

Engineered CRISPR-Cas9 system enables noiseless, fine-tuned and multiplexed repression of bacterial genes

Antoine Vigouroux^{1,2}, Enno Oldewurtel², Lun Cui¹, Sven van Teeffelen^{2*} and David Bikard^{1*}

1. Synthetic Biology laboratory, Institut Pasteur, Paris, France

2. Morphogenesis and Microbial Growth, Institut Pasteur, Paris, France

*Correspondence should be addressed to:

Sven van Teeffelen: sven.van-teeffelen@pasteur.fr

David Bikard: david.bikard@pasteur.fr

Abstract

Over the past few years, tools that make use of the Cas9 nuclease have led to many breakthroughs, including in the control of gene expression. The catalytically dead variant of Cas9 known as dCas9 can be guided by small RNAs to block transcription of target genes, in a strategy also known as CRISPRi. Here, we reveal that the level of complementarity between the guide RNA and the target controls the rate at which dCas9 successfully blocks the RNA polymerase. We use this mechanism to precisely and robustly reduce gene expression by defined relative amounts. We demonstrate broad applicability of this method to the study of genetic regulation and cellular physiology. First, we characterize feedback strength of a model auto-repressor. Second, we study the impact of copy-number variations of cell-wall synthesizing enzymes on cell morphology. Finally, we demonstrate that this system can be multiplexed to obtain any combination of fractional repression of two genes.

Introduction

A powerful tool to investigate genes and their regulation is to vary their expression level and investigate the response of the cell. Recently, it has been shown that genes can be knocked down from their native locus to varying degrees using CRISPR technology^{1,2}. The RNA-guided Cas9 nuclease from *Streptococcus pyogenes* can be easily reprogrammed to bind any position of interest on the chromosome, with the requirement of an “NGG” protospacer adjacent motif (PAM). The catalytic mutant dCas9 is unable to cleave target DNA, but still binds DNA strongly. It is able to block transcription initiation when binding a promoter region as well as transcription elongation when targeting downstream of a promoter.

Target search begins by probing DNA for the presence of a PAM followed by DNA melting and RNA strand invasion^{3,4}. This last step is extremely sensitive to mutations in the PAM-proximal region known as the seed sequence. Chromatin immunoprecipitation and sequencing assays (ChIP-seq) have revealed that several mismatches in the PAM-distal region can be tolerated by dCas9 whose binding can be detected with as little as 5 base pairs (bp) of homology between the seed sequence and the target^{5,6}. These observations are consistent with CRISPRi assays performed in bacteria, which show the ability to reduce transcription of a target gene with as little as 8 bp of homology in the PAM-proximal region¹. Accordingly, the degree of gene repression can be controlled quantitatively in two different ways: First, by changing the level of dCas9 expression from an inducible promoter, which then impacts the probability of dCas9 binding to target DNA. This has recently been demonstrated in *Bacillus subtilis* where dCas9 was placed under the control of a xylose-inducible promoter⁷, as well as in an *E. coli* strain modified to enable tunable control of expression from a P_{BAD} promoter⁸. Second, by introducing mismatches between the guide RNA and the target DNA, as demonstrated in *E. coli*¹. While a perfectly matched guide RNA leads to very strong repression, decreasing complementarity in the PAM-distal region progressively reduces the repression strength¹.

Here we compare these two repression strategies by characterizing the properties of dCas9-mediated repression at the single-cell level. This enables us to propose a novel physical model of dCas9-mediated repression. It was previously assumed that decreased levels of guide RNA complementarity would decrease repression strength by virtue of reduced occupancy of the target by dCas9⁹. Here, we demonstrate a different mechanism: If the target is inside an open reading frame (ORF), complementarity determines the rate at which RNA polymerase (RNAP) kicks out dCas9 during transcription attempts. If dCas9 levels are high enough to saturate the target this mechanism alone determines repression strength. This leads to desirable properties: First, relative repression strength is independent of native expression levels. Second, repression does not add any extrinsic noise to gene expression. On the contrary, tuning gene expression by changing the level of dCas9 expression is inherently noisy and depends on the promoter strength of the target.

We propose to use complementarity-based CRISPR knock-down in combination with fluorescent-protein reporters inserted upstream of a gene of interest to

precisely and robustly control its expression. The use of reporter gene fusions rather than direct targeting of the gene of interest yields a predictable repression fold as characterized in this study and provides an easy way to monitor expression levels in single cells. We demonstrate the versatility of our approach using two examples: First, the accurate control of the rate at which the RNAP kicks out dCas9 enables us to quantify the degree of feedback in a model auto-repressor by measuring how much actual gene expression differs from the controlled rate. Second, we take advantage of the ability to obtain a precise degree of repression during steady-state growth to investigate the impact of expression level of an operon coding for two essential cell-wall synthesis enzymes of the 'rod' complex, PBP2 and RodA. Finally, we demonstrate that this system can be easily and robustly multiplexed to obtain any combination of the fractional repression of two genes.

Results

Varying levels of guide RNA-target complementarity enables noiseless control of gene expression

To quantify how CRISPR/dCas9 modulates gene expression at the single-cell level, we integrated expression cassettes for two constitutively expressed reporters, *sfgfp* coding for the superfolder green fluorescent protein (GFP) and *mCherry* coding for a red fluorescent protein (RFP) at two different chromosomal loci of *E. coli* strain MG1655. To repress either of these genes using CRISPR knock-down we integrated the *dcas9* gene from *S. pyogenes* under a P_{tet} promoter, inducible by the addition of anhydrotetracycline (aTc)². We then guided the dCas9 protein to target the coding strand of GFP- and RFP-coding ORFs using a constitutively expressed CRISPR array coding for the guide RNAs and the necessary tracrRNA, which form a complex together with dCas9¹⁰.

In this system, repression strength can be tuned in two different ways: either by modulating dCas9 expression level using different aTc concentrations, or by modulating spacer complementarity to the target gene using different numbers of mismatches at the 5' side of the spacer (Fig. 1, a and b). We employed these two different strategies to repress GFP by different amounts and measured GFP concentration at the single cell level by high-throughput microscopy. As expected, average GFP levels decreased with increasing aTc concentration or increasing spacer complementarity. However, the distributions of single-cell GFP concentrations differed significantly between the two different modes of repression modulation. Specifically, using a perfectly matched guide RNA and varying aTc concentrations led to large cell-to-cell fluctuations, presumably due to variations in dCas9 levels. On the contrary, using high levels of dCas9 expression and varying the degree of guide RNA complementarity maintained the noise (standard deviation over the mean) of single-cell GFP concentration almost constant (Fig. 1, c-d and Supplementary Fig. 1). Noise increased only for very high repression strength, presumably largely due to fluorescence measurement error and possibly due to the

