

ARTICLE

DOI: 10.1038/s41467-018-04901-6

OPEN

Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria

Chen Dong¹, Jason Fontana², Anika Patel¹, James M. Carothers^{2,3,4}  & Jesse G. Zalatan^{1,2,4} 

Methods to regulate gene expression programs in bacterial cells are limited by the absence of effective gene activators. To address this challenge, we have developed synthetic bacterial transcriptional activators in *E. coli* by linking activation domains to programmable CRISPR-Cas DNA binding domains. Effective gene activation requires target sites situated in a narrow region just upstream of the transcription start site, in sharp contrast to the relatively flexible target site requirements for gene activation in eukaryotic cells. Together with existing tools for CRISPRi gene repression, these bacterial activators enable programmable control over multiple genes with simultaneous activation and repression. Further, the entire gene expression program can be switched on by inducing expression of the CRISPR-Cas system. This work will provide a foundation for engineering synthetic bacterial cellular devices with applications including diagnostics, therapeutics, and industrial biosynthesis.

¹Department of Chemistry, University of Washington, Seattle, WA 98195, USA. ²Molecular Engineering & Sciences Institute, University of Washington, Seattle, WA 98195, USA. ³Department of Chemical Engineering, University of Washington, Seattle, WA 98195, USA. ⁴Center for Synthetic Biology, University of Washington, Seattle, WA 98195, USA. Correspondence and requests for materials should be addressed to J.M.C. (email: jcaroth@uw.edu) or to J.G.Z. (email: zalatan@uw.edu)

Bacteria are attractive targets for a wide variety of engineering applications. Bacterial strains with the ability to utilize carbon sources like CO₂, CO, methane, or lignocellulose, and alternative energy sources such as light or H₂ could provide the foundation for cost-effective and environmentally-friendly industrial biosynthesis^{1,2}. Microbial communities, such as those that reside in the human gut, play an important role in human health and disease, and tools to engineer these bacteria have great potential as both diagnostics and therapeutics^{3–5}. To harness, regulate, and modify the behavior of these and other bacteria, there is a compelling need to develop genetic tools to control gene expression and implement complex, multi-gene regulatory programs. Ideally, we want to build circuits that can regulate many genes at once, dynamically respond to external inputs or the internal state of the cell, and be easily reprogrammed to explore different functional architectures. While capabilities to edit and modify genomes are rapidly expanding, our ability to encode a precisely-defined and dynamically-responsive gene expression program with cis-regulatory sequences at the DNA level remains difficult. Thus, we sought to develop synthetic transcription factors in bacteria, which could be coupled to programmable DNA binding domains and controlled by inducible promoters to engineer complex, dynamically-responsive multi-gene expression programs.

Synthetic control of gene expression has recently become much more straightforward with the emergence of programmable transcription factors using the CRISPR-Cas system (Fig. 1). A catalytically-inactive Cas9 (dCas9) protein can be used to target specific DNA sequences with guide RNAs (gRNAs) that recognize their targets based on predictable Watson-Crick base pairing. This approach can be used to repress genes by physically blocking RNA polymerase (CRISPR interference or CRISPRi)^{6,7}. To activate genes (CRISPR activation or CRISPRa), the CRISPR complex can be linked to a transcriptional activator by direct fusion to dCas9 or via recruitment domains on the gRNA^{7–11}. In bacteria, however, there are very few transcriptional activation domains that have been reported to be effective when fused to modular DNA binding domains. Bacterial two-hybrid systems have been constructed with pairs of candidate interacting proteins separately fused to RNA polymerase subunits and DNA binding proteins. It is also possible to fuse RNA polymerase subunits directly to DNA binding domains to activate transcription^{12–14}. One of the RNA polymerase subunits, RpoZ, has been coupled to the CRISPR system to activate gene expression^{7,15–18}. For comparison, in eukaryotic systems there are many effective activators and CRISPRa has been extensively used in a variety of applications¹⁹. The paucity of reports of CRISPRa

in bacteria suggests that RpoZ may not be effective as a general activator of transcription, or that we lack a complete understanding of the design rules to predictably activate gene expression in bacteria.

To develop an improved toolkit for gene activation in bacteria, we screened a broad set of candidate proteins for transcriptional activity in *E. coli*. We identified several proteins that can effectively activate gene expression when recruited via the CRISPR-Cas system. Using the most effective activator, SoxS, we can activate one target gene while simultaneously repressing a different target gene with CRISPRi, and we can control the entire multi-gene expression program with inducible promoters driving CRISPR-Cas system components. We find that gene activation in *E. coli* is highly sensitive to the location of the gRNA target site, consistent with prior results⁷, and suggesting a possible explanation for why it has been difficult to develop synthetic activators in bacteria. Finally, we show that bacterial CRISPRa can be used to increase the output of a heterologous ethanol biosynthesis pathway. These results provide a framework for implementing CRISPRa in bacteria with a wide variety of potential applications. Further, because SoxS interacts with a highly conserved site on RNA polymerase, our bacterial CRISPRa toolkit may be portable to a broad range of bacterial species.

Results

Identifying transcriptional activation domains for bacteria. To recruit transcriptional activators to the CRISPR-Cas system, we used gRNAs that are extended with hairpin sequences to recruit RNA binding proteins (RBPs), which are in turn fused to candidate activators (Fig. 1a)¹⁰. These modified gRNAs, termed scaffold RNAs (scRNAs), encode both the target sequence and the regulatory action to execute at that target. By using scRNAs to recruit activators (CRISPRa) to some genes and gRNAs to physically block RNA polymerase (CRISPRi) at other genes, we can encode complex expression programs where some genes are activated and others are repressed (Fig. 1b)¹⁰.

To screen for potential activators, we first constructed an *E. coli* strain with a genomically-integrated, weakly-expressed GFP reporter gene. The upstream region of the reporter gene includes a number of potential gRNA target sites and is identical to that used previously to evaluate the dCas9-RpoZ fusion protein (see Supplementary Methods)⁷. We targeted the CRISPR-Cas complex to the W108 gRNA site located 91 bases upstream of the transcriptional start site (TSS), as this site previously demonstrated the strongest activation with dCas9-RpoZ⁷. For candidate activators we chose endogenous transcriptional regulators SoxS,

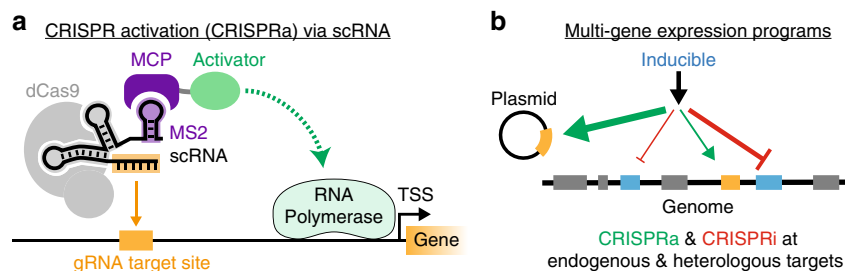


Fig. 1 CRISPR activation in bacteria enables complex multi-gene expression programs. **a** To activate gene expression, we target a CRISPR-Cas complex upstream of a target gene. dCas9 binds a scaffold RNA (scRNA), which is a modified gRNA that encodes both the target sequence and an RNA hairpin to recruit effector proteins that interact with RNA polymerase. The schematic depicts a 1x MS2 scRNA containing an MS2 RNA hairpin, which binds the MS2 coat protein (MCP) that is fused to candidate activator proteins¹⁰. **b** Combining CRISPRi with CRISPRa enables multi-gene expression programs for simultaneous activation and repression. scRNAs that recruit activators can target genes for activation, while gRNAs targeted within a gene result in CRISPRi-based repression. If the CRISPR-Cas system components are controlled by inducible promoters, the entire gene expression program can be dynamically regulated

