecules
r	Mature reporter protein per-cell level	Molecules
α_m	mRNA transcription rate per cell	Molecules time ⁻¹
α_s	sRNA transcription rate per cell	Molecules time ⁻¹
β_m	mRNA degradation rate	time ⁻¹
β_s	sRNA degradation rate	time ⁻¹
k_+	Kinetic constant for mRNA and sRNA association to form the mRNA–sRNA complex	Molecules ⁻¹ time ⁻¹
k_-	Kinetic constant for mRNA–sRNA complex dissociation	time ⁻¹
β_c	mRNA–sRNA complex degradation rate	time ⁻¹
ρ	Protein translation rate per RNA unit per cell	time ⁻¹
z	Protein maturation rate	time ⁻¹
μ	Cell growth rate	time ⁻¹

This equation describes the S_{cell} value of an RFP expression system in presence of the sRFP silencer. In this equation, $a = \frac{b \cdot \alpha_m}{\beta_m}$ is proportional to the transcription rate of the target mRNA and it is identical to the S_{cell} value in absence of sRNA, $a_s = \frac{b \cdot \alpha_s}{\beta_m}$ is proportional to the transcription rate of sRNA, and $a_\lambda = \frac{b \cdot \lambda}{\beta_m}$, where $\lambda = \frac{\beta_m \cdot \beta_s}{k}$ has been previously defined as the leakage rate (since its value affects the threshold-linear response of the target mRNA, as a function of its transcription rate for a given sRNA level value; Levine et al. 2007) and $k = \frac{\beta_c \cdot k_+}{k_- + \beta_c}$.

The analytical solution of the model allows to study the S_{cell} output value, at the steady-state, of systems including a target gene (RFP) and a specific sRNA (sRFP) for different RFP (a) and sRFP levels (a_s). In this study, the RFP level is tuned by inducing the P_{lux} promoter upstream of the RFP gene through different HSL concentrations (with the Plux-R construct), while the sRFP level is tuned by changing the copy number of the plasmid containing the sRNA expression cassette. Equation 10 was used to fit experimental data (S_{cell} value in the non-repressed condition in the x-axis, S_{cell} value in the repressed conditions tested in the y-axis, for different HSL concentrations) via the *lsqnonlin* Matlab routine. A different a_s parameter value was estimated for each sRNA level tested, while a single a_λ parameter value for all the sRNA levels was estimated, as described in (Levine et al. 2007).

Results

RFP silencing in a single-gene cassette

The constructs with reporter target gene RFP driven by the P_{lux} inducible promoter (Plux-R, Fig. 1c) were tested, both in the TOP10 and W strains, in presence of no silencer and of the sRFP silencer in medium-copy (pSB3K3 vector) and high-copy (pSB1A2 and pUC57-Simple vectors) contexts. The RFP synthesis rate per cell was measured as systems output that reflects gene silencing. Such tests allowed the study of specific silencing as a function of intracellular concentration of sRNA, which is regulated by changing the intracellular copy number of the sRNA expression system.

Results, reported in Fig. 2a, show that the sRFP silencer works as expected in both strains, since RFP is repressed only in presence of its silencer. Statistical analysis of the S_{cell} values showed a significant difference between the condition with the silencer and the condition without the silencer, in both strains and in all conditions (p value < 0.05 , Kruskal–Wallis and multiple comparisons). Figure 2a shows that the silencer represses RFP by up to 92 % in TOP10 and 68 % in the W strain. Repression values were systematically higher in TOP10 than in W, with sRFP in medium-copy giving the lowest Eff% value and the pUC context giving the highest value, for a given strain. Doubling times were similar among the tested conditions for each of the two strains (see Figure S2A).

Fig. 2 Silencing results for RFP or GFP expressed by a single-gene cassette driven by P_{lux} (Plux-R or Plux-G). **a** Specific silencing of the target gene (RFP) via sRFP in TOP10 and W. **b** Unspecific silencing of RFP or GFP via different sRNAs in TOP10 and W. Bars represent the mean S_{cell} value computed on at least three biological replicates. Error bars represent the 95 % confidence intervals of the mean value.

Asterisks indicate that the S_{cell} value in the condition is statistically different from the S_{cell} of the expression cassette without sRNA (Plux-R and Plux-G conditions for RFP and GFP, respectively). Percentages represent the Eff% values. When S_{cell} in a given condition is higher than S_{cell} without sRNA, Eff% value is set to zero

The per-cell copy numbers of pSB3K3, pSB1A2 and pUC57-Simple in the TOP10 and W strains were estimated to investigate the difference of repression values between strains and among the conditions of a given strain. Results are shown in Table 2.