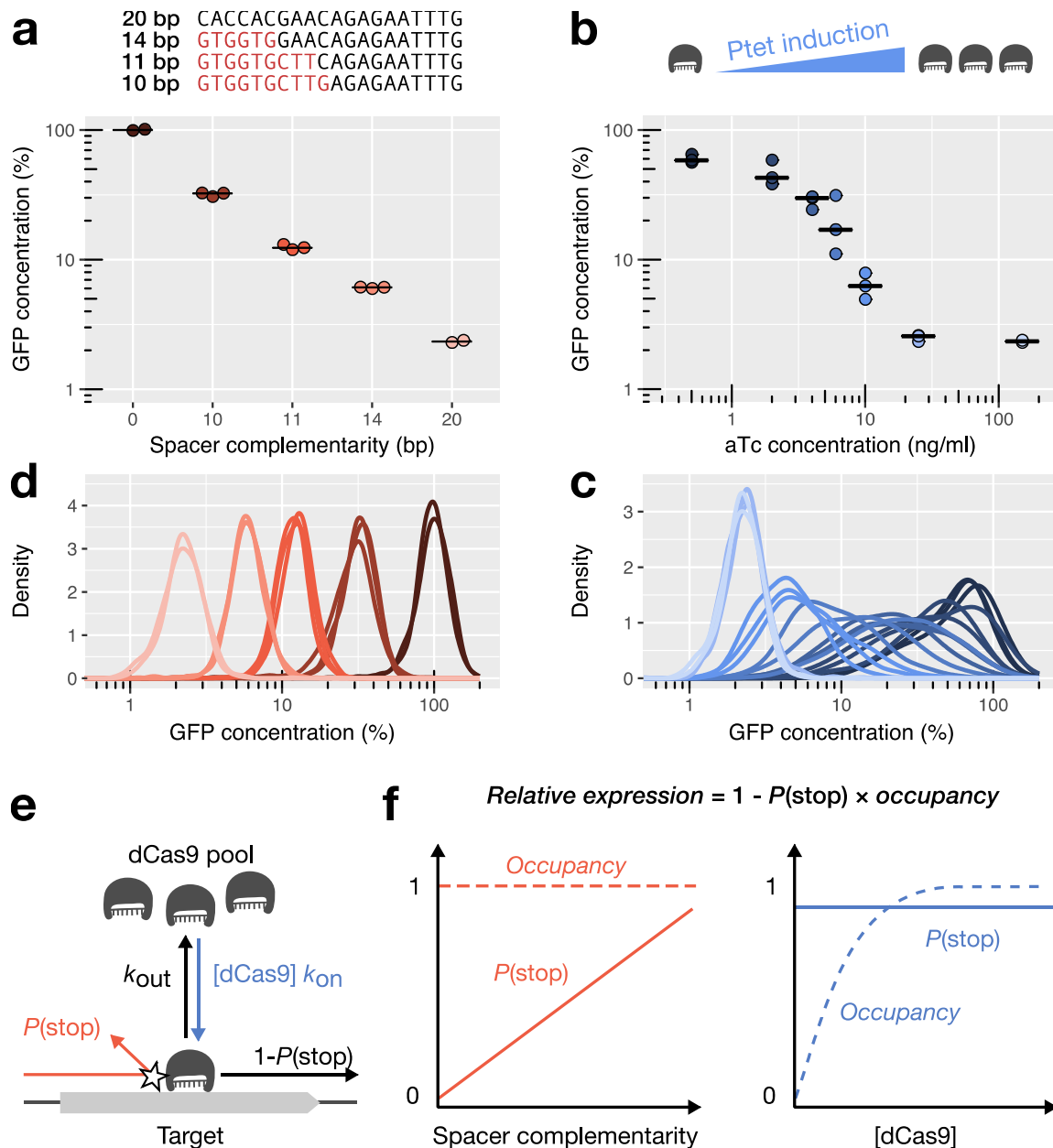


Figure 1 - In saturating conditions, CRISPR knock-down can modulate gene expression over a large dynamic range without generating noise

a, b: Average cellular GFP concentration obtained (a) by changing guide RNA-target complementarity at a constant high dCas9 concentration or (b) by varying dCas9 levels with increasing concentration of the aTc inducer. Relative GFP concentrations are obtained by high-throughput microscopy and given relatively to the non-targeting spacer at high dCas9 expression. Individual points represent independent replicates. Horizontal bars represent the median of three replicates.

c, d: Distribution of GFP concentrations for each experiment in panels a and b. Curves of the same color represent replicates of the same condition.

e: Mechanistic model of dCas9-mediated repression. The expression level of a dCas9-targeted gene is reduced by the product of two probabilities: the probability $P(stop)$ of dCas9 blocking RNAP upon collision if occupying the target, and the probability of dCas9 occupying the target (termed occupancy). If dCas9 does not block RNAP during a collision between dCas9 and RNAP, dCas9 is kicked out of the target site (with probability $1-P(stop)$). The occupancy is determined by binding constant k_{on} , dCas9 concentration $[dCas9]$, and dCas9 unbinding rate k_{off} . The unbinding rate k_{off} , in turn, is the sum of equilibrium unbinding rate and kick-out rate due to collision with the RNAP (see supplements for details).

f: The two panels schematically illustrate the behavior of the probability of dCas9 blocking RNAP $P(stop)$ and dCas9 occupancy if repression strength is controlled by guide-RNA complementarity (left) or dCas9 concentration (right), respectively.

stochastic nature of transcription referred to as intrinsic noise, which is only relevant at very low transcription rates¹¹. The plateau value of the expression noise of about 0.3 (corresponding to cell-to-cell variations of 30%) is similar to measurements made by others for constitutive genes in wild-type *E. coli*¹¹. This noise-less repression is qualitatively different from gene repression using transcriptional repressors, which can increase the extrinsic part of the noise of their targets by 5-fold as compared to the unrepressed case¹². Accordingly, a similar increase of noise is observed if repression is modulated by inducer concentration (see Supplementary Fig. 1 and also reference ⁸). The alternative system proposed here thus enables to tune expression levels with high precision in single cells.

To explain these observations, we formulate a model of dCas9 repression according to which, for a high enough intracellular concentration of dCas9, the target becomes saturated by dCas9 but transcription may still occur when dCas9 is kicked out by the RNAP. In this model, mismatches affect the repression level by increasing the probability that dCas9 is ejected upon collision (see Supplementary Information). Our model thus makes the prediction that gene repression should be independent of dCas9 concentration even in the presence of mismatches between the guide RNA and the target, as long as there are enough dCas9 complexes to saturate the target. We could indeed verify this prediction by reducing the fraction of active dCas9