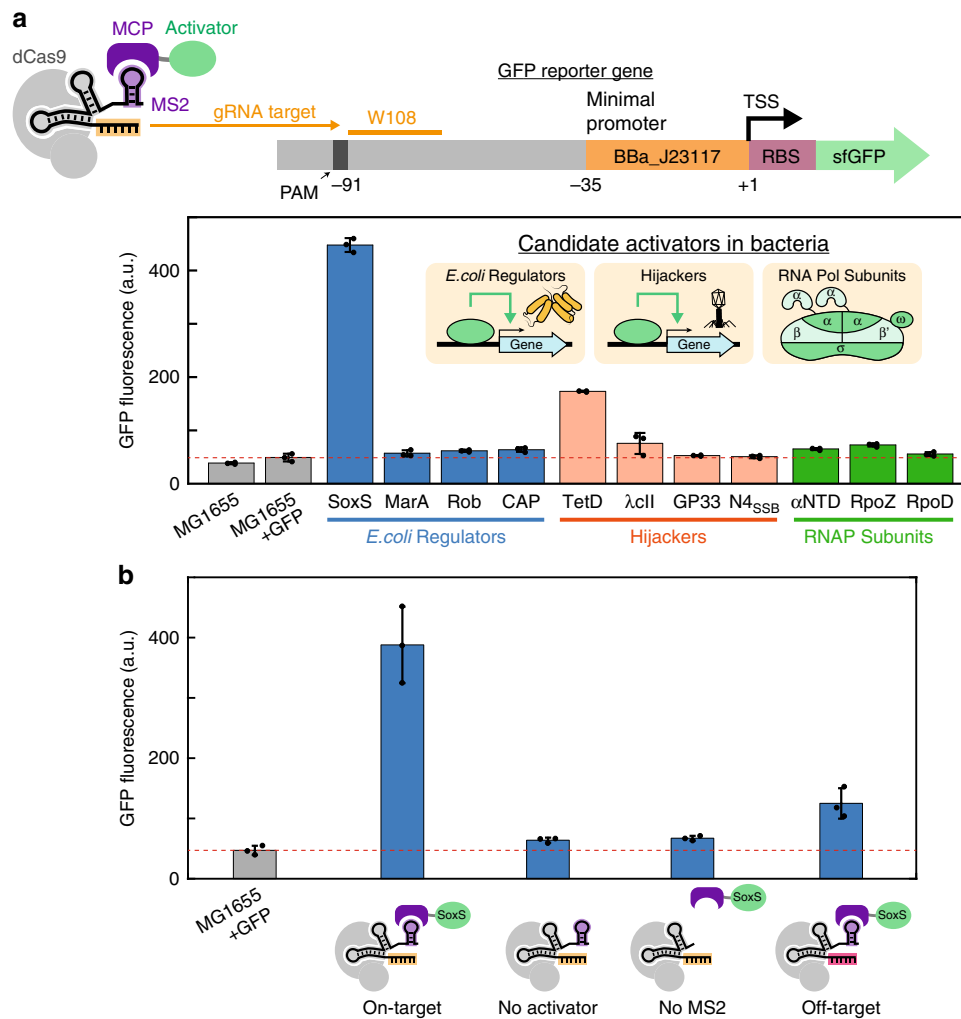


Fig. 2 Effector proteins can activate reporter gene expression in *E. coli*. **a** The CRISPRa complex targets a GFP reporter gene (sfGFP, superfolder GFP) driven by a weak BBa_J23117 promoter. The gRNA target is the W108 sequence, located 91 bases upstream of the transcription start site (TSS). Several candidate activator proteins fused to MCP result in significant increases in GFP expression, including SoxS and TetD. The dotted red line indicates the background fluorescence level observed in the parental MG1655 *E. coli* strain containing the GFP reporter gene. GFP levels were measured in late stationary phase. Similar trends with smaller overall effects were observed in exponential phase. **b** All components of the CRISPRa complex must be present for gene activation. Significant GFP expression is observed when dCas9, the 1x MS2 scRNA, and MCP-SoxS are all expressed. When MCP-SoxS is omitted (no activator) or the MS2 hairpin is removed (no MS2), there is no significant GFP expression. Modest GFP expression is observed when an off-target scRNA is used that has no target site in this strain (RR2, Supplementary Table 2), suggesting that overexpression of SoxS may have some off-target gene activation effects. Values reported are GFP fluorescence levels measured by flow cytometry. Values are median±s.d. for at least three biological replicates (specific values are indicated by black dots)

MarA, Rob, and CAP^{20,21}; hijackers (i.e. bacteriophage or transposon effectors) TetD, λcII, GP33, N4_{SSB}, and AsiA^{22–26}; and RNA polymerase subunits RpoZ (ω), RpoD (σ⁷⁰), and the N-terminal domain of RpoA (αNTD)^{12,13,27}. Of these candidates, αNTD¹², RpoZ^{7,13}, and AsiA²⁶ have been previously reported to activate transcription when fused to heterologous DNA binding domains, and other candidates were selected based on literature reports suggesting that they could recruit RNA polymerase to activate transcription. Each of these candidate proteins was fused to the MS2 coat protein (MCP), which binds to an MS2 hairpin on the scRNA (Fig. 2a)¹⁰.

Several candidate activators produced significant GFP reporter expression in late stationary phase cultures, with the largest effects from SoxS and TetD, and smaller but still detectable GFP expression from λcII, αNTD, and RpoZ (Fig. 2a). Activation with MCP-RpoZ is significantly increased in a Δ*rpoZ* host strain, consistent with that observed previously for other RpoZ fusion

proteins, including dCas9-RpoZ (Supplementary Fig. 1a)^{7,13}. We also obtained significant GFP expression with the T4 bacteriophage activator AsiA (Supplementary Fig. 1b) by co-transforming with a σ⁷⁰ F563Y mutant, which prevents toxicity that occurs when AsiA is expressed alone and inhibits the endogenous σ⁷⁰ subunit²⁶. The most effective activator, SoxS, is a member of the AraC family of transcription factors. SoxS is expressed during oxidative stress and activates expression of a number of genes by binding to RNA polymerase in a pre-recruitment complex and scanning the genome to find DNA targets located in promoter regions^{20,28}. This mechanism of action could explain the effectiveness of SoxS as a candidate activator for CRISPRa.

To confirm that the observed GFP expression arises from recruitment of the candidate activator upstream of the GFP reporter gene, we performed several negative controls with the SoxS activator. First, we demonstrated that activation requires the