The obtained copy number values for TOP10 strain are comparable to the values of the literature for similar laboratory strains (Lutz and Bujard 1997; Guido et al. 2006), whereas for the W strain data are not available. Copy numbers are systematically higher in the TOP10 strain than in the W strain. Such copy number values are highly correlated with the RFP repression efficiencies in the tested conditions with a Pearson correlation coefficient of 0.98. These results suggest that the silencer copy number is the

main responsible of the repression efficiency variation among the tested sRFP plasmid contexts and strains.

To evaluate if the obtained repression efficiencies were actually due to the specific action of sRFP, we studied the effect of different unspecific silencers on the target genes RFP and GFP driven by the P_{lux} inducible promoter (Plux-R and Plux-G, Fig. 1c, d), both in TOP10 and W. In particular, RFP or GFP repression was tested in presence of: (1) no silencer, (2) the highest-copy number vector without expression cassettes (pUC-RING), and (3) a set of silencers (sLDH, sACK, sFRD and sPFL, see “Materials and methods”), in different copy numbers, designed to target specific genes involved in the *E. coli* fermentation pathway.

Table 2 Estimated copy numbers for the pSB4C5 low-copy (from the literature), pSB3K3 medium-copy, pSB1A2 and pUC57-Simple high-copy vectors

Vector	Replication origin	Copy number in the TOP10 strain	Copy number in the W strain
pSB4C5	pSC101	5 (fixed)	5 (fixed)
pSB3K3	p15A	26	12
pSB1A2	Mutated pMB1	39	16
pUC57-Simple	Mutated pMB1	52	28

Both high-copy replication origins are noted as “Mutated pMB1”, but their sequences are different in a single nucleotide mismatch, therefore they can be considered as different origins, since they can be characterized by a quantitatively different copy number

Results (Fig. 2b) showed that unspecific silencers have a low repression capability towards both the RFP and GFP gene. In particular, the highest unspecific repression values of RFP (31 % in the TOP10 strain and 26 % in the W strain) were obtained when sRNA expression systems are in the high copy number pUC57-Simple vector. On the other hand, the highest repression values observed for GFP were 35 % for the TOP10 strain (sRFP silencer in the pUC57-Simple vector) and 21 % for the W strain (sRFP silencer in the pSB1A2 vector). In general, the entity of the observed unspecific repressions is lower than the specific silencing percent values obtained above (compare Fig. 2a, b) and the highest values corresponded to conditions where silencer is placed in a high copy plasmid. The doubling times of the recombinant strains in the illustrated conditions did not significantly correlate with unspecific repression values (see Figure S2B) and are similar to the ones shown above for specific silencing (Figure S2A). Statistical analysis of the unspecific silencing data showed that no significant difference between the condition without sRNA and the conditions with plasmid-borne sRNA occurs.

Model-based characterization of RFP silencing

A mathematical model, previously developed to describe silencing efficiency as a function of sRNA and mRNA levels, was used for the characterization of the sRFP-dependent RFP gene repression. The sRFP level and the RFP

Fig. 3 RFP silencing as a function of target mRNA and sRNA levels in TOP10 with the Plux-R construct. Data points represent average S_{cell} values in presence of sRFP in three different copy number conditions (corresponding to three different sRFP levels), as a function of a , corresponding to the average S_{cell} in absence of sRFP that is proportional to the mRNA level. Lines represent model fitting. Estimated parameters are reported in Table 3. Average S_{cell} values are computed on at least three biological replicates

transcript level were varied by means of plasmid copy number (as in previous section) and by tuning the P_{lux} promoter transcriptional activity, respectively. The resulting data are shown in Fig. 3 and they were fitted with the steady-state solution of the kinetic model (Eq. 10). The fitted curves showed that the model was able to describe RFP silencing according to different sRNA/mRNA levels, with the trend reported previously (Levine et al. 2007). The estimated model parameters are reported in Table 3. The a_s parameters, corresponding to the sRNA levels for each of the three copy number conditions, were consistent with the estimated plasmid copy number reported in the previous section (see Table 2): the estimated a_s value in the pSB1A2 context was 1.6-fold higher than in the pSB3K3 context, consistent with the data of Table 2 where a 1.5-fold variation is observed. The estimated a_s value in the pUC context was higher than the values of the two other plasmids, as expected, but it was 2.8- and 4.6-fold higher than a_s in pSB1A2 and pSB3K3, while data of Table 2 showed a smaller fold-change (1.3 and 2, respectively). This could be due to saturation phenomena in the measurement of the copy number through RFP; in fact, protein expression may not change in a linear fashion at high per-cell copy numbers and measured values can be underestimated (Zucca et al. 2012).