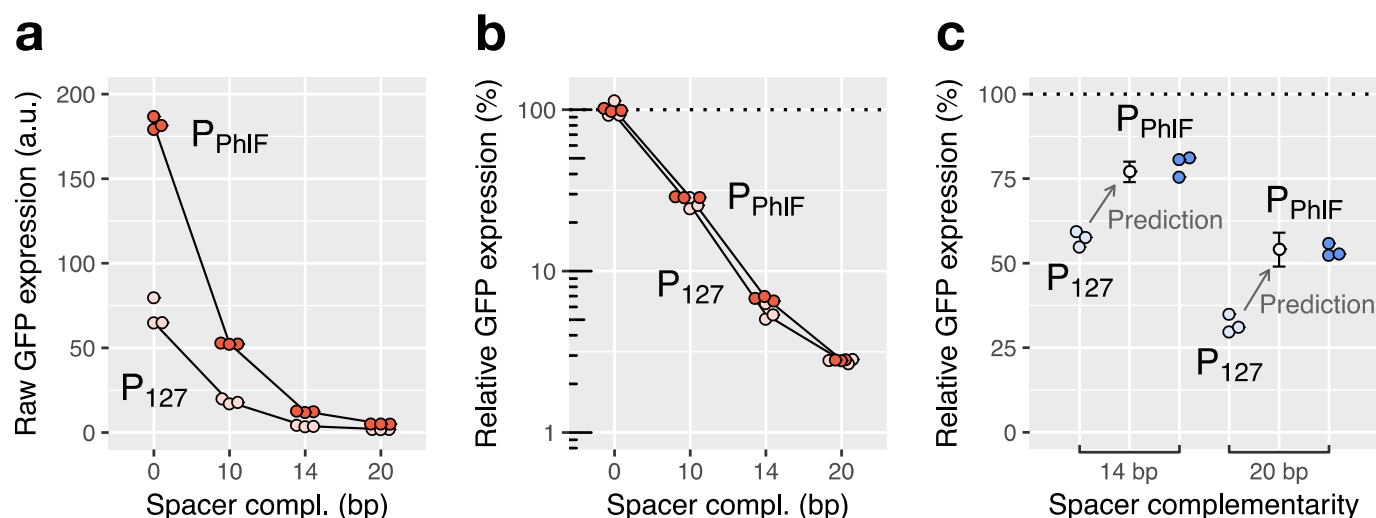


Figure 2 – Relative repression by dCas9 is independent of promoter strength only in saturating conditions.

Relative GFP expression measured by flow cytometry for two promoters of different strengths (P_{127} and P_{PhIF}) and repressed using the same set of spacers for saturating (a, b) and non-saturating (c) dCas9 concentrations. **a, b:** Raw GFP expression (a) and relative GFP expression with respect to a non-targeting spacer (b) for a saturating dCas9 concentration. While P_{PhIF} is about 3 times stronger than P_{127} , the relative expression levels after repression are similar for both promoters.

c: Experimental and predicted relative GFP expression for a non-saturating dCas9 concentration (using a 40-times lower concentration of aTc). Repression is weaker for the stronger P_{PhIF} promoter for up to 6 mismatches on the guide RNA, in quantitative agreement with the kick-out model (see Supplementary Information). Error bars: standard error of the mean of the computational prediction.

complexes using a decoy guide RNA (Supplementary Fig. 2). When repression strength is controlled by mismatches rather than dCas9 concentration, the cell-to-cell fluctuations of dCas9 concentration no longer affect the repression of the target, thus explaining the low noise obtained in Figure 1c and Supplementary Figure 1.

The relative repression level is robust with respect to changes in the target's promoter strength

To use CRISPR knock-down on genes with different native expression levels, it is important to know whether the transcription rate of the target has an influence on the relative repression. Our model predicts that repression strength should not be measurably affected for promoters of different strengths, if dCas9 is saturating the target as already indicated above. To verify this prediction, we put *sfgfp* under the control of two promoters of different strengths (P_{127} and P_{PhlF}) and blocked expression using four different guide RNAs with an increasing number of mismatches. While the strain with P_{PhlF} expressed about three times more GFP than the strain with P_{127} (Fig. 2a), the repression fold with regard to the promoter's initial expression level was identical in each case (Fig. 2b). These observations confirm that repression by mismatched guide RNAs in saturating conditions is independent of promoter strength. On the contrary, if repression is controlled by varying dCas9 concentration, our model predicts that repression fold is dependent on promoter strength: a stronger promoter is expected to kick out dCas9 more frequently, reducing the occupancy of the target. Indeed, we observed a weaker repression of P_{PhlF} compared to P_{127} at low dCas9 concentrations (Fig. 2c) that quantitatively agrees with our model prediction (Fig. 2c and Supplementary Information).

dCas9 ejection probability increases with temperature

It was recently reported¹³ that dCas9 is no longer active at 42°C, suggesting that repression strength might decrease with increasing temperature. This observation also bears the possibility that our system becomes less robust with respect to dCas9-copy number fluctuations and promoter strength with increasing temperature, if the condition of target saturation was not fulfilled. To quantify the temperature dependence of repression and test for robustness, we measured the repression of RFP by guide RNAs with 11 bp or 20 bp of

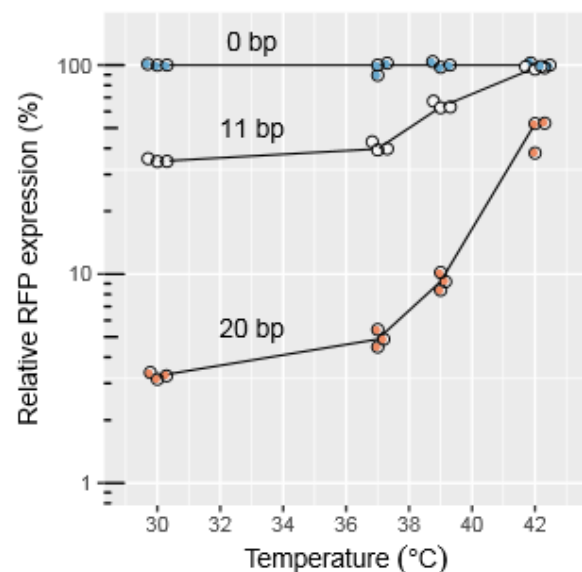


Figure 3 – The efficiency of CRISPR knock-down is affected by high temperatures.

Relative RFP expression measured by flow cytometry upon repression with different levels of complementarity and at different temperatures. The values are normalized with respect to the non-targeting spacer at each temperature.

complementarity at temperatures ranging from 30 to 42°C. The repression strength decayed continuously with increasing temperature (Fig. 3), displaying a sharp decrease of repression between 37 and 42°C. Regardless of the temperature, repression strength was not affected by dCas9 complex concentration (Supplementary Fig. 3). From our model we can thus conclude that increasing temperature does not affect dCas9 occupancy but increases the probability of dCas9 being kicked out by the RNAP. This also indicates that our system should work independently of promoter strength at all temperatures tested.