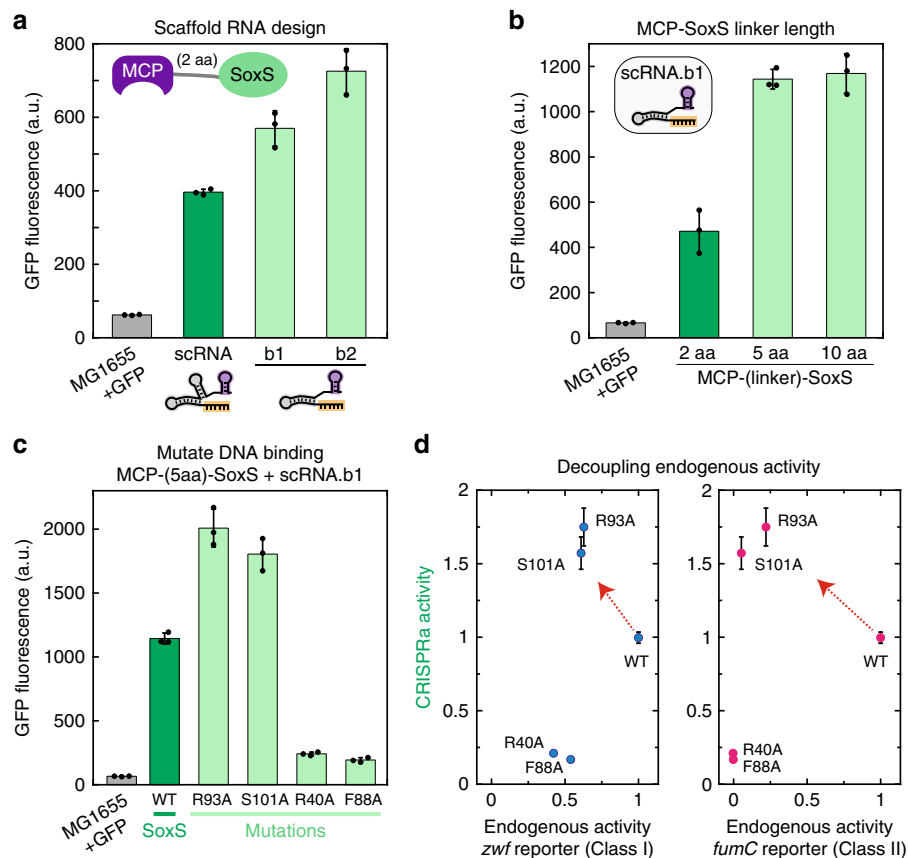


Fig. 3 Optimization of gene activation. **a** Modifying the scRNA design to remove the tracr terminator hairpin results in a 1.5–2-fold increase in gene expression. Complete sequences of the original scRNA, scRNA.b1, and scRNA.b2 designs are included in the Supplementary Methods. scRNA.b1 and scRNA.b2 differ by one base from the 5' end of the terminator hairpin. **b** Increasing the linker length between MCP and SoxS from 2 to 5 or 10 amino acids increases GFP expression by 2–3-fold. **c** Mutations in SoxS at the binding interface to endogenous DNA targets have variable effects on activity in a CRISPRa assay. Point mutations at SoxS R93A or S101A result in a 2-fold increase in GFP expression, while SoxS R40A and F88A lead to substantial decreases in GFP expression. Based on the structure of the SoxS homolog MarA³³, residues R93, S101, and R40 are surface exposed while the F88 side chain points into the hydrophobic core. **d** Plots of CRISPRa activity vs endogenous activity for wild type (wt) and mutant SoxS proteins indicate that transcriptional activation can be decoupled from binding to endogenous targets. CRISPRa activity values are GFP fluorescence levels (**c**) normalized to the value obtained for wt SoxS. Endogenous activity values are from LacZ reporter assays with endogenous SoxS promoters for the *zwf* and *fumC* genes (Supplementary Fig. 3a, b), corrected for background reporter activity in the absence of SoxS and normalized to the value obtained for wt SoxS. *zwf* is representative of class I SoxS target genes in which the SoxS site is located upstream of the –35 site, and *fumC* is representative of class II SoxS target genes in which the SoxS site overlaps the –35 site (Supplementary Fig. 3a). The data shown for CRISPRa and endogenous activity were obtained in separate reporter strains. Similar results were obtained when CRISPRa values were measured in strains with both the GFP and LacZ reporters (Supplementary Fig. 3c). Values reported are GFP fluorescence levels measured by flow cytometry. Values are median±s.d. for at least three biological replicates (specific values are indicated by black dots)

presence of the activator protein. When dCas9 and the 1x MS2 scRNA are expressed without MCP-SoxS, the CRISPR-Cas complex can bind upstream of the GFP reporter, but there is no significant GFP expression (Fig. 2b). Similarly, when the gRNA lacks the 1x MS2 recruitment hairpin, there is no significant GFP activation. Finally, we expressed dCas9, MCP-SoxS, and an off-target 1x MS2 scRNA and observed weak but detectable GFP expression (Fig. 2b). This result suggests that a small fraction of the GFP expression observed with MCP-SoxS arises from non-specific gene activation, which is plausible given that SoxS is an endogenous regulator of transcription in *E. coli* with many gene targets and a relatively degenerate binding site^{29,30}.

Optimizing activity of the SoxS activator. To optimize the activity of CRISPRa with SoxS and minimize off-target effects at endogenous SoxS gene targets, we systematically varied several design parameters of our system. First, we modified the scRNA

structure to optimize it for bacterial expression. In our original design, the MS2 hairpin is appended at the 3' end of the gRNA sequence¹⁰, just downstream of an endogenous tracr terminator hairpin³¹. Removing this terminator hairpin while retaining the 3' MS2 hairpin leads to a 1.5–2-fold increase in GFP expression, likely due to increased steady-state levels of the full length scRNA (Fig. 3a). For comparison, in eukaryotic cells the same modified scRNA lacking the terminator reduces CRISPRa-mediated reporter gene expression ~2-fold (Supplementary Fig. 2a), possibly because the terminator hairpin interacts with Cas9³². Presumably, the same detrimental effect is present in bacteria, but it is outweighed by the benefit from ensuring that the entire scRNA construct is expressed. Using this optimized scRNA design (scRNA.b1), we then varied the length of the amino acid linker connecting MCP and SoxS. 5 and 10 amino acid linkers increased GFP expression by ~2–3-fold compared to our original 2 amino acid linker (Fig. 3b). To increase activity further, we tested an scRNA design with two MS2 hairpins (2x MS2), but observed no significant increase in GFP expression

compared to the 1x MS2 scRNA (Supplementary Fig. 2b). We also tested an alternative gRNA design in which MS2 hairpins are embedded within existing hairpins in the gRNA (sgRNA 2.0)¹¹. We observed no significant GFP expression with sgRNA 2.0 (Supplementary Fig. 2c), which was surprising given that this design is highly effective in eukaryotic cells.

To minimize potential off-target effects at endogenous SoxS gene targets, we attempted to decouple its endogenous DNA binding activity from its transcriptional activation function. We identified several candidate residues at the SoxS DNA binding interface that are known to disrupt activity at endogenous sites when mutated^{33,34}. We expected that if transcriptional activation could be decoupled from endogenous DNA binding, we would retain activity with our CRISPRa system even with SoxS mutants that are defective for activity at endogenous promoters. For two SoxS mutants, R93A and S101A, we observed no loss in activity in a CRISPRa assay; rather, we observed a 2-fold increase in GFP expression compared to wild type (wt) SoxS (Fig. 3c). The increased activity could arise if less SoxS is sequestered at endogenous DNA sites in these mutants. In contrast, the mutations R40A and F88A result in a substantial loss of GFP fluorescence in a CRISPRa assay, which could result from perturbations to protein structure or stability. To assess whether we effectively decoupled transcriptional activation from endogenous DNA binding, we plotted CRISPRa versus endogenous promoter activity for the series of SoxS mutants (Fig. 3d). To directly measure endogenous activity at SoxS gene targets when MCP-SoxS fusion proteins are expressed, we used LacZ fusion constructs for two different SoxS target genes, *zwf* and *fumC* (Supplementary Fig. 3), which are representative of class I and class II SoxS target genes, respectively³⁴. The mutants that retain CRISPRa function, R93A and S101A, are ~2-fold weaker in endogenous activity than wt SoxS at *zwf* and >4-fold weaker at *fumC*. Thus, with SoxS R93A or S101A, we can partially decouple transcriptional activation from endogenous DNA binding. It may be possible to further decouple transcriptional activation from DNA binding with additional SoxS mutations.

We proceeded to test a fully optimized system with the 1xMS2 scRNA.b1 design, a longer 5 amino acid linker between MCP and SoxS, and the SoxS R93A mutant to reduce non-specific activity at endogenous SoxS targets. We observed a substantial increase in GFP fluorescence using an on-target guide RNA and a significant decrease in background activity with an off-target guide RNA (Supplementary Fig. 4) compared to the original, unoptimized system (Fig. 2b). Using RT-qPCR, we observed a 50-fold increase in GFP mRNA levels relative to negative control strains (Supplementary Fig. 5). Importantly, we observed no significant growth burdens associated with expression of the CRISPRa system (Supplementary Fig. 6a, b). Further, while we typically measured GFP levels in late stationary phase, we found similar trends in gene activation with smaller overall effects when GFP levels were measured in exponential phase (Supplementary Fig. 6c). We therefore proceeded with MCP-(5aa)-SoxS_{R93A} and 1x MS2 scRNA.b1 in future experiments. In some cases, we also used the functionally equivalent 1x MS2 scRNA.b2 design (Fig. 3a), which differs by one base (see Supplementary Methods). In direct comparisons, this fully optimized SoxS-based CRISPRa system performs better than the previously reported dCas9-RpoZ bacterial CRISPRa system⁷. Activation with SoxS is effective in an MG1655 strain where dCas9-RpoZ activity is undetectable, and in a $\Delta rpoZ$ MG1655 strain the SoxS-based system outperforms dCas9-RpoZ by >2-fold (Supplementary Fig. 7).

Inducible and simultaneous control of multiple genes. With CRISPR-Cas transcriptional programs, it is straightforward

to target multiple genes for activation or repression^{6,8,10,35}, although simultaneous activation and repression has not yet been demonstrated in bacteria. Using our optimized bacterial CRISPRa system, we tested whether we could simultaneously activate one gene while repressing another. We constructed an *E. coli* strain with two genomically-integrated fluorescent reporter genes, a weakly expressed GFP and a strongly expressed RFP. We expressed dCas9, MCP-(5aa)-SoxS_{R93A}, a 1x MS2 scRNA.b1 for GFP activation, and a gRNA targeting RFP for CRISPRi repression (Fig. 4a). When both the scRNA and the gRNA are expressed, we observe a simultaneous increase in GFP expression and decrease in RFP expression, and the magnitudes of the effects are indistinguishable from those observed when each gRNA is expressed alone (Fig. 4a). Thus, when all components are expressed constitutively, we can effectively control multiple genes simultaneously and have different effects at individual gene targets.