The obtained a_s and k parameter units depend on our acquisition system (see Methods section) and for this reason they are not immediately comparable with published values. In order to enable such comparisons, we computed the α_s (per DNA copy) and k values in absolute units as $\frac{\text{RNA molecules}}{s}$ and $\frac{1}{\text{nM} \cdot \text{min}}$, respectively. We considered the a_s value in the medium copy context ($\alpha_s = 65$) and we assumed: a transcriptional activity of $0.03 \frac{\text{RNA molecules}}{s}$ per DNA copy for $a = 1$ (corresponding to the activity of the BBa_J23101 promoter; Kelly et al. 2009), plasmid copy numbers of 5 for pSB4C5 and 26 for pSB3K3, an *E. coli* cell volume of $1 \mu\text{m}^3$ (Milo 2013), and a half-life of 6.8 min (Selinger et al. 2003) for both mRNA and sRNA molecules.

We found an sRNA transcription rate of about $0.042 \frac{\text{RNA molecules}}{s}$, which is consistent with the activity of the P_R promoter, previously found to have an about 2.5-fold higher activity than the BBa_J23101 promoter (Pasotti et al. 2012). We found a k value of $0.0007 \frac{1}{\text{nM} \cdot \text{min}}$, which is about 30-fold lower than typical k values found in literature for naturally occurring regulatory RNAs ($0.02 \frac{1}{\text{nM} \cdot \text{min}}$; Levine et al. 2007). This result highlights that, under the hypotheses above, the sRNA designed in this work following the guidelines of Na et al. (2013) resulted to be functional but with a lower binding rate, k , than observed in nature, thus showing a lower repression efficiency (Levine et al. 2007).

Table 3 Definitions and estimated values of model parameters. Parameters refer to the analytical solution of the model (Eq. 10)

Parameter	Description	Estimated value (AU)
a_{λ}	Leakage rate	1.0198
a_{s-3K3}	Transcription rate of sRNA placed in the medium-copy vector (pSB3K3)	7.2256
a_{s-1A2}	Transcription rate of sRNA placed in the high-copy vector (pSB1A2)	11.8460
a_{s-pUC}	Transcription rate of sRNA placed in the high-copy vector (pUC-Simple)	33.1899

RFP silencing in different expression systems

While the previous sections focused on RFP repression when produced by a P_{lux} -driven expression system, here we considered a different expression cassette for RFP, which is driven by the constitutive BBa_J23101 promoter (J101-R construct of Fig. 1a). As above, the sRFP-mediated

silencing was tested in TOP10 and W in different copy number contexts and the RFP expression system was kept in a low-copy plasmid. Results, shown in Fig. 4a, were consistent with the ones obtained for the P_{lux} -driven cassette (see Fig. 2a): silencing efficiencies were copy number- and strain-dependent, with systematically higher repression for TOP10 than W. Again, repression efficiencies were highly correlated with the estimated copy numbers of Table 2 (Pearson correlation coefficient of 0.96). Statistical analysis showed that RFP repression in most of the conditions with the sRFP silencer was significantly different from the sRNA-free condition. In principle, the tested condition is equivalent to setting the target mRNA to 1 (i.e., the S_{cell} value of the J101-R construct) in the mathematical model, while leaving all the other parameters unchanged, since mRNA and sRNA sequences were the same, and the tested copy number context and strains were identical. However, according to the prediction of the described mathematical model (i.e., the S_{cell} values of the silenced systems as a function of mRNA level and for three different sRNA levels), reported in Fig. 3, higher repression efficiencies were expected for this mRNA level (see Fig. 4b). The observed differences could be due to the slightly different mRNA sequences between the BBa_J23101- and the P_{lux} -driven expression cassettes; in fact, the transcription start site (TSS) of the two promoters is different (Wang et al. 2011; Kosuri et al. 2013). For this reason, the mRNA sequence of the BBa_J23101-driven mRNA has the ACTAGAG sequence upstream of the BBa_B0034 RBS, while the corresponding sequence for P_{lux} has 4 additional nucleotides and it is AAATACTAGAG. Despite this structural difference, the local secondary structure free energy was predicted to be the same between the two sequences and the reasons determining the unexpected difference are unclear. Free energy was computed as described in (Kosuri et al. 2013), by analyzing the entire 5'-UTR and 30 nucleotides of the RFP coding sequence, via the UNAFold web server (Markham and Zuker 2005). Doubling times for the conditions tested above are reported in Figure S3 and are consistent with the ones obtained for the other tested systems (see Figure S2A and B).