CRISPR knock-down in combination with fluorescent-protein insertions can be used to repress and monitor genes in their native contexts

Precision, robustness, and large dynamic range make complementarity-based CRISPR knock-down a versatile repression strategy. To repress genes of interest in their native context, we propose to insert *sfgfp* or *mCherry* reporters as transcriptional or translational fusions upstream of the gene. A library of CRISPR plasmids can then be introduced to repress the fusions to the desired levels by targeting the *sfgfp* or *mCherry* coding sequences. The method thus allows taking

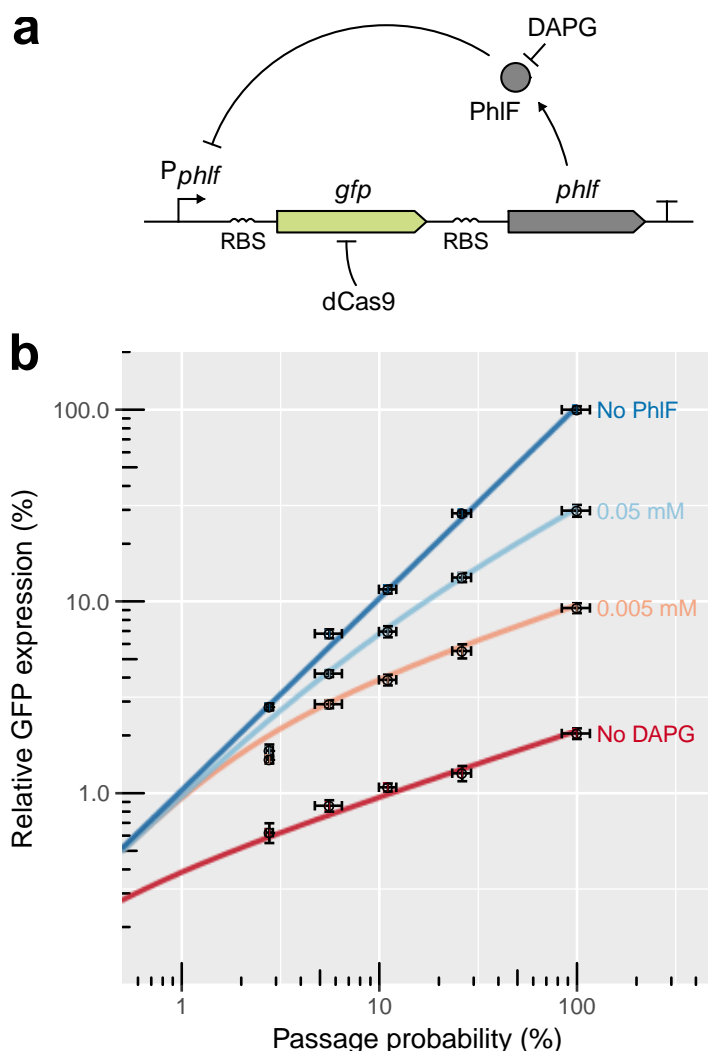


Figure 4 – CRISPR knock-down can be used to quantitatively characterize feedback loops

a: Schematic of the synthetic feedback loop constructed for this experiment. The strength of the feedback can be modulated by addition of DAPG, an inhibitor of PhlF. RBS: Ribosome Binding Site. T: Transcription terminator.

b: Flow cytometry measurements and fits to a theoretical model of relative GFP expression levels, where GFP is expressed from the artificial feedback loop presented in panel a. GFP expression is normalized by the maximal level of GFP expressed constitutively from the P_{PhlF} promoter alone (indicated as “No feedback”). The GFP is repressed using 4 different guide RNAs with respectively 10, 11, 14 and 20 bp of complementarity. The passage probability $1-P(\text{stop})$ associated with each of these guide RNAs was measured in parallel on a strain expressing GFP constitutively from the P_{127} promoter. Adding different amounts of DAPG to the medium reduces the strength of the feedback, causing the steady-state level to increase and repression to become more efficient. The colored lines represent the GFP expression as predicted by a mathematical model that was fitted to the data (see Supplementary Information). For each DAPG concentration, a binding constant characterizing the strength of the feedback and a Hill coefficient were determined. Error bars: 95% confidence interval of the mean based on 3 biological replicates.

advantage of the measured repression levels for constitutive promoters established above. Furthermore, gene expression can be measured at the single-cell level, revealing cell-to-cell variations. In the following, we demonstrate that this system has broad applicability for the study of genetic regulation and cellular physiology: first, we study the regulation of a model transcriptional feedback circuit, and second we quantify the effect of fractional protein repression on cell morphology during steady-state growth.

CRISPR knock-down can be used to uncover and characterize genetic feedback

To demonstrate the versatility of our system for the study of genetic circuits, we chose the previously described PhlF auto-repressor from *Pseudomonas fluorescens*¹⁵ as a model system: We constructed a synthetic operon consisting in the P_{PhlF} promoter followed by the *sfgfp* and *phlF* genes in a single operon (Fig. 4a). The PhlF repressor binds to the P_{PhlF} promoter and decreases transcription initiation, thus creating an artificial negative feedback loop. The strength of this feedback can be externally reduced by adding the chemical inducer 2,4-diacetyl-phloroglucinol (DAPG) that blocks binding of PhlF to the promoter (Supplementary Fig. 5). Accordingly, higher DAPG concentrations lead to higher steady-state concentrations of PhlF and GFP. To quantify the feedback strength for different DAPG concentrations we targeted the *sfgfp* ORF using spacers with variable degrees of complementarity (Fig. 4). CRISPR knock-down of GFP should lead to an increased transcription initiation rate of the promoter. As a consequence the fold change of expression during CRISPR knock-down should be lower in the case of feedback than without feedback. The quantitative difference between the two situations can then be used to quantify the feedback strength.

As anticipated, expression of GFP decreased with increasing complementarity and the relative reduction of expression was less pronounced with feedback than without feedback (Fig. 4b). We then fit the expression data to a mathematical model of gene repression (Fig. 4b and Supplementary Information) to calculate for each DAPG concentration the binding constant of the repressor K and a Hill coefficient n . We observe that a Hill coefficient of $n=2$ describes our data for low DAPG concentrations (0 μ M and 5 μ M), while a Hill coefficient of $n=1$ was required to describe our observations at 50 μ M. PhlF proteins dimerize *in vitro* and are thought to bind the operator as a dimer¹⁶. We thus hypothesized that PhlF is predominantly found as monomers at low DAPG concentrations, where PhlF expression is also low, and as dimers at high DAPG concentrations, where PhlF expression is high. The role of DAPG concentration in this transition is discussed in the Supplementary Information.

The detailed insights obtained here demonstrate the usefulness of precisely controlling the rate at which the RNAP is blocked by dCas9, while monitoring residual expression with a fluorescent reporter. The same method can be applied to other and more complex problems of gene regulation, e.g., by monitoring the response of one gene to the precisely tuned levels of another gene repressed by CRISPR knock-down.

CRISPR knock-down reveals how cells adapt their shapes to low levels of an essential cell-wall synthesis operon

We then used our approach to explore the morphological response of cells to different expression levels of two essential proteins for peptidoglycan cell-wall synthesis encoded by the *mrdAB* operon. PBP2 (encoded by *mrdA*) and RodA (encoded by *mrdB*) are inner membrane proteins with respectively transglycosylase¹⁷ and transpeptidase activity¹⁸. The two highly conserved enzymes are part of the multi-enzyme 'rod' complex, which is essential for cell-wall synthesis during cell elongation¹⁹.