To develop the capability to dynamically regulate multi-gene programs, we first attempted to control the CRISPRa system with inducible promoters. In previous work, using Tet-inducible promoters to control CRISPRi in bacteria required pTet controlling both dCas9 and the gRNA⁶. We tested whether a similar strategy could effectively control CRISPRa. When dCas9 and the 1x MS2 scRNA are controlled by pTet and MCP-(5aa)-SoxS_{R93A} is expressed constitutively, however, we observed leaky GFP expression in the absence of the anhydrotetracycline (aTc) inducer (Fig. 4b), likely due to leaky expression from pTet³⁶. In contrast, when either MCP-(5aa)-SoxS_{R93A} alone or all three components of the CRISPRa system are controlled by pTet, we observed no leaky GFP expression and pTet-inducible GFP levels comparable to that observed with constitutive expression (Fig. 4b). GFP levels for pTet-inductions were measured in late stationary phase cultures; in early stationary phase cultures we observed modestly weaker overall GFP induction compared to that observed with constitutive promoters.

We also tested arabinose-inducible pBAD promoters as an alternative to pTet. With pBAD controlling all three components of the CRISPRa system, we observed inducible GFP expression, but at a level twofold weaker than that observed with constitutive expression. Alternatively, with pBAD controlling both dCas9 and the 1x MS2 scRNA, and constitutive expression of MCP-(5aa)-SoxS_{R93A}, we observed strong inducible GFP expression with no significant leaky expression in the absence of arabinose (Fig. 4b). These results were obtained with cultures in early stationary phase, unlike the results with pTet which were obtained in late stationary phase. In preliminary experiments with arabinose-inducible CRISPRi, we observed leaky repression in late stationary phase cultures, possibly because glucose depletion can relieve catabolite repression of the pBAD promoter³⁷. We therefore performed the arabinose-inducible experiments in early stationary phase cultures, where no leaky CRISPRi repression was observed. Taken together, these results indicate that effective inducible control of CRISPRa can be achieved with pTet controlling MCP-(5aa)-SoxS_{R93A}, or with pBAD controlling dCas9 and the 1x MS2 scRNA.

With inducible control over CRISPRa and CRISPRi, we can implement a two-gene switch from ON/OFF to OFF/ON. We used separate pTet promoters to control all components of the system: dCas9, MCP-(5aa)-SoxS_{R93A}, the 1x MS2 scRNA for GFP, and the gRNA for RFP. Upon addition of aTc, GFP expression increases while RFP expression decreases (Fig. 4c). Alternatively, with pBAD promoters controlling dCas9, the 1x MS2 scRNA for GFP, and the gRNA for RFP, and constitutively expressed MCP-(5aa)-SoxS_{R93A}, we could also inducibly switch GFP on and RFP off (Fig. 4c).

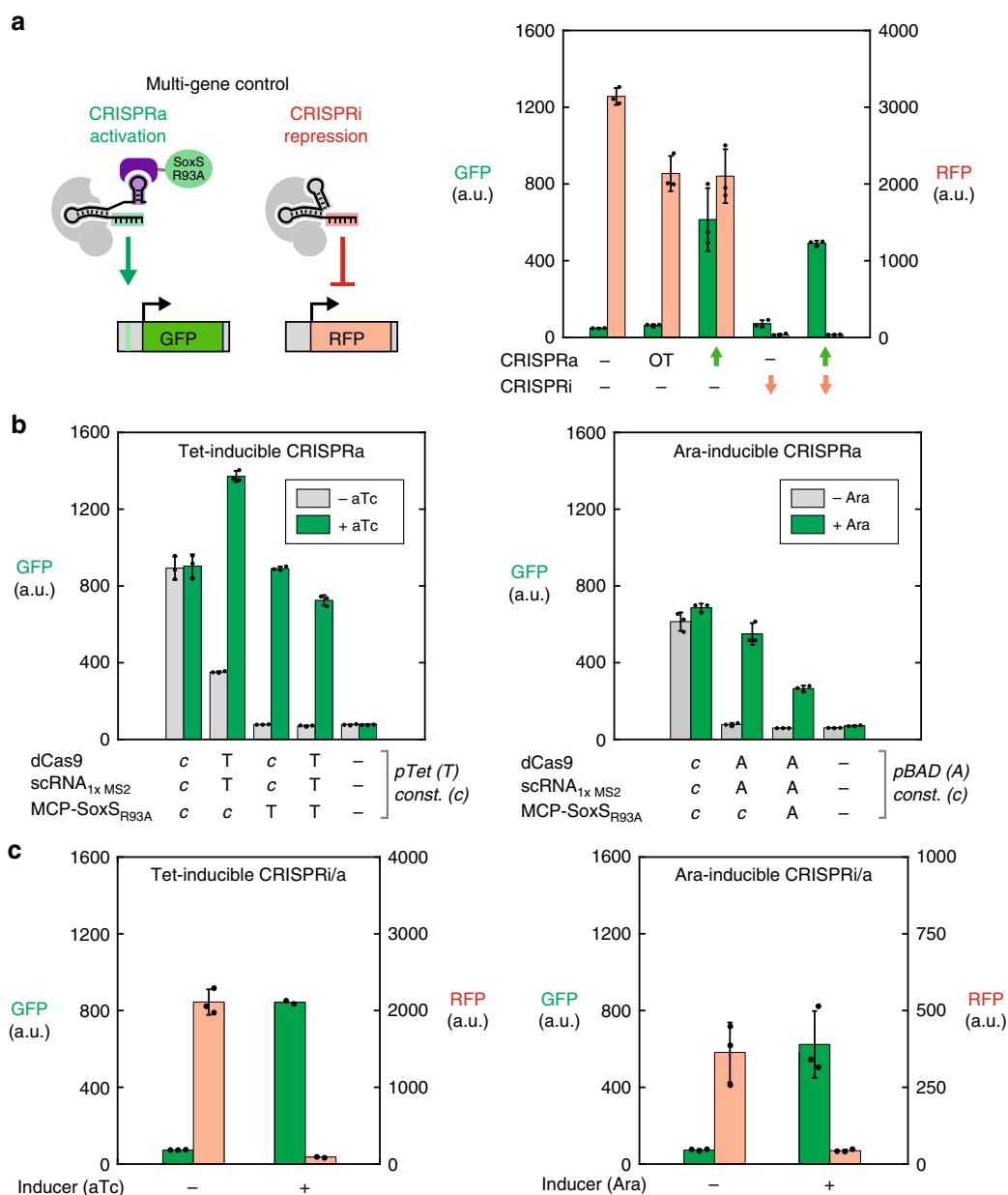


Fig. 4 Multi-gene inducible control. **a** Expressing multiple gRNAs allows simultaneous regulation of multiple genes. Using an *E. coli* strain with integrated GFP and RFP reporters, a 1x MS2 scRNA targets GFP for activation by recruiting MCP-(5aa)-SoxS_{R93A}, while an unmodified gRNA targets RFP for repression. When both gRNAs are expressed, GFP expression increases while RFP expression decreases, and the observed effects are similar to those observed when each gRNA is expressed alone. OT indicates an off-target control for the J106 target site (Supplementary Table 2), which is not present in this strain. Values reported are GFP fluorescence levels measured by flow cytometry. Values are median±s.d. for three biological replicates (specific values are indicated by black dots). **b** CRISPRa can be inducibly controlled with pTet or pBAD (Ara) promoters. In each case, different components of the CRISPRa system (i.e. dCas9, the 1x MS2 scRNA, or MCP-(5aa)-SoxS_{R93A}) are controlled by constitutive (c) or inducible (pTet, T or pBAD, A) promoters. pTet is induced with aTc and pBAD is induced with arabinose. For pTet inductions, cultures were harvested after overnight growth (late stationary phase), while for arabinose inductions cultures were harvested after 6 h (early stationary phase). Values reported are GFP or RFP fluorescence levels measured by flow cytometry. Values are median±s.d. for three biological replicates (specific values are indicated by black dots). **c** CRISPRa and CRISPRi can be simultaneously induced with pTet or pBAD (Ara) promoters. For pTet-induction, all components of the CRISPR system are controlled by pTet. For arabinose induction the pBAD promoter controls dCas9 and the guide RNAs, while MCP-(5aa)-SoxS_{R93A} is constitutively expressed. The RFP axis in the Ara panel is smaller than all other RFP axes in this figure because we observed consistently lower fluorescent protein expression in early stationary phase cultures versus late stationary phase cultures, even in the absence of CRISPRi. Values reported are GFP or RFP fluorescence levels measured by flow cytometry. Values are median±s.d. for three biological replicates, except for the +aTc tet inductions, for which two biological replicates were obtained (specific values are indicated by black dots)