Fig. 4 Silencing results for RFP expressed by a single-gene cassette driven by BBa_J23101 (J101-R). **a** Specific silencing of the target gene (RFP) via the specific silencing device (sRFP) in TOP10 and W. **b** Experimental S_{cell} results and values predicted by the mathematical model in the different copy number conditions of sRFP in TOP10. Bars represent the mean S_{cell} value computed on at least three biological replicates. Error bars represent the 95 % confidence intervals of the mean value. Asterisks indicate that the S_{cell} value in the condition is statistically different from the S_{cell} of the expression cassette without sRNA (J101-R). Percentages represent the Eff% values

A similar RFP expression system (J101-R32 construct of Fig. 1b), which has the BBa_B0032 RBS upstream of RFP instead of the stronger BBa_B0034 RBS, was tested in the TOP10 strain. Results (see Figure S4) demonstrated that sRFP was also functional in a different RBS context and the quantitative repression values were similar to the ones obtained with J101-R. Since the RBS sequence has been recently shown to exert a complex effect, the tested context of J101-R32 could not be used to draw strong conclusions on the RBS-dependent functioning of rationally designed sRNAs and further investigations are needed.

The obtained results showed that an sRNA designed with the guidelines of Na et al. (2013) can work in several contexts (different promoters and RBSs for the target gene, and different strains), with qualitatively expected strain and copy number dependence, although the precise repression values could not be predicted, thus highlighting the need for additional studies and the importance of evaluating sRNA efficiency on different measurement constructs.

Silencing of a target gene in a synthetic operon

We used synthetic two-gene operons, including RFP and GFP under the control of the P_{lux} promoter, to complete the characterization of the sRFP silencer. Specifically, the repression capability of the sRNA designed in this work was evaluated when targeting a specific gene present in a polycistronic transcript. The Plux-RG and Plux-GR constructs (Fig. 1e, f), in a low-copy plasmid, were used as model systems in the TOP10 and W strains, and the sRFP expression system, in pSB3K3, pSB1A2 or pUC57-Simple, was co-transformed. Red and green fluorescence signals were simultaneously quantified to study the protein synthesis rate for the target and non-target gene, respectively. The fluorescence acquisition system used in this study was previously characterized and a negligible crosstalk was found to occur between the red and green fluorescence signals (Pasotti et al. 2012). Considering the TOP10 strain, as observed for single-gene cassettes, RFP repression correctly worked (although statistical differences were detected only for the pSB1A2 and pUC57-Simple contexts in both operon systems) and it was dependent on the copy number of the specific sRNA (see Fig. 5a). Such experiments showed that RFP repression occurred when the RFP target gene was present as both the first (Plux-RG construct) and the second (Plux-GR construct) operon gene. In particular, repression values were systematically higher when RFP was present in the second position. A direct comparison among repression values in operons and single-gene cassette contexts is not trivial to carry out, since the stability of mRNA molecules with different sequence can be different (Levin-Karp et al. 2013). GFP was expected not to be repressed as the RFP target protein.

Unfortunately, the GFP signal could not be detected for Plux-RG. The operon context is known to be highly unpredictable (Pasotti and Zucca 2014; Levin-Karp et al. 2013), thus preventing synthetic biological systems designers to infer the translation efficiency of operon genes when their order is changed. GFP could be successfully detected when present as the first gene of the operon (Plux-GR construct). Results showed that an sRFP expression-dependent repression occurred for GFP, although it was much smaller than the one observed for RFP (see Fig. 5a). In particular, while RFP in Plux-GR was repressed by 73, 86 and 93 % in the pSB3K3, pSB1A2 and pUC57-Simple contexts, a repression was observed for GFP in the pSB1A2 and pUC57-Simple conditions (15 and 34 % respectively), although only the latter was found to be statistically significant.

As in the case of single-gene cassettes, the analysis of unspecific silencing was carried out by testing a set of sRNAs, in different plasmids, targeting genes that were different from RFP and GFP. The GFP–RFP operon (Plux-GR construct) was considered for the unspecific silencing study. Results (see Fig. 5b) showed that RFP production was repressed up to 28 %, in the sLDH-pUC context, although a statistically significant RFP repression was not detected for any of the tested conditions. On the other hand, GFP was significantly repressed, up to 47 %, in the two tested conditions where an sRNA (sLDH and sFRD) expression cassette was present in the pUC57-Simple vector.

The obtained results indicate that the RFP target gene is specifically repressed also in operon context, while the non-target gene in the operon mRNA was not affected. According to the unspecific silencing data (Fig. 5b), the observed repression of GFP was most probably due to the metabolic overload of the host strain, caused by the presence of two plasmids with an operon and an sRNA expression cassettes. The observation of a considerable repression of similar entity for both RFP (27–28 %) and GFP (47 %, statistically significant) by unspecific silencers only in the conditions in which sRNAs are expressed in the pUC57-Simple plasmid supports this statement. Doubling times, reported in Figure S5, were consistent with the ones reported for single-gene cassettes. Moreover, since RFP is efficiently repressed while GFP is not, the data suggest that the designed sRFP silencer acts as a repressor of protein synthesis, not increasing the mRNA decay rate like other natural sRNAs, although sRNAs can always affect target mRNA stability.