Previous depletion experiments suggest that PBP2 expression is buffered against large copy-number fluctuations, as cells grow for multiple generations before showing a reduction of growth rate²⁰. The drawback of depletion experiments is that they do not allow studying the effect of protein abundance in the steady state. To quantify the relation between PBP2 levels and morphological response during steady-state growth we constructed a translational protein fusion by seamlessly integrating *mCherry* in front of the *mrdA* open reading frame in the native chromosomal *mrdAB* locus (Fig. 5a).

The mCherry-PBP2 fusion is fully functional, similarly to a fusion constructed previously²⁰. We then introduced a chromosomal P_{tet}-*dCas9* cassette and different pCRRNA plasmids programmed to target *mCherry* with 0, 11, 18 or 20 bp of complementarity in order to obtain a range of transcription rates for the operon. These strains were induced for dCas9 expression and grown until protein levels and cell dimensions reached steady state (Supplementary Fig. 6). Lowering expression of the *mrdAB* operon led to increasing cell width, with a sharp rise of cell width below ~20% of the native expression level (Fig. 5b, 5c and Supplementary Fig. 7a), while cell length was largely unaffected except for the highest repression strength (Supplementary Figs. 8b and 9). Cells presumably increased their diameters to lower their surface-to-volume ratio in response to reduced levels of cell-wall synthesis²¹. In support of this hypothesis, cell-to-cell fluctuations in the intracellular density of mCherry-PBP2 negatively correlated with cell diameter for intermediate *mrdAB* repression (at 11 bp guide RNA/target complementarity) (Fig. 5d). Such a correlation was not observed when the operon was not repressed (Fig. 5e), indicating that the cells buffer natural fluctuations of *mrdAB* and thus avoid fluctuations of cell morphology, as previously suggested²⁰. By gradually lowering the levels of PBP2 and RodA, we were able to take the cells out of the buffering regime at about 30% of native expression. Together, these experiments demonstrate that cells buffer gene expression of an essential operon against copy-number fluctuations of about threefold and that cells cope with even stronger fluctuations by adjusting their surface-to-volume ratio. However, once expression levels are reduced by more than fivefold, cells show severe growth defects.

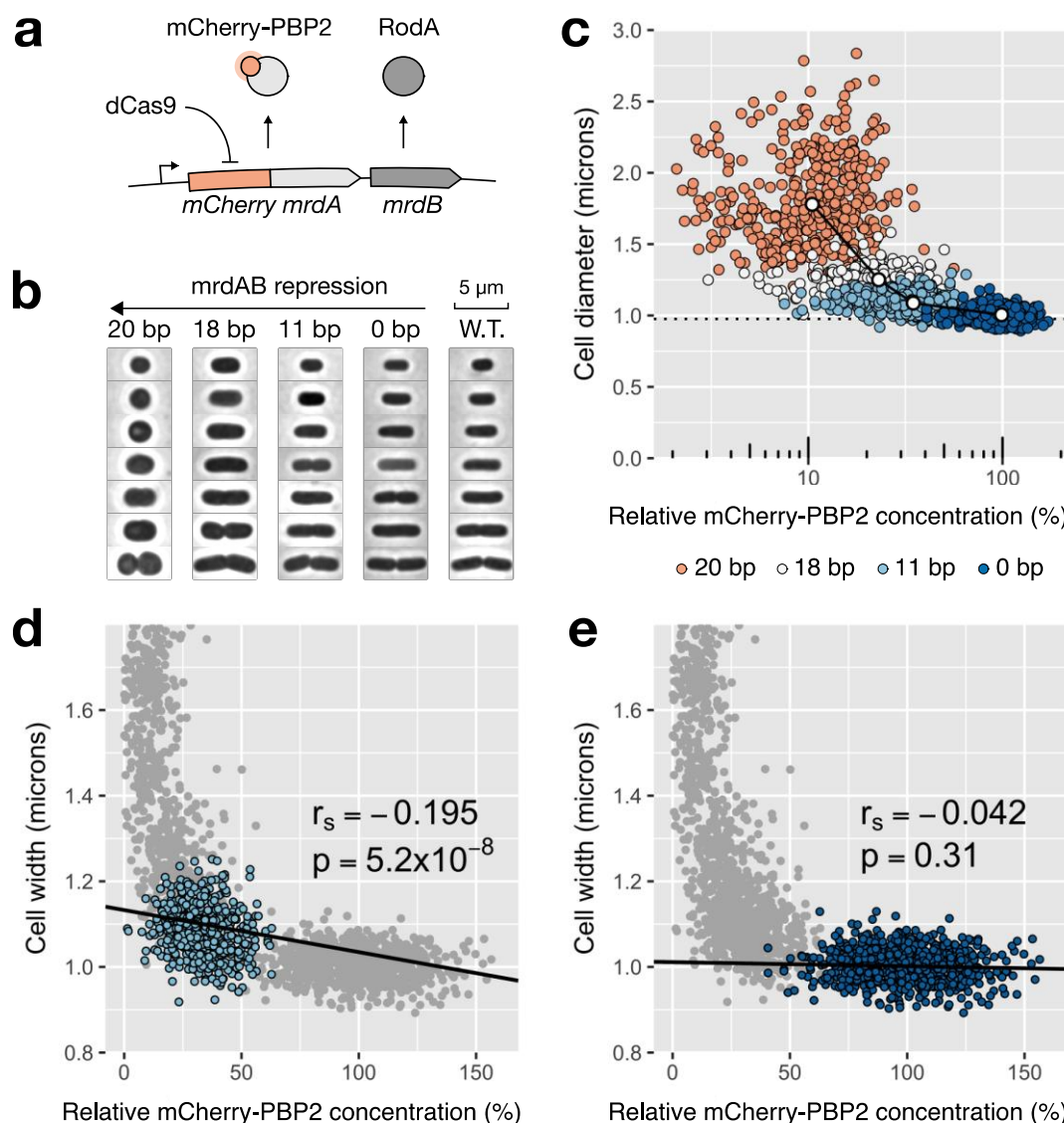


Figure 5 – CRISPR knock-down of the *mrdAB* operon increases cell width at high repression strengths.

a: Schematic of the modified chromosomal locus of the *mrdAB* operon in strain AV08.

b: Cell shapes observed by phase-contrast microscopy for cells grown in M63 minimal medium. Different repression levels of the *mrdAB* operon are compared to wild-type *E. coli*. Cells with different cell lengths were picked at random and images were rotated numerically.

c: Cell width as a function of the mCherry-PBP2 concentration measured by fluorescence microscopy. Each point represents a cell and colors represent different levels of spacer complementarity. The connected white dots represent the population averages (mean of 3 biological replicates). The dotted line represents the average cell width for wild-type *E. coli* (mean of 3 replicates). The values are normalized with respect to the non-targeting spacer.

d, e: Linear regression between mCherry-PBP2 concentration level and cell width, for the strains repressed with 11 bp (panel d) and 0 bp (no repression, panel e). r_s is the Spearman correlation coefficient (median of 3 biological replicates). The negative value indicates that cells with a lower level of PBP2/RodA tend to be wider. The p-values (two-sided F-test) measure the certainty that the slope is different from 0.