Bacterial CRISPRa is highly sensitive to gRNA target site. In prior reports of bacterial CRISPRa with dCas9-RpoZ, reporter gene expression depended on the distance of the gRNA target from the TSS, with peak expression occurring at a site located ~90 bases upstream⁷. To determine if a similar dependence applies to different activators recruited via 1x MS2 scRNA.b2 to the CRISPR-Cas complex, we designed a synthetic promoter (J1) that has gRNA target sites with the necessary PAM sequences every 10 bases on both strands upstream of a weak BBa_J23117 promoter driving an RFP reporter gene (Fig. 5a). With the MCP-(5aa)-SoxSR_{93A} activator, we observed a sharp dependence of RFP expression on target site position, with significant RFP expression occurring for sites at 80–90 bases upstream of the TSS on the non-template strand and at 60–80 bases upstream on the template strand (Fig. 5b). To determine if activators with weak activity in our initial assays (Fig. 2a) might be more effective at

different target sites, we tested MCP fusion proteins of TetD, λ cII, and α NTD with this promoter. We observed peak activities for these activators at target site positions similar to the most effective sites for SoxS (Fig. 5b), and in no case did we observe dramatic increases in activity relative to our initial assays (Fig. 2a). Similarly, we tested whether an alternative gRNA design (sgRNA 2.0)¹¹ with MS2 sites embedded within internal hairpins in the gRNA and the MCP-(5aa)-SoxSR_{93A} activator would be more effective at different target sites; we observed no detectable CRISPRa activity with this gRNA design at any target site. Finally, we tested an alternative RNA hairpin (1x PP7 scRNA.b1), that recruits the PP7 coat protein (PCP) fused to SoxSR_{93A}. MCP and PCP are structurally homologous proteins, but the MS2 and PP7 RNA hairpins differ significantly and interact in distinct orientations with their cognate binding proteins³⁸. Thus, we do not expect activators recruited by MS2 or PP7 to be positioned in

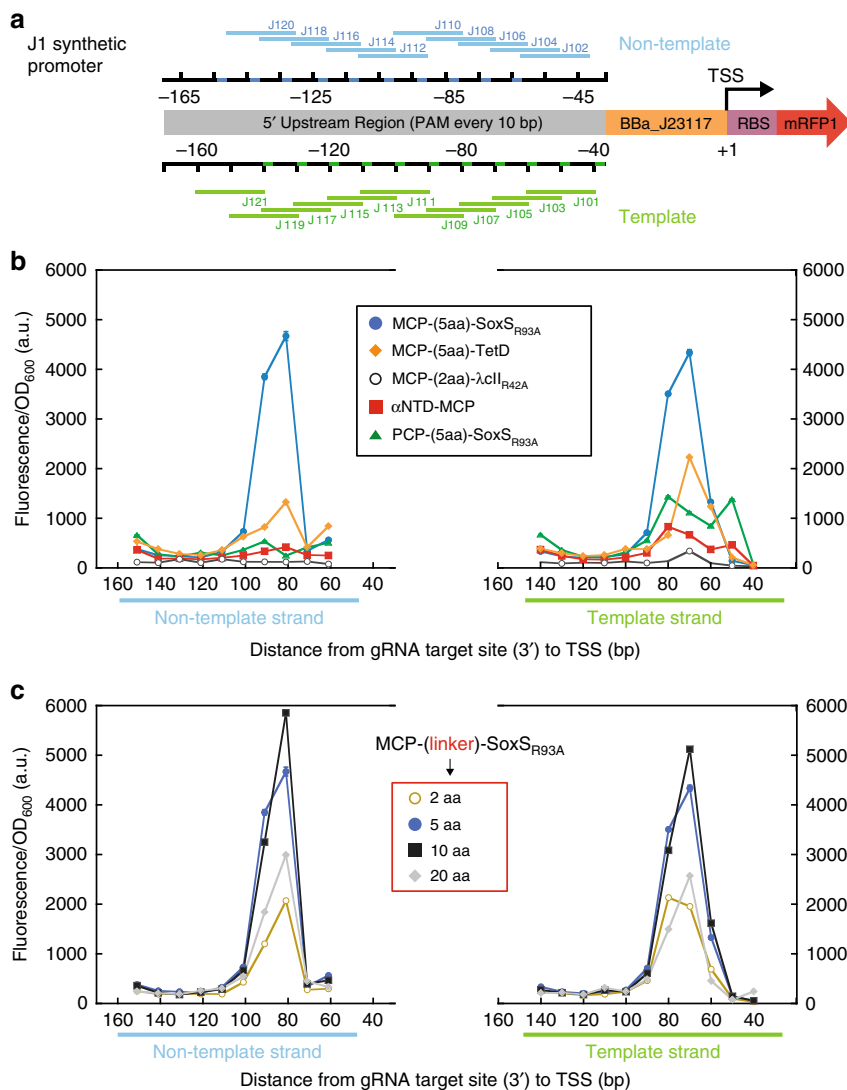


Fig. 5 Gene activation is highly sensitive to CRISPR target site position. **a** The J1 synthetic promoter has potential gRNA target sites with an appropriately positioned PAM sequence every 10 bases on both strands upstream of a weak BBa_J23117 minimal promoter driving an RFP reporter gene. **b** A plot of RFP expression level versus target site position indicates a narrow region for effective CRISPRa from -80 to -90 upstream of the TSS on the non-template strand and from -50 to -80 on the template strand. The plot shows CRISPRa with several different activators recruited by a 1x MS2 scRNA.b2, including MCP-(5aa)-SoxSR_{93A}, MCP-(5aa)-TetD (see Supplementary Fig. 2d), MCP-(2aa)- λ cII_{R42A} (see Supplementary Fig. 2e), and α NTD-MCP. We also used a 1x PP7 scRNA.b1 to recruit PCP-(5aa)-SoxSR_{93A}. **c** Increasing the linker length between MCP and SoxSR_{93A} does not extend the range of gRNA target sites that are effective for CRISPRa. Maximal activity is observed with 5 and 10 amino acid linkers, while both 2 and 20 amino acid linkers show reduced activity but no significant difference in the position of effective target sites. Values reported are RFP fluorescence levels normalized by the cell density (fluorescence/OD₆₀₀) measured in a plate reader. Values are mean \pm s.d. for at least three biological replicates

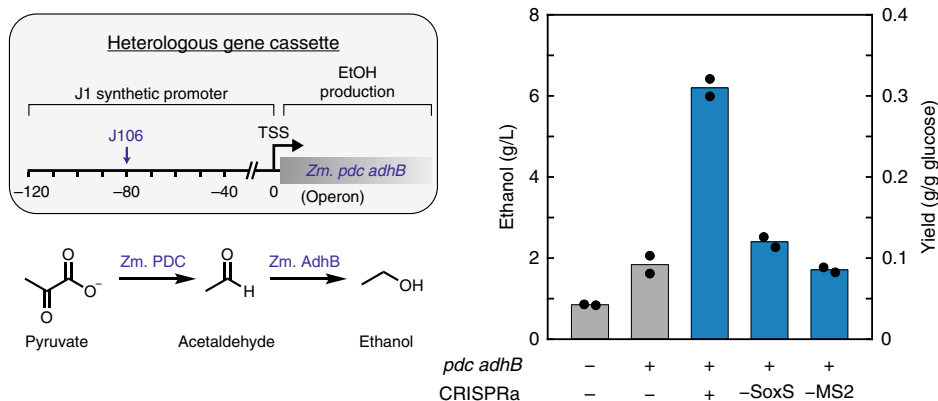


Fig. 6 CRISPRa-mediated ethanol production. A heterologous gene cassette for pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase (*adhB*) from *Z. mobilis* converts pyruvate to ethanol in *E. coli*. The gene cassette is controlled by a weak promoter (see Supplementary Methods). When a CRISPRa complex (dCas9, MCP-(5aa)-SoxS_{R93A}, and 1x MS2 scRNA.b1) is targeted to the J106 site upstream of the promoter, ethanol production increases ~3-fold. In the absence of MCP-(5aa)-SoxS_{R93A} (-SoxS) or with a gRNA lacking the MS2 hairpin (-MS2), ethanol levels are indistinguishable from background strain containing just the *pdc adhB* gene cassette. The left y-axis is ethanol titer (g/L) and the right y-axis is the same data represented as yield (g ethanol/g glucose). Yield is calculated relative to the initial 2% glucose (20 g/L). Values are the mean for two biological replicates (specific values are indicated by black dots)

similar orientations when recruited to the CRISPR-Cas complex. Perhaps because of this structural difference, PCP-(5aa)-SoxS_{R93A} behaved very differently than MCP-(5aa)-SoxS_{R93A}. The PCP construct was relatively ineffective at all sites on the non-template strand and moderately effective on the template strand (Fig. 5b). These results suggest that bacterial CRISPRa is highly sensitive to target position upstream of the TSS, and that it may be possible to find different combinations of activators, recruitment domains, and target sites that perform better than the most effective SoxS constructs tested here. Because all of the activators tested here interact directly with RNA polymerase, albeit at different protein interfaces, we suggest that strong activation requires a CRISPR target site that allows RNA polymerase to bind to the minimal promoter (i.e. the -35 and -10 sites) and the activator simultaneously, which puts a significant constraint on the upstream target sites that can be used for CRISPRa.