The GFP–RFP operon (Plux-GR construct) was also tested in the W strain in presence of the specific RFP target gene repressor, sRFP, in pSB3K3, pSB1A2 and pUC57-Simple, or with unspecific sRNAs. Specific silencing results (see Figure S6A) showed that RFP

Fig. 5 Silencing results for RFP and GFP expressed by the Plux-RG and Plux-GR synthetic operons in the TOP10 strain.

a Silencing of the target gene (RFP) and the non-target gene (GFP) via the silencing device sRFP. **b** Unspecific silencing of RFP and GFP, in the Plux-GR construct, via different sRNAs. Bars represent the mean S_{cell} value computed on at least three biological replicates. Error bars represent the 95 % confidence intervals of the mean value. Asterisks indicate that the S_{cell} value of RFP or GFP in the condition is statistically different from the S_{cell} of the operon without sRNA (Plux-RG or Plux-GR). Percentages represent the Eff% values. When S_{cell} in a given condition is higher than S_{cell} without sRNA, Eff% value is set to zero

silencing in operon also works in W, with similar repression efficiency to the single-gene context (see Figs. 2a, 4a). Repression values of RFP were systematically lower than in TOP10 tested with the same plasmid (see Fig. 5a), as expected from the lower sRNA plasmid copy number in the W strain. While RFP is repressed in a copy number-dependent fashion, as expected, reaching up to 62 % repression, GFP is never repressed by sRFP. GFP expression unexpectedly increased up to 1.7-fold when the operon was co-transformed with the sRFP-pUC context, compared to the operon without sRNA expression cassettes. Unspecific silencing experiments (see Figure S6B) showed that RFP and GFP were not repressed by sRNAs different from sRFP. However, these data showed a highly variable RFP and GFP expression, which were highly correlated. Importantly, doubling times analysis showed that in all the conditions with Plux-GR in W growth is clearly slower than in conditions with single-gene cassettes (see Figure S7). This slow growth occurs even when the operon was tested without sRNAs, demonstrating that the operon

itself is responsible of the high doubling time and this was not due to the presence of a co-transformed sRNA expression cassette. This effect suggests a metabolic burden of recombinant strains with the operon. The highly variable RFP and GFP expression may be explained by the metabolic burden of the strains in such conditions, which could result in copy number variation of the medium- and high-copy number plasmids, as previously reported (Lee et al. 2011).

Overall, the results obtained in W were consistent with the ones obtained in TOP10 and confirmed the conclusions drawn above.

Finally, we attempted to overcome the GFP detection limit problem in the RFP–GFP operon (Plux-RG) by constructing and studying novel operons with a stronger RBS (BBa_B0030 instead of BBa_B0032) upstream of GFP (Plux-RG30 and Plux-G30R constructs of Fig. 1g, h). Unfortunately, they resulted in slow, highly variable doubling times (see Figure S8) and significant unspecific silencing (see Figure S9). For this reason, the obtained results cannot be considered to draw robust conclusions.

Silencing of the endogenous lactate dehydrogenase

One of the sRNAs used to evaluate the RFP and GFP unspecific silencing, sLDH, was used to study the specific silencing of the endogenous *ldhA* gene, encoding for a lactate dehydrogenase (LdhA) involved in the fermentation pathway of *E. coli*.

The change of LdhA activity in presence of the sLDH silencer was studied through enzymatic assay (see Methods) in the TOP10 and W strains. In this case, the sRFP silencer was used as unspecific sRNA to evaluate the non-specific repression of LdhA activity. Both sLDH and sRFP were tested in the high-copy number pUC57-Simple vector. Results (see Fig. 6) showed that sLDH significantly repressed LdhA activity, with 50 and 72 % repression values in TOP10 and W, respectively. LdhA repression was very low and not statistically significant (14 and 16 %, respectively) for TOP10 and W with the unspecific silencer sRFP. Such results demonstrated that the sLDH silencer, designed in this study according to the guidelines of Na et al. (2013), is functional. The observed LdhA repression difference between the TOP10 and W strains was unexpected, since the pUC57-Simple plasmid is maintained at higher copy number in TOP10 than in W. However, differences in the *ldhA* gene expression and regulation between the two strains, not investigated in this work, can occur and might explain the observed repression values. Such differences can be due to differences in nucleotide sequences of the *ldhA* gene or promoter in the genomes of

TOP10 and W, even if the 24-bp binding sequence of sLDH is identical.