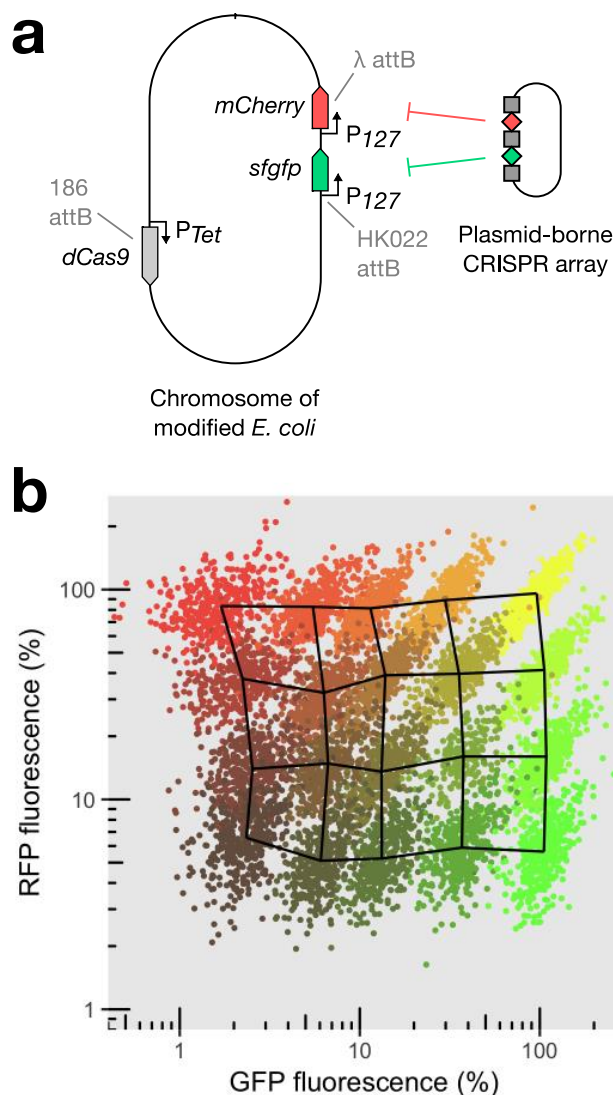


Figure 6 – CRISPR knock-down can be multiplexed to modulate expression of two genes without cross-talk

a: Schematic of the strain expressing two reporters and P_{Tet}-*dCas9* integrated in the chromosome at phage attachment sites. The levels of the two reporters can be controlled using a plasmid-borne CRISPR array coding for guide RNAs (diamonds) interspaced with CRISPR repeat motifs (squares), and also carrying the *tracrRNA* sequence (not shown).

b: Relative GFP and RFP concentration given relatively to the non-targeting spacer measured by high-throughput microscopy for a collection of 20 CRISPR plasmids. Each point represents a single cell, and each color represents the population obtained with one CRISPR plasmid. The overlaid meshwork connects the median values of the different populations.

CRISPR knock-down can be used to modulate the expression of two genes

With our method, the fractional repression level of any target gene is controlled genetically rather than chemically by the concentration of an inducer. It can thus be used to modify expression of multiple genes independently. To assess this potential, we built a library of CRISPR arrays containing two spacers, one targeting *sfgfp* and the other *mCherry*. We selected five spacers with varying levels of complementarity to each of the target, spanning a large range of expression, from 2% to 100% of the initial level. We combined these spacers to form 20 CRISPR arrays that cover the entire space of expression and used them to control the concentrations of GFP and RFP expressed from the chromosome (Fig. 6a). As expected, the repression of one gene is independent of the repression of the other (Fig. 6b). Strong correlations between GFP and RFP in single cells targeted with the same combinations of guide RNAs are due to common sources of extrinsic noise¹². We anticipate this to be a useful tool to study interactions of genes and specifically the effect of stoichiometry in genetic networks.

Discussion

Here, we demonstrate that tuning gene expression through complementarity between guide RNA and target works robustly at the single-cell level, with two specific advantages over previous methods relying on inducible promoters: First, relative repression strength is independent of native expression levels, making the system applicable to study genes of vastly different promoter strengths. Second, the system preserves endogenous expression noise of the repressed gene. This allows studying the impact of gene repression on cellular physiology without generating stochastic cell-to-cell variability, which is known to have important downstream consequences for processes such as cell differentiation or the emergence of spatial structure in populations^{22,23}.

The ability to control gene expression level through guide RNA complementarity rather than the concentration of an inducer has other advantages: For example, it enables differential control of different cells within the same culture. This could prove useful in pooled screens or competition assays. Furthermore, the strategy can be multiplexed to enable the simultaneous control of multiple genes independently without requiring multiple chemical inducers. We demonstrated this ability with two targets (GFP and RFP; Fig. 6), which can be inserted in front of genes of interest. The strategy can easily be extended to include more than two fluorescent reporters as targets.

Alternatively to using fluorescent-protein fusions it is also possible to guide dCas9 directly to the gene of interest, but this comes with the disadvantage of uncertainty about the exact repression strength due to two reasons: First, the rate at which dCas9 blocks the RNAP is dependent on the specific target sequence. In the future it might thus be desirable to develop computational means to predict target repression based on sequence alone²⁴. Second, any feedback controlling the expression of the target could lead to altered transcription-initiation rates (Fig. 4). Therefore, using fluorescent-protein fusions has the advantage to report the exact expression level.

The properties of dCas9 repression described in this study originate from the mechanism of dCas9 binding to DNA inside ORFs. We show that at high concentrations, dCas9 is saturating the target site even when using guide RNAs with large numbers of mismatches, where repression of the target gene is weak. We explain these observations by a 'kick-out' model of repression, according to which RNAP kicks out dCas9 with a probability that can be tuned by spacer complementarity (Supplementary information). The model predicts that repression is promoter-strength dependent in non-saturating conditions, in quantitative agreement with experiments (Fig. 2c). Our model also predicts a low rate of spontaneous unbinding of dCas9 from the target for full and intermediate (14 bp) complementarity (Supplementary Fig. 10). For full complementarity, our model is compatible with the hypothesis that dCas9 never leaves the target spontaneously but gets kicked out either by the RNAP or during DNA replication. This prediction is consistent with the long half-life of dCas9 binding *in vitro* previously reported³. A recent high-throughput study of dCas9 off-target binding and unbinding suggests

that mutations in the PAM-distal region control the unbinding kinetics of dCas9²⁴. Since unbinding is dominated by RNAP-dCas9 collisions for the promoters tested here and as rebinding to the target is fast at high dCas9 concentrations, spontaneous unbinding plays no significant role for gene repression in our model system. However, if dCas9 targets the promoter rather than the coding region, spontaneous unbinding might be the only mode allowing residual gene expression. This view is supported by the higher repression strength observed when targeting the promoter region^{1,2}.

Our results also suggest that dCas9 ejection does not lead to bursts of transcription, but that instead dCas9 returns to the target site after ejection in a time that is small with respect to the typical time interval between transcription initiations. It is still conceivable that such bursts may occur for transcription rates higher than the strongest promoter we used.