To extend the range of effective target sites, we hypothesized that a longer, flexible linker between MCP and SoxS_{R93A} might allow activators to recruit RNA polymerase to the promoter from a broader range of target sites. We therefore tested linkers with 2, 5, 10, or 20 amino acids. Extending the original 5 amino acid linker to 20 amino acids should lengthen the linker by ~57 Å (assuming a worm-like chain model for a flexible peptide linker with 3.8 Å/residue)³⁹. Because target sites on the J1 promoter are spaced ~33 Å apart (10 bp spacing with 3.3 Å/bp for B-DNA), we expect that the extended linker should broaden the range of target sites by at least 10–20 bp in either direction from the optimal range observed with the 5 amino acid linker. Surprisingly, we observed no significant broadening of the effective target site range with longer linkers (Fig. 5c). We observed maximal activity with 5 and 10 amino acid linkers, while both 2 and 20 amino acid linkers show reduced activity but no significant difference in the position of effective target sites. While it remains to be seen if it is possible to design a bacterial CRISPRa complex that functions over a broader target range, simply extending linker lengths between MCP and the SoxS_{R93A} activator was not effective.

Activating a metabolic gene cluster for ethanol biosynthesis.

An important practical application of bacterial CRISPRa will be to enable complex and dynamic multi-gene circuits to control biosynthetic pathways. The capability to express pathway enzymes while repressing competing enzymes may lead to

significant improvements in product yields⁴⁰. As a proof of concept, we tested whether we could upregulate a heterologous *pdc adhB* gene cassette from *Zymomonas mobilis*, which converts pyruvate to ethanol and has previously been used for ethanol production in *E. coli* (Fig. 6)^{41,42}. When we targeted the heterologous gene cassette with our CRISPRa system, we observed a threefold increase in ethanol production relative to cells without CRISPRa (Fig. 6). This result suggests that CRISPRa-based transcriptional control can be used to activate biosynthetic pathways. We expect that combining CRISPRa-based control of heterologous biosynthetic genes with CRISPRi on competing endogenous genes will enable rapid explorations of a large space of genetic circuit architectures to improve biosynthesis yields, and further improvements may be realized by targeting additional metabolic genes for activation or repression, and by coupling the system to dynamically-regulated or inducible promoters.

Discussion

In this work, we have identified multiple synthetic transcriptional activators compatible with CRISPRa in *E. coli*, including SoxS, TetD, and AsiA. The most effective activator, SoxS, is substantially stronger than the previously reported RpoZ-based synthetic activator and does not require additional host-strain modifications for function. SoxS may be an effective activator in part because it interacts with the C-terminal domain of RpoA, which is connected to RNA polymerase by a relatively flexible tether^{12,43}. Using SoxS and other activators, we find a surprisingly sharp dependence on the CRISPR target site distance from the TSS for effective gene activation, with the optimal gRNA sites positioned within a narrow window roughly 60–90 bases upstream of the transcription start site of a heterologous reporter gene. This range is strikingly smaller than the broad target site range observed for effective CRISPRa in eukaryotic cells, where many target sites in the 1–500 base range upstream of the TSS are effective⁴⁴. One notable difference between bacteria and eukaryotes is that the effective bacterial activators identified in this and other work all appear to directly interact with RNA polymerase, while in eukaryotic cells transcription factors may act over longer distances via indirect chromatin modifications. Nevertheless, even given these constraints the ability to perform effective CRISPRa in bacteria will open significant avenues for engineering bacterial systems. Most importantly, we expect that

controlling branch points in metabolic networks with simultaneous gene activation and repression will improve biosynthetic yields beyond that obtained simply by constitutively over-expressing heterologous pathways. Further, we can use inducible promoters to control the timing of gene expression, and we can build more sophisticated dynamic gene expression programs using protein or RNA biosensors to sense and respond to cellular metabolic states^{45,46}.

Two future challenges remain for broad applicability of CRISPRa in bacteria. First, it will be necessary to identify predictive rules for targeting and activating gene expression at endogenous sites. In preliminary experiments, we have observed only modest increases in gene expression at endogenous sites, and the position of effective target sites does not necessarily correspond to the optimal target sites that we observed for heterologous promoter activation. It is possible that our observed distance dependence is not generalizable to different promoters, that endogenous regulatory factors interfere with CRISPRa, or that different types of promoters require distinct transcriptional activation domains. In practice, if predictive rules for targeting arbitrary endogenous genes remain elusive, we envision using gene editing to introduce heterologous promoters at target sites of interest, which will enable us to regulate endogenous genes as part of a multi-gene CRISPRi/a control program.

A second outstanding challenge for bacterial CRISPRa is to develop a system that is portable across bacterial species of high commercial and industrial value, such as strains that have the ability to utilize carbon sources like CO₂, CO, methane, or lignocellulose, and alternative energy sources such as H₂ or light^{1,2}. For gene repression, CRISPRi has been successfully used in a broad range of bacterial species^{16,47–50}, suggesting that the programmable DNA targeting component of the system is portable. For CRISPRa, many of our candidate gene activators interact with motifs on bacterial RNA polymerase that are highly conserved across a broad range of bacteria⁵¹, including several that are relevant for industrial biosynthesis or microbiome engineering. In particular, SoxS interacts with a surface on the RNA polymerase α subunit that is well-conserved in gammaproteobacteria, alphaproteobacteria, bacteroides, gram-positive bacteria, and even to a lesser extent in cyanobacteria (Supplementary Fig. 8). This conserved interface may allow CRISPRa systems developed in *E. coli* to be ported to non-model bacteria with a wide range of useful biological functions.

Methods

Bacterial strain construction and manipulation. Plasmid constructs were cloned and *E. coli* cells were cultured using standard molecular biology methods. Gene knockouts and integrations of fluorescent reporters were performed by recombineering in a strain with a genomically-integrated lambda red system under the control of a temperature-sensitive promoter (NM700)^{52–54}. Modified genome fragments were then transferred to MG1655 by P1 transduction⁵⁵. The sfGFP reporter was integrated at the *rfsA* locus. The mRFP reporter (for CRISPRi experiments) was integrated at *rbsAR*⁵⁶. The Δ rpoZ strain was constructed by recombineering a pKD13-derived KanR linear PCR cassette with 36 base homology overhangs to flanking sites at the *rpoZ* locus. The Δ rpoZ::KanR knockout genomic fragment was transferred to MG1655 by P1 transduction, and the FRT-flanked KanR cassette was eliminated by transformation with the pCP20 helper plasmid to express FLP recombinase, followed by incubation at 42 °C to cure the plasmid. CRISPR-Cas system components were delivered on plasmids as described in Supplementary Table 3. *S. pyogenes* dCas9 was expressed from its endogenous *S. pyogenes* Cas9 promoter (cloned from pWJ66, addgene #46570). Candidate effector proteins fused to RBPs (i.e. MCP and PCP) were expressed with the medium-strength BBa_J23107 promoter. Guide RNAs were expressed from the strong BBa_J23119 promoter⁶. BBa sequences are available from the Repository of Standard Biological Parts (<http://parts.igem.org>). *zwf*-LacZ and *fumC*-LacZ reporter genes were constructed following previously described designs (see Supplementary Methods)^{57,58}. Complete annotated sequences of the reporter genes and CRISPR-Cas system components are included in the Supplementary Methods along with a list the engineered *E. coli* strains (Supplementary Table 1). Guide RNA target sites are listed in Supplementary Table 2.

Flow cytometry. Cells were inoculated in EZ-RDM (Teknova) supplemented with appropriate antibiotics and grown in 96-deep-well plates at 37 °C, 220 RPM overnight. Late stationary phase cultures were then diluted 1:40 in PBS and analyzed on a MACSQuant VYB flow cytometer (Miltenyi Biotec). To enrich for single cells, a side scatter threshold trigger (SSC-H) was applied. To gate for single bacterial cells, we first selected events along the diagonal of the SSC-H vs. SSC-A plot⁵⁹. We then excluded events that appeared on the edges of the SSC-A vs. FSC-A plot, and events that appeared on the edge of the fluorescence histogram (Supplementary Fig. 9).