Discussion

In this work, the silencing capability of sRNAs designed with recently proposed guidelines was characterized. In particular, key features that are typically investigated in quantitative studies on natural sRNA have been herein measured to evaluate the performance of synthetic sRNAs in several contexts and to enable the comparison with natural sRNAs. Since synthetic sRNAs designed with the guidelines of Na et al. (2013) have never been tested in different contexts, such as different target mRNA/sRNA levels, when the target gene is in operon architecture and in gene expression cassettes driven by different promoters (Yoo et al. 2013), this study is of wide importance in the bottom-up design of artificial sRNAs. Our study is mainly focused on quantitative performance evaluation for a synthetic sRNA targeting a reporter gene, but data on another sRNA, targeting a gene of interest in metabolic engineering, are also reported.

A synthetic sRNA targeting the reporter gene RFP, here called sRFP, was designed and used in most of the performed experiments. sRFP was expressed at different levels by tuning the copy number of the plasmid bearing the sRNA expression system, while the target gene was produced via different constitutive or inducible expression

Fig. 6 Lactate dehydrogenase assay results for sLDH characterization in TOP10 and W. Specific enzymatic activity of LdhA in the strain without sRNA, in the strain with an unspecific sRNA (sRFP) and in the strain with the sRNA targeting LdhA (sLDH). Bars represent the mean activity value computed on at least three

biological replicates. Error bars represent the 95 % confidence intervals of the mean value. Asterisks indicate that the value in the condition is statistically different from the value without sRNA (first bar). Percentages represent the Eff% values computed on activity values

systems. Statistical analysis was carried out for all the performed experiments to highlight the conditions where significant repression was present, compared to the recombinant strain without sRNA. Unspecific silencing analysis was also carried out to decouple specific gene silencing from other non-specific silencing mechanisms. To this aim, sRNAs designed to target genes different from RFP were used, as well as the GFP gene, which is not targeted by sRFP and is easily detectable.

When RFP was present in a single-gene expression cassette, driven by the P_{lux} promoter at full induction, sRFP showed to work as expected in two different *E. coli* strains (TOP10 and W) and repressed red fluorescence in a copy number-dependent fashion, reaching silencing levels up to 92 %, in the highest copy number condition (pUC57-Simple plasmid in TOP10) of sRFP. The same construct (called Plux-R) was used to study RFP silencing for different mRNA (tuned via P_{lux} induction with HSL) and sRNA (tuned via plasmid copy number, as above) levels and results showed the expected mRNA-dependent trend (Levine et al. 2007). By using a previously proposed kinetic model of target gene silencing with sRNA, we fitted the experimental data and estimated the binding affinity parameter of sRFP, under different assumptions and considering the Plux-R construct. It resulted to be about 30-fold lower than the one of RyhB, an extensively studied natural sRNA involved in iron metabolism in *E. coli*.

When RFP was constitutively expressed via a single-gene cassette driven by the BBa_J23101 promoter, sRFP also worked as expected in both strains, but the repression values were lower than the ones predicted by the mathematical model that was trained on data from Plux-R. Such observed deviation could be due to the different target mRNA sequence, which is 4 nucleotide longer in transcripts produced by P_{lux} than by BBa_J23101.

Synthetic operons including the target RFP gene and a non-target reporter gene (GFP) were used to study gene silencing, via sRFP, in polycistronic transcripts. RFP repression successfully worked, in an sRNA copy number-dependent fashion, and it was higher when RFP was present as the second operon gene. This result could be due to different mRNA decay rates of the two operons, which affects the steady-state level, or by the mRNA folding which differently exposes the binding sequence. RFP repression level reached values up to 93 % in the highest copy number condition (pUC57-Simple plasmid in TOP10), while the non-target gene, GFP, was not considerably repressed. Experimental data of operon systems suggested that sRFP silencing acts only at the translation level, not by enhancing the decay rate of the whole transcript. GFP signal could not be detected in one of the tested operons (Plux-RG), while in the other operon (Plux-GR)

specific and unspecific silencing could be fully studied. The latter operon was also successfully tested in the W strain, although it caused a slow growth for the host. Additional operons were also constructed (Plux-RG30 and Plux-G30R) with a stronger RBS upstream of GFP, than Plux-RF and Plux-GR. However, although both of them worked as expected, they resulted in slow growth and could not be used to draw sound conclusions, since unspecific silencing was considerable probably due to the metabolic burden exerted by the operon.

Overall, the obtained results on reporter genes demonstrated the importance of the target sequence that could affect gene silencing, and the difficulty of characterizing gene silencing in operon via ad-hoc constructed model systems.