Our strategy enables to precisely control gene expression without introducing cell-to-cell variability, and should be useful for any quantitative measurements that depend on the expression level of a gene. By taking advantage of the ability to precisely control the rate at which dCas9 blocks the RNAP we could characterize a synthetic feedback loop, revealing unexpected properties of Phlf repression activity. In a second example, we took advantage of the ability to fine-tune expression levels at the steady state to quantitatively measure cell shape as a function of the levels of PBP2 and RodA. The level of precision achieved here would be hard to establish with conventional methods. Accordingly, this is the first study to establish a quantitative relationship between the abundance of cell-wall synthesizing proteins and cell morphology at the population and single-cell levels. We anticipate that our method will be useful to study many other systems and in particular genetic circuits that include high levels of noise, such as stochastic switches, or other noise-dependent processes, where preservation of a well-defined level of expression noise is desirable.

Materials and methods

Genome modifications

All the strains used for measurements derive from *E. coli* MG1655. Table S1 details the construction of the strains used in this study.

For integration of cassettes at phage attachment sites, we used the “clonetegration” method²⁵. Integrated backbones were excised by expressing a flippase from pE-FLP. Plasmid pLC97 (addgene number: to be announced) can be used to easily integrate P_{Tet}-dCas9 at the lambda attB site.

For scarless integration of *mCherry-mrdA* in the native *mrdAB* operon we used the pCas / pTarget system²⁶. The PAmCherry-PBP2 protein fusion present in strain TKL130^{27,20} was replaced by a translational fusion with *mCherry* extracted from plasmid pFB262²⁸. To this end the pAV06 variant of pTarget was constructed and genome editing was performed as described in reference ²⁶.

The sequences of the *sfGFP* and *mCherry* genes used in this study can be found on the GenBank database with accession codes KT192141.2 and JX155246.1, respectively.

Plasmid design and construction

The CRISPR targets were chosen next to the beginning of the ORF, but at least 50 bp away from the initiation codon, in order to preclude unwanted interactions with the native regulation of transcription. None of spacers used in this study have any off-target position with more than 8 bp of complementarity in the PAM-proximal region. Spacers were cloned into the CRISPR array of plasmid pCRRNA using Golden-Gate assembly as previously described^{29,30}. The oligonucleotide sequences are available in table S4.

The other plasmids from this study were constructed by Gibson assembly³¹. The fragments are described in table S2 and the primer sequences in table S3. DNA constructions were electroporated in *E. coli* strain DH5 α or Pi1 for *pir*-dependent origins of replication³².

Media and reagents

For all flow cytometry measurements, the cells were grown in Luria-Bertani (LB) broth. As a minimal medium for the *mrdAB* measurements, we used M63 medium supplemented with 2 g/L of glucose, 10 mg/L of thiamine, 10 mM of MgSO₄ and 1 g/L of casaminoacids. When needed, we used various antibiotics (25 μ g/ml chloramphenicol, 100 μ g/ml carbenicillin, 50 μ g/ml kanamycin, 100 μ g/ml spectinomycin). Di-acetyl-phloroglucinol (DAPG) and anhydrotetracycline (aTc) were used respectively for induction of P_{PhlF} and P_{Tet} promoters. All oligonucleotides were obtained from Eurofins Genomics.

Preparation of steady-state exponential cultures

Unless stated otherwise, all cultures were grown at 30°C. Strains were first re-streaked from a freezer stock. Independent single colonies were picked for each replicate. Cells were then grown overnight in 96 deep-wells plates using a table-top shaker in 1 ml of medium with 100 ng/ml of aTc and 50 μ g/ml of kanamycin (Eppendorf). The day of the measurement, cultures were back-diluted 250 times in fresh medium with aTc and kanamycin, and grown for 1h45 into exponential phase. We then fixed the cells with 4% formaldehyde (30 min on ice) and washed with PBS.

Flow cytometry

Fluorescence of single cells was recorded using a Miltenyi MACSquant flux cytometer. 10,000 events were recorded per replicate. In all cases, the AV01 strain (with no reporters) carrying a non-targeting pCRRNA plasmid was used to measure the auto-fluorescence background. We calculated the mean fluorescence signal of each population and subtracted the mean autofluorescence signal. To test whether

differences in expression were significant we performed Student's t-test on the natural logarithms of the average fluorescence³³.

High throughput microscopy (imaging cytometry)

An Amnis ImageStreamX (EMD Millipore) imaging cytometer was used to image the cells in high throughput in brightfield, GFP, and RFP channels. Images were analyzed using the IDEAS® (EMD Millipore) software suite. For each condition, at least 10,000 events were recorded per replicate. Cells that were out of focus or tilted were identified by calculating the average gradient of a pixel normalized for variations in intensity levels (*Gradient RMS* feature in IDEAS®). Additionally, we used the Feature Finder script of IDEAS® to remove contaminating particles, images with multiple cells and beads. After filtering, at least 2000 images remained per sample. The fluorescence channels were not used for filtering. A color compensation matrix was calculated to account for spectral overlap of GFP and RFP emission spectra, so cultures of AV02 (GFP only) and AV04 (RFP only) would each have a null signal on the converse channel. As a proxy for the reporter's intracellular concentration we used the average image intensity inside the area corresponding to each cell. The cell area was determined by using a threshold on the bright field images. Single points located more than 3 standard deviations away from the population average were discarded as outliers, as they can disrupt the noise computations. The average fluorescence μ of each sample was calculated by taking the mean of the single-cell fluorescence. The noise was defined as σ/μ , with $\sigma = \sqrt{\sigma_{\text{sample}}^2 - \sigma_{\text{blank}}^2}$ where σ_{sample} is the standard deviation of the intracellular average intensity of the sample, and σ_{blank} is the standard deviation of a sample with no fluorescent reporter (noise from auto-fluorescence).

Fluorescence and phase-contrast microscopy

Fixed cells were transferred to Phosphate Buffered Saline (PBS) microscopy pads with 1.5% agar and imaged using an inverted microscope (TI-E, Nikon Inc.) equipped with a 100x phase contrast objective (CFI PlanApo LambdaDM100x 1.4NA, Nikon Inc.), a solid-state light source (Spectra X, Lumencor Inc.), a multiband dichroic (69002bs, Chroma Technology Corp.). mCherry fluorescence was measured using excitation (560/32) and emission (632/60) filters. Images were acquired using a sCMOS camera (Orca Flash 4.0, Hamamatsu) with an effective pixel size of 65 nm.

MatLab code adapted from the Morphometrics package³⁴ was used to find cell contours from phase-contrast images. Background intensity, uneven illumination and cell auto-fluorescence were accounted for in the analysis. For the analysis of fluorescence signal, we corrected the raw mCherry values for uneven illumination, background intensity and cell auto-fluorescence. Intracellular protein concentration was obtained as the mean pixel intensity inside the cell area. Total regression was used to find the major axis of the cell. Cell width was defined as the average distance between the cell contour and this axis, excluding the poles.