For inducible CRISPR system construction with pTet or pBAD promoters, strains were inoculated in 3 mL LB medium supplemented with antibiotics and grown overnight at 37 °C, 220 RPM. On the next day, they were diluted 1:100 in 500 μ L EZ-RDM supplemented with antibiotics and induced with 1 μ M anhydrotetracycline (aTc) for pTet or 100 mM L-arabinose for pBAD. Non-induced controls were prepared for each strain. For pTet, cells were grown at 37 °C, 220 RPM overnight. For pBAD, cells were grown at 37 °C, 220 RPM for 6 h. In both cases, fluorescence was assayed via flow cytometry as described above.

LacZ reporter assays. LacZ reporter assays for SoxS activation of *zwf* and *fumC* promoters were performed following a previously reported protocol with minor modifications^{60,61}. Specifically, cultures were grown for 18 h in EZ-RDM supplemented with antibiotics. Absorbance at 600 nm (OD₆₀₀) was measured from 150 μ L samples in a Biotek Synergy HTX plate reader. 20 μ L aliquot from each culture were added to 80 μ L permeabilization solution [100 mM Na₂HPO₄, 20 mM KCl, 2 mM MgSO₄, 0.8 mg/mL hexadecyltrimethylammonium bromide, 0.4 mg/mL sodium deoxycholate, 5.4 μ L/mL β -mercaptoethanol]. Samples were incubated at 30 °C for 20 min. To initiate the reaction, 600 μ L of substrate solution [60 mM Na₂HPO₄, 40 mM NaH₂PO₄, 1 mg/mL *o*-nitrophenyl- β -D-galactopyranoside, 2.7 μ L/mL β -mercaptoethanol] was added. Samples were incubated at 30 °C until visible color developed, at which point 700 μ L of stop solution [1 M Na₂CO₃] was added and the reaction time was recorded. Samples were centrifuged for 10 min to pellet cell debris, and the supernatants were removed. Absorbance at 420 nm (*A*₄₂₀) was measured from 150 μ L samples in a plate reader. LacZ activity was calculated according to the formula: Miller Units = (1000 \times *A*₄₂₀) / (OD₆₀₀ \times 0.02 mL \times time).

J1 reporter sequence design. Custom Python scripts were used to generate a tiling array containing an NGG PAM site every 10 nucleotides on each strand according to the following base unit: 5'-NNNCCNNNGG-3'. 10000 sequences of length 500 nt were generated by randomly sampling N nucleotides, adjusted for endogenous GC content in *E. coli* (BNID '100528 [<http://bionumbers.hms.harvard.edu/bionumber.aspx?id=100528>])⁶². We discarded any sequences with four consecutive identical nucleotides. Sequences containing known transcription factor binding sites (using the experimental TF binding site dataset from RegulonDB)⁶³ were also discarded. Of the remaining sequences, we arbitrarily chose one, labeled as J1, and confirmed that it had no detectable homology to the *E. coli* genome. We placed a 170 bp fragment of this sequence upstream of the weak BBa_J23117 constitutive promoter (<http://parts.igem.org>).

Plate reader experiments. For fluorescent reporter experiments with the J1 reporter, cells were inoculated in 2 mL EZ-RDM supplemented with appropriate antibiotics and grown at 37 °C, 220 RPM overnight. OD₆₀₀ and observed fluorescence values were measured in a Biotek Synergy HTX plate reader using 150 μ L of the overnight culture in flat, clear-bottomed 96-well plates (Corning). For mRFP detection, the excitation wavelength was 540 nm and the emission wavelength was 600 nm.

Ethanol fermentations. *E. coli* MG1655 cells were transformed with the pCD355 plasmid containing the *pdC adhB* gene cassette from *Z. mobilis* under the control of the CRISPRa system. Transformed cells were inoculated in 5 mL LB (2% glucose) supplemented with appropriate antibiotics and grown overnight at 37 °C, 220 RPM. OD₆₀₀ measurements were taken from the overnight cultures, and 2.5 OD₆₀₀-mL of the culture was diluted into 50 mL LB supplemented with antibiotics. Cultures were grown aerobically for 18 h at 37 °C, 220 RPM. OD₆₀₀ measurements were taken, and 150 OD₆₀₀-mL was pelleted and resuspended into 25 mL M9 media with 2% dextrose and antibiotics in a 125 mL conical flask. Flasks were sealed with rubber septa with a needle connected to a balloon to relieve pressure from CO₂ production. Cultures were incubated at 37 °C without shaking for 4 days. Ethanol concentrations in supernatants were measured with an Ethanol assay kit (R-Biopharm). We counted viable cells by plating on selective media before and after fermentation and observed no decrease in viable cells after the fermentation. To evaluate plasmid stability over the course of the fermentation, we performed minipreps on cells after the fermentation. Analytical restriction digests indicated no large-scale recombination events, and no point mutations were detected by sequencing.

Quantitative RT-PCR. Strains were inoculated in 5 mL LB supplemented with appropriate antibiotics and grown overnight at 37 °C, 220 RPM. Cultures were then

diluted 1:100 in 5 mL EZ-RDM supplemented with antibiotics and grown to OD₆₀₀ 0.5 (using 150 µL samples in flat clear bottomed 96-well plates in a Biotek Synergy HTX plate reader). Cultures were pelleted, flash-frozen in liquid nitrogen, and stored at -80 °C. Total RNA was extracted using an Aurum Total RNA Mini Kit (Bio-Rad). 1 µg of RNA was converted to cDNA using iScript reverse transcriptase (Bio-Rad) in 20 µL reactions. qPCR was performed using SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) with a 58 °C annealing temperature and 10 µL reaction volumes. qPCR reactions were performed in triplicate on a CFX Connect (Bio-Rad) using 0.5 ng of cDNA, 400 nM primer concentration, and 15 s extension time. 16 S rRNA was used for normalization. Control samples without a template and without reverse transcriptase were analyzed to confirm gene-specific amplification and absence of gDNA contamination. Primer sequences are listed in Supplementary Table 4. Expression levels for each gene were calculated relative to 16 S using the $\Delta\Delta C_T$ method¹⁶⁴.

Code availability. Custom Python code to generate tiling arrays with PAM sequences (detailed in “J1 reporter sequence design” section of Methods) is available upon request.

Data availability. All data from this study are available upon request.