Experimental data were also presented for a second sRNA, targeting the endogenous *ldhA* gene in the TOP10 and W strains. Repression efficiency was 50 and 72 %, respectively, while unspecific silencing, via sRFP, did not result in significant repression values, thus confirming that another rationally designed silencer works as expected. In this case, the quantitative repression values obtained could be explained by measuring the expression level of the target *ldhA* gene.

The rational design of sRNAs with predictable performance is a key feature in synthetic biology and the guidelines proposed by Na et al. can be successfully used without relying on trial-and-error searches. Quantitative characterization studies, like the one proposed in this work, will strongly support the predictability of sRNA performances in different contexts. Although the measured variables can have an sRNA- and target gene-specific behaviour, the reported procedure and results support the future characterization of novel sRNAs, confirm the effectiveness of the design via the used guidelines, and elucidate quantitative performance-related and context-dependent features never investigated before for such synthetic silencers. Synthetic sRNAs will enable to face different problems in synthetic biology, such as the simultaneous silencing of different pathways in metabolic engineering studies (Na et al. 2013), the silencing of essential or non-essential genes for bacterial physiology research studies (Man et al. 2011; Sharma et al. 2013) and the tuning of synthetic circuits (Na et al. 2013) to engineer repression systems with low-fluctuations or noise in the regulation of target proteins (Levine et al. 2007), which is an important feature to control cell-to-cell variability (Politi et al. 2015; Zucca et al. 2015).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aiba H (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* 10(2):134–139
- Archer CT, Kim JF, Jeong H, Park JH, Vickers CE, Lee SY, Nielsen LK (2011) The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC Genomics* 12:9
- Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res* 41(15):7429–7437
- Canton B, Labno A, Endy D (2008) Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol* 26:787–793
- Chen S, Zhang A, Blyn LB, Storz G (2004) MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J Bacteriol* 186(20):6689–6697
- Choudhary E, Thakur P, Pareek M, Agarwal N (2015) Gene silencing by CRISPR interference in mycobacteria. *Nat Commun* 6:6267
- Desnoyers G, Morissette A, Prevost K, Masse E (2009) Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J* 28(11):1551–1561
- Gay RJ, McComb RB, Bowers GN (1968) Optimum reaction conditions for human lactate dehydrogenase isoenzymes as they affect total lactate dehydrogenase activity. *Clin Chem* 14(8):740–753
- Gottesman S (2004) The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu Rev Microbiol* 58:303–328
- Gottesman S (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet* 21(7):399–404
- Guido NJ, Wang X, Adalsteinsson D, McMillen D, Hasty J, Cantor CR, Elston TC, Collins JJ (2006) A bottom-up approach to gene regulation. *Nature* 439(7078):856–860
- Jost D, Nowojewski A, Levine E (2011) Small RNA biology is systems biology. *BMB Rep* 44(1):11–21
- Keene JD, Tenenbaum SA (2002) Eukaryotic mRNPs may represent posttranscriptional operons. *Mol Cell* 9(6):1161–1167
- Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, Czar MJ, de Mora K, Gliberman AL, Monie DD, Endy D (2009) Measuring the activity of BioBrick promoters using an in vivo reference standard. *J Biol Eng* 3:4
- Knight T (2003) Idempotent vector design for standard assembly of BioBricks. MIT DSpace 2003. <http://hdl.handle.net/1721.1/21168>
- Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y, Arkin AP, Endy D, Church GM (2013) Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *PNAS* 110(34):14024–14029
- Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS (2013) CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* 8(11):2180–2196
- Lavi-Itzkovitz A, Peterman N, Jost D, Levine E (2014) Quantitative effect of target translation in small RNA efficacy reveals a novel mode of interaction. *Nucleic Acids Res* 42(19):12200–12211
- Lee TS, Krupa RA, Zhang F, Hajimorad M, Holtz WJ, Prasad N, Lee SK, Keasling JD (2011) BglBrick vectors and datasheets: a synthetic biology platform for gene expression. *J Biol Eng* 5:12
- Levine E, Zhang Z, Kuhlman T, Hwa T (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* 5(9):e229
- Levin-Karp A, Barenholz U, Bareia T, Dayagi M, Zelcbuch L, Antonovsky N, Noor E, Milo R (2013) Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth Biol* 2(6):327–336
- Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 25(6):1203–1210
- Man S, Cheng R, Miao C, Gong Q, Gu Y, Lu X, Han F, Yu W (2011) Artificial trans-encoded small non-coding RNAs specifically silence the selected gene expression in bacteria. *Nucleic Acids Res* 39(8):e50
- Markham MR, Zuker M (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res* 33:W577–W581
- Masse E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *PNAS* 99(7):4620–4625
- Masse E, Escorcia FE, Gottesman S (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* 17(19):2374–2383
- Masse E, Vanderpool CK, Gottesman S (2005) Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187(20):6962–6971
- Milo R (2013) What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* 35(12):1050–1055
- Moller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 16(13):1696–1706
- Morita T, Mochizuki Y, Aiba H (2006) Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *PNAS* 103(13):4858–4863
- Mutalik VK, Guimaraes JC, Cambray G, Mai QA, Christoffersen MJ, Martin L, Yu A, Lam C, Rodriguez C, Bennett G, Keasling JD, Endy D, Arkin AP (2013) Quantitative estimation of activity and quality for collections of functional genetic elements. *Nat Methods* 10(4):347–353
- Na D, Yoo SM, Chung H, Park H, Park JH, Lee SY (2013) Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nat Biotechnol* 31(2):170–174
- Nakashima N, Tamura T (2009) Conditional gene silencing of multiple genes with antisense RNAs and generation of a mutator strain of *Escherichia coli*. *Nucleic Acids Res* 37(15):e103
- Nakashima N, Tamura T, Good L (2006) Paired termini stabilize antisense RNAs and enhance conditional gene silencing in *Escherichia coli*. *Nucleic Acids Res* 34(20):e138
- Pasotti L, Zucca S (2014) Advances and computational tools towards predictable design in biological engineering. *Comput Math Methods Med* 2014:369681. doi:10.1155/2014/369681
- Pasotti L, Politi N, Zucca S, Cusella De Angelis MG, Magni P (2012) Bottom-up engineering of biological systems through standard bricks: a modularity study on basic parts and devices. *PLoS ONE* 7(7):e39407
- Pasotti L, Zucca S, Magni P (2013) Modelling for synthetic biology. In: Modeling methodology for physiology and medicine: second edition. Elsevier, pp 545–564. doi:10.1016/B978-0-12-411557-6.00023-9
- Peterman N, Lavi-Itzkovitz A, Levine E (2014) Large-scale mapping of sequence-function relations in small regulatory RNAs reveals plasticity and modularity. *Nucleic Acids Res* 42(19):12177–12188
- Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16(8):840–846
- Politi N, Pasotti L, Zucca S, Casanova M, Micoli G, Cusella De Angelis MG, Magni P (2014) Half-life measurements of chemical inducers for recombinant gene expression. *J Biol Eng* 8:5