Measurements of cell morphology during steady-state exponential growth (Supplementary Fig. 6) were performed after overnight induction of dCas9, followed by 1/250 dilution of the culture in fresh M63 medium with aTc and while the culture was kept in exponential growth phase at an optical density below 0.1. Samples were taken from the culture, fixed and imaged after 2h and 4h.

Northern blot

Total RNA was extracted from cultures in early stationary phase using Trizol. Electrophoresis on Novex® TBE-Urea Gels (10% polyacrylamide gels containing 7 M urea, Invitrogen) was used to separate RNAs. The gels were blotted onto Nylon membranes (Invitrogen), which were subsequently cross-linked with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC, Thermo Scientific) buffer³⁵. The probes were labeled as follow: 100 pmol of oligonucleotide were heat denatured, labeled and phosphorylated by mixing 40 µCi of 32P-γ-ATP (PerkinElmer) and T4 PNK (NEB) reagents. A labeled probe specific to the guide RNA R20 (5' GCATAGCTCTAAACTCCGTATGAAGGCACCCAGA 3') was column purified (Macherey-Nagel PCR cleanup kit) and used for overnight hybridization. The intensity of the shortest band, corresponding to the fully processed guide RNA, was quantified using the Fiji software package.

Acknowledgements

We thank T. K. Lee and K.C. Huang for providing strain TKL130, F. Bendezù for providing plasmid pFB262, H. Cho and T. Bernhardt for providing plasmid pHc942 and D. Mazel for providing plasmid pSW23t. We thank J. Fernández-Rodríguez for providing the *P_{phlF}-sfGFP-phlF* DNA fragment. We thank E. Brambilla and E. Oldewurtel for support in microscopy, as well as A. Soler and M. Hasan for support regarding flux cytometry and imaging cytometry. We acknowledge the Technology Core of the Center for Translational Science (CRT) at Institut Pasteur for support in conducting this study. This work was supported by the European Research Council (ERC) under the Europe Union's Horizon 2020 research and innovation program (grant agreement No [677823] to DB and No [679980] to SvT); the French Government's Investissement d'Avenir program Laboratoire d'Excellence 'Integrative Biology of Emerging Infectious Diseases' [ANR-10-LABX-62-IBEID] to SvT, DB, and AV; the Marie de Paris "Emergence(s)" program to SvT; and the Pasteur-Weizmann consortium to DB.

Author contributions

DB, SvT and AV designed the study. AV constructed the strains and plasmids and performed all measurements and data analysis. SvT developed the mathematical models. LC took part in strain construction and EO performed the compensation of uneven illumination for the measurement of mCherry-PBP2 by microscopy. AV, DB and SvT wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Bikard, D. *et al.* Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41**, 7429–7437 (2013).
2. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
3. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–7 (2014).
4. Szczelkun, M. D. *et al.* Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc Natl Acad Sci U A* **111**, 9798–803 (2014).
5. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* (2014). doi:10.1038/nbt.2889
6. Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **32**, 677–683 (2014).
7. Peters, J. M. *et al.* A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165**, 1493–1506 (2016).
8. Li, X. *et al.* tCRISPRi: tunable and reversible, one-step control of gene expression. *Sci. Rep.* **6**, 39076 (2016).
9. Farasat, I. & Salis, H. M. A Biophysical Model of CRISPR/Cas9 Activity for Rational Design of Genome Editing and Gene Regulation. *PLOS Comput. Biol.* **12**, e1004724 (2016).
10. Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**, 1262–1278 (2014).
11. Taniguchi, Y. *et al.* Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538 (2010).
12. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic Gene Expression in a Single Cell. *Science* **297**, 1183–1186 (2002).
13. Wiktor, J., Lesterlin, C., Sherratt, D. J. & Dekker, C. CRISPR-mediated control of the bacterial initiation of replication. *Nucleic Acids Res.* **44**, 3801–3810 (2016).
14. Pósfai, G., Kolisnychenko, V., Bereczki, Z. & Blattner, F. R. Markerless gene replacement in Escherichia coli stimulated by a double-strand break in the chromosome. *Nucleic Acids Res.* **27**, 4409–4415 (1999).
15. Stanton, B. C. *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nat. Chem. Biol.* **10**, 99–105 (2014).
16. Abbas, A. *et al.* Characterization of Interactions between the Transcriptional Repressor PhlF and Its Binding Site at the phlA Promoter in Pseudomonas fluorescens F113. *J. Bacteriol.* **184**, 3008–3016 (2002).

17. Meeske, A. J. *et al.* SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nature* **537**, 634–638 (2016).
18. Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A. & Charlier, P. The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiol. Rev.* **32**, 234–258 (2008).
19. Cho, H. *et al.* Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nat. Microbiol.* **1**, 16172 (2016).
20. Lee, T. K. *et al.* A dynamically assembled cell wall synthesis machinery buffers cell growth. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4554–4559 (2014).
21. Harris, L. K. & Theriot, J. A. Relative Rates of Surface and Volume Synthesis Set Bacterial Cell Size. *Cell* **165**, 1479–1492 (2016).
22. Waite, A. J. *et al.* Non-genetic diversity modulates population performance. *Mol. Syst. Biol.* **12**, 895 (2016).
23. Çağatay, T., Turcotte, M., Elowitz, M. B., Garcia-Ojalvo, J. & Süel, G. M. Architecture-Dependent Noise Discriminates Functionally Analogous Differentiation Circuits. *Cell* **139**, 512–522 (2009).
24. Boyle, E. A. *et al.* High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proc. Natl. Acad. Sci.* 201700557 (2017). doi:10.1073/pnas.1700557114
25. St-Pierre, F. *et al.* One-Step Cloning and Chromosomal Integration of DNA. *ACS Synth. Biol.* **2**, 537–541 (2013).
26. Jiang, Y. *et al.* Multigene Editing in the Escherichia coli Genome via the CRISPR-Cas9 System. *Appl. Environ. Microbiol.* **81**, 2506–2514 (2015).
27. Subach, F. V. *et al.* Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods* **6**, 153–159 (2009).
28. Bendezú, F. O., Hale, C. A., Bernhardt, T. G. & de Boer, P. A. J. RodZ (YfgA) is required for proper assembly of the MreB actin cytoskeleton and cell shape in E. coli. *EMBO J.* **28**, 193–204 (2009).
29. Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat Biotechnol* **32**, 1146–50 (2014).
30. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of Escherichia coli. *Nucleic Acids Res.* **44**, 4243–4251 (2016).
31. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**, 343–5 (2009).
32. Shafferman, A. & Helinski, D. R. Structural properties of the beta origin of replication of plasmid R6K. *J. Biol. Chem.* **258**, 4083–4090 (1983).
33. Beal, J. *et al.* Reproducibility of Fluorescent Expression from Engineered Biological Constructs in E. coli. *PLOS ONE* **11**, e0150182 (2016).
34. Ursell, T. *et al.* Rapid, precise quantification of bacterial cellular dimensions across a genomic-scale knockout library. *BMC Biol.* **15**, 17 (2017).
35. Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection of small RNA. *Nat. Protoc.* **3**, 1077–1084 (2008).