Received: 5 February 2018 Accepted: 1 June 2018

Published online: 27 June 2018

References

- Haynes, C. A. & Gonzalez, R. Rethinking biological activation of methane and conversion to liquid fuels. *Nat. Chem. Biol.* **10**, 331–339 (2014).
- Lan, E. I. & Liao, J. C. Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources. *Bioresour. Technol.* **135**, 339–349 (2013).
- Mimee, M., Tucker, A. C., Voigt, C. A. & Lu, T. K. Programming a human commensal bacterium, *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota. *Cell Syst.* **1**, 62–71 (2015).
- Whitaker, W. R., Shepherd, E. S. & Sonnenburg, J. L. Tunable expression tools enable single-cell strain distinction in the gut microbiome. *Cell* **169**, 538–546. e12 (2017).
- Riglar, D. T. et al. Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* **35**, 653–658 (2017).
- Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
- 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).
- Gilbert, L. A. et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451 (2013).
- Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. *Nat. Methods* **10**, 957–963 (2013).
- Zalatan, J. G. et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339–350 (2015).
- Konermann, S. et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).
- Dove, S. L., Joung, J. K. & Hochschild, A. Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**, 627–630 (1997).
- Dove, S. L. & Hochschild, A. Conversion of the omega subunit of *Escherichia coli* RNA polymerase into a transcriptional activator or an activation target. *Genes. Dev.* **12**, 745–754 (1998).
- Dove, S. L. & Hochschild, A. A bacterial two-hybrid system based on transcription activation. *Methods Mol. Biol.* **261**, 231–246 (2004).
- Otupal, P. B., Erickson, K. E., Escalas-Bordoy, A. & Chatterjee, A. CRISPR perturbation of gene expression alters bacterial fitness under stress and reveals underlying epistatic constraints. *ACS Synth. Biol.* **6**, 94–107 (2017).
- Peters, J. M. et al. Bacterial CRISPR: accomplishments and prospects. *Curr. Opin. Microbiol.* **27**, 121–126 (2015).
- Choi, K. R. & Lee, S. Y. CRISPR technologies for bacterial systems: Current achievements and future directions. *Biotechnol. Adv.* **34**, 1180–1209 (2016).
- Peng, R. et al. CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in *Mycococcus xanthus*. *Microb. Cell. Fact.* **17**, 15 (2018).
- Dominguez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* **17**, 5–15 (2016).
- Martin, R. G., Gillette, W. K., Martin, N. I. & Rosner, J. L. Complex formation between activator and RNA polymerase as the basis for transcriptional activation by MarA and SoxS in *Escherichia coli*. *Mol. Microbiol.* **43**, 355–370 (2002).
- Busby, S. & Ebricht, R. H. Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* **293**, 199–213 (1999).
- Griffith, K. L., Becker, S. M. & Wolf, R. E. Characterization of TetD as a transcriptional activator of a subset of genes of the *Escherichia coli* SoxS/MarA/Rob regulon. *Mol. Microbiol.* **56**, 1103–1117 (2005).
- Jain, D. et al. Crystal structure of bacteriophage lambda cII and its DNA complex. *Mol. Cell* **19**, 259–269 (2005).
- Twist, K.-A. F. et al. Crystal structure of the bacteriophage T4 late-transcription coactivator gp33 with the β -subunit flap domain of *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA* **108**, 19961–19966 (2011).
- Miller, A., Wood, D., Ebricht, R. H. & Rothman-Denes, L. B. RNA polymerase β' subunit: a target of DNA binding-independent activation. *Science* **275**, 1655–1657 (1997).
- Gregory, B. D., Deighan, P. & Hochschild, A. An artificial activator that contacts a normally occluded surface of the RNA polymerase holoenzyme. *J. Mol. Biol.* **353**, 497–506 (2005).
- Gruber, T. M. & Gross, C. A. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* **57**, 441–466 (2003).
- Griffith, K. L. & Wolf, R. E. Genetic evidence for pre-recruitment as the mechanism of transcription activation by SoxS of *Escherichia coli*: the dominance of DNA binding mutations of SoxS. *J. Mol. Biol.* **344**, 1–10 (2004).
- Pomposiello, P. J., Bennik, M. H. & Demple, B. Genome-wide transcriptional profiling of the *Escherichia coli* responses to superoxide stress and sodium salicylate. *J. Bacteriol.* **183**, 3890–3902 (2001).
- Griffith, K. L. & Wolf, R. E. Systematic mutagenesis of the DNA binding sites for SoxS in the *Escherichia coli* *zwf* and *fpr* promoters: identifying nucleotides required for DNA binding and transcription activation. *Mol. Microbiol.* **40**, 1141–1154 (2001).
- Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
- Nishimasu, H. et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).
- Rhee, S., Martin, R. G., Rosner, J. L. & Davies, D. R. A novel DNA-binding motif in MarA: the first structure for an AraC family transcriptional activator. *Proc. Natl Acad. Sci. USA* **95**, 10413–10418 (1998).
- Griffith, K. L. & Wolf, R. E. A comprehensive alanine scanning mutagenesis of the *Escherichia coli* transcriptional activator SoxS: identifying amino acids important for DNA binding and transcription activation. *J. Mol. Biol.* **322**, 237–257 (2002).
- Cheng, A. W. et al. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* **23**, 1163–1171 (2013).
- Fontana, J., Dong, C., Ham, J. Y., Zalatan, J. G. & Carothers, J. M. Regulated expression of sgRNAs tunes CRISPRi in *E. coli*. *Biotechnol. J.* **13**, 1800069 (2018).
- Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **177**, 4121–4130 (1995).
- Chao, J. A., Patskovsky, Y., Almo, S. C. & Singer, R. H. Structural basis for the coevolution of a viral RNA-protein complex. *Nat. Struct. Mol. Biol.* **15**, 103–105 (2008).
- Zhou, H.-X. Polymer models of protein stability, folding, and interactions. *Biochemistry* **43**, 2141–2154 (2004).
- Tan, S. Z. & Prather, K. L. Dynamic pathway regulation: recent advances and methods of construction. *Curr. Opin. Chem. Biol.* **41**, 28–35 (2017).
- Ingram, L. O., Conway, T., Clark, D. P., Sewell, G. W. & Preston, J. F. Genetic engineering of ethanol production in *Escherichia coli*. *Appl. Environ. Microbiol.* **53**, 2420–2425 (1987).
- Alterthum, F. & Ingram, L. O. Efficient ethanol production from glucose, lactose, and xylose by recombinant *Escherichia coli*. *Appl. Environ. Microbiol.* **55**, 1943–1948 (1989).
- Blatter, E. E., Ross, W., Tang, H., Gourse, R. L. & Ebricht, R. H. Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**, 889–896 (1994).
- Gilbert, L. A. et al. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647–661 (2014).
- Zhang, F., Carothers, J. M. & Keasling, J. D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* **30**, 354–359 (2012).
- Skjoedt, M. L. et al. Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat. Chem. Biol.* **12**, 951–958 (2016).
- Peters, J. M. et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* **165**, 1493–1506 (2016).

48. Yao, L., Cengic, I., Anfelt, J. & Hudson, E. P. Multiple gene repression in cyanobacteria using CRISPRi. *ACS Synth. Biol.* **5**, 207–212 (2016).
49. Cleto, S., Jensen, J. V., Wendisch, V. F. & Lu, T. K. *Corynebacterium glutamicum* metabolic engineering with CRISPR interference (CRISPRi). *ACS Synth. Biol.* **5**, 375–385 (2016).
50. Tan, S. Z., Reisch, C. R. & Prather, K. L. J. A robust CRISPR interference gene repression system in *Pseudomonas*. *J. Bacteriol.* **200**, e00575–17 (2018).
51. Murakami, K. S. Structural biology of bacterial RNA polymerase. *Biomolecules* **5**, 848–864 (2015).
52. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
53. Yu, D. et al. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **97**, 5978–5983 (2000).
54. Court, D. L. et al. Mini-lambda: a tractable system for chromosome and BAC engineering. *Gene* **315**, 63–69 (2003).
55. Thomason, L. C., Costantino, N. & Court, D. L. *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.* <https://doi.org/10.1002/0471142727.mb0117s79> (2007).
56. Sabri, S., Steen, J. A., Bongers, M., Nielsen, L. K. & Vickers, C. E. Knock-in/ Knock-out (KIKO) vectors for rapid integration of large DNA sequences, including whole metabolic pathways, onto the *Escherichia coli* chromosome at well-characterised loci. *Microb. Cell Fact.* **12**, 60 (2013).
57. Fawcett, W. P. & Wolf, R. E. Genetic definition of the *Escherichia coli* *zwf*⁺soxbox⁺, the DNA binding site for SoxS-mediated induction of glucose 6-phosphate dehydrogenase in response to superoxide. *J. Bacteriol.* **177**, 1742–1750 (1995).
58. Fawcett, W. P. & Wolf, R. E. Purification of a MalE-SoxS fusion protein and identification of the control sites of *Escherichia coli* superoxide-inducible genes. *Mol. Microbiol.* **14**, 669–679 (1994).
59. Shapiro, H. M. Multiparameter flow cytometry of bacteria: implications for diagnostics and therapeutics. *Cytometry* **43**, 223–226 (2001).
60. Zhang, X. & Bremer, H. Control of the *Escherichia coli* *rrnB* P1 promoter strength by ppGpp. *J. Biol. Chem.* **270**, 11181–11189 (1995).
61. Martens, A. T. Beta-Galactosidase Assay (A better Miller). *OpenWetWare* [https://openwetware.org/mediawiki/index.php?title=Beta-Galactosidase_Assay_\(A_better_Miller\)&oldid=620416](https://openwetware.org/mediawiki/index.php?title=Beta-Galactosidase_Assay_(A_better_Miller)&oldid=620416) (2012).
62. Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers - the database of key numbers in molecular and cell biology. *Nucleic Acids Res.* **38**, D750–D753 (2010).
63. Gama-Castro, S. et al. RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res.* **44**, D133–D143 (2016).
64. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method. *Methods* **25**, 402–408 (2001).

Acknowledgements

We thank Maureen Thomason, Mary Lidstrom, Frances Chu, Willy Voje, Jason Stevens, Chuhern Hwang, and members of the Zalatan and Carothers groups for technical assistance, advice, and helpful discussions. The NM700 strain was a gift from Nadim Majdalani and Susan Gottesman. This work was supported by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund (J.G.Z.), an NSF Award MCB 1517052 (J.M.C.), and a University of Washington Presidential Innovation Award (J.M.C.)

Author contributions

J.G.Z. conceived the project. C.D., J.F., J.M.C., and J.G.Z. designed experiments, analyzed data, and wrote the manuscript. C.D., J.F., and A.P. performed experiments.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-04901-6>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018