- Politi N, Pasotti L, Zucca S, Magni P (2015) Modelling the effects of cell-to-cell variability on the output of interconnected gene networks in bacterial populations. *BMC Syst Biol* 9(Suppl 3):S6. doi:10.1186/1752-0509-9-S3-S6
- Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152(5):1173–1183
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Selinger DW, Saxena RM, Cheung KJ, Church GM, Rosenow C (2003) Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res* 13(2):216–223
- Sharma V, Yamamura A, Yokobayashi Y (2012) Engineering artificial small RNAs for conditional gene silencing in *Escherichia coli*. *ACS Synth Biol* 1(1):6–13
- Sharma V, Sakai Y, Smythe KA, Yokobayashi Y (2013) Knockdown of *recA* gene expression by artificial small RNAs in *Escherichia coli*. *Biochem Biophys Res Commun* 430(1):256–259
- Shetty R, Endy D, Knight T (2008) Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* 2:5
- Shimoni Y, Friedlander G, Hetzroni G, Niv G, Altuvia S, Biham O, Margalit H (2007) Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* 3:138
- Tummala SB, Welker NE, Papoutsakis ET (2003) Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. *J Bacteriol* 185(6):1923–1934
- Urban JH, Vogel J (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res* 35(3):1018–1037
- Wang B, Kitney RI, Joly N, Buck M (2011) Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat Commun* 2:508
- Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* 15(13):1637–1651
- Yoo SM, Na D, Lee SY (2013) Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. *Nat Protoc* 8(9):1694–1707
- Zucca S, Pasotti L, Mazzini G, De Angelis MG, Magni P (2012) Characterization of an inducible promoter in different DNA copy number conditions. *BMC Bioinform* 13(Suppl 4):S11
- Zucca S, Pasotti L, Politi N, Cusella De Angelis MG, Magni P (2013) A standard vector for the chromosomal integration and characterization of BioBrick™ parts in *Escherichia coli*. *J Biol Eng* 7:12
- Zucca S, Pasotti L, Politi N, Casanova M, Mazzini G, Cusella De Angelis MG, Magni P (2015) Multi-faceted characterization of a novel LuxR-repressible promoter library for *Escherichia coli*. *PLoS ONE* 10(5):e0126264