

Rational engineering of a modular bacterial CRISPR–Cas activation platform with expanded target range

Maria Claudia Villegas Kcam¹, Annette J. Tsong¹ and James Chappell^{1,2,*}

¹Department of BioSciences, Rice University, 6100 Main Street, MS 140, Houston, TX 77005, USA and ²Department of Bioengineering, Rice University, 6100 Main Street, MS 142, Houston, TX 77005, USA

Received November 24, 2020; Revised March 10, 2021; Editorial Decision March 11, 2021; Accepted March 19, 2021

ABSTRACT

CRISPR–Cas activator (CRISPRa) systems that selectively turn on transcription of a target gene are a potentially transformative technology for programming cellular function. While in eukaryotes versatile CRISPRa systems exist, in bacteria these systems suffer from a limited ability to activate different genes due to strict distance-dependent requirements of functional target binding sites, and require greater customization to optimize performance in different genetic and cellular contexts. To address this, we apply a rational protein engineering approach to create a new CRISPRa platform that is highly modular to allow for easy customization and has increased targeting flexibility through harnessing engineered Cas proteins. We first demonstrate that transcription activation domains can be recruited by CRISPR–Cas through noncovalent protein–protein interactions, which allows each component to be encoded on separate and easily interchangeable plasmid elements. We then exploit this modularity to rapidly screen a library of different activation domains, creating new systems with distinct regulatory properties. Furthermore, we demonstrate that by harnessing a library of circularly permuted Cas proteins, we can create CRISPRa systems that have different target binding site requirements, which together, allow for expanded target range.

INTRODUCTION

Since their discovery, RNA-guided CRISPR–Cas ribonuclease systems have transformed many aspects of biological sciences, biotechnology and medicine (1–3). In particular, engineering of CRISPR–Cas systems has led to a powerful suite of *trans*-acting gene regulatory tools able to precisely program gene expression in a diversity of cellular con-

texts. Owing to the ease of designing and synthesizing guide RNAs (gRNAs) that determine target specificity, CRISPR–Cas regulators have found broad utility for performing targeted gene perturbation studies to uncover gene function (2,4–10) and connectivity (2,4–6,11,12), and to rapidly optimize strains for bioprocesses (13–20). Additionally, creation of orthogonal and composable CRISPR–Cas controlled promoter elements has enabled construction of genetic circuits able to perform cellular computations and signal processing (21–26).

In bacteria, the first iteration of CRISPR–Cas regulators was CRISPR interference (CRISPRi), which uses a nuclease-dead Cas9 (dCas9) and a single guide RNA (sgRNA) complex to sterically block transcription initiation and elongation (27,28). Extending this, further innovations have sought to use the CRISPR–dCas9 complex as a localization platform and to control gene expression through recruitment of protein effector domains. For example, CRISPR–Cas activators (CRISPRa) have been created by harnessing CRISPR–dCas9 systems to recruit protein activation domains (ADs) that stimulate transcription when localized in proximity to promoter elements. Different engineering approaches have been used to obtain CRISPRa systems, such as performing fusions of ADs directly to dCas9 (28,29) or engineering sgRNAs to recruit ADs (30).

While in eukaryotes CRISPRa systems have provided robust tools for activating transcription that have resulted in the widespread use of these technologies, in bacteria they have not had the same impact because of several technical limitations. First, it is increasingly clear that currently available ADs offer distinct regulatory properties and trade-offs, and different ADs are likely to be optimal depending on the exact gene target or application. For example, some bacterial ADs, such as those derived from the transcription factor SoxS (31,32) or RNAP subunits (28), show highest activation when localized immediately upstream (60–100 bp) of a target gene's transcription start site (TSS), while other ADs, such as those based on the AsiA anti-sigma factor (33), operate at longer distances (>100–190 bp upstream). ADs can also have different trade-offs between fold of activation and

*To whom correspondence should be addressed. Tel: +1 713 348 3781; Email: jc125@rice.edu

other performance criteria, such as off-target effects, growth inhibition and requirement of knockout strains (28,31,33). Specialized ADs are also required to activate promoter elements that utilize alternative sigma factors (34) or to allow creation of CRISPRa systems for different bacterial species (35,36). Unlike eukaryotic CRISPRa systems that have largely relied upon a handful of ADs (29,37,38) to achieve activation in diverse genetic and cellular contexts (37), bacterial CRISPRa systems likely require a greater repertoire of components and it remains unclear if an optimal AD has been identified. A second challenge of bacterial CRISPRa systems are distance-dependent activation patterns that significantly limit functional CRISPRa binding sites. Specifically, it was recently shown that bacterial CRISPRa systems that use sgRNA-based recruitment have activation patterns that are periodical, in which strong activation is only seen when the system is targeted within a 2–4 bp window that repeats every 10–11 bp from the TSS (32,34). This significantly limits the targeting range as only genes with correctly positioned upstream binding sites can be activated.

Bacterial CRISPRa systems represent a potentially transformative technology; however, work remains to be done to create a versatile tool. Towards this goal, we apply a rational protein engineering approach to create a new CRISPRa platform that is highly modular, allowing for facile exchange of new ADs, and presents expanded target range, through the exchange of different engineered dCas9 variants with different binding site requirements. First, recognizing that different ADs have distinct properties, we characterize a panel of ADs and identify the N-terminal domain of the alpha subunit of RNAP (α NTD) to robustly activate compared to previously reported ADs. Next, we explore modular approaches to recruit ADs to dCas9 through noncovalent protein-protein interaction sequences called SYNZIP domains (39,40). This modular approach presents comparable fold activation levels as using covalent linkers while allowing ADs to be encoded on separate and interchangeable plasmids, therefore allowing for facile exchange with dCas9-encoding plasmids. We exploit this enhanced modularity to rapidly characterize a panel of α NTD ADs derived from diverse bacterial species. Finally, to overcome the limitations in targeting range, we utilize circularly permuted variants of dCas9 (cpdCas9) that allow localization of ADs to distinct positions within the tertiary structure, and as a result, shift activation patterns. By screening a library of cpdCas9 variants, we identify novel CRISPRa systems that can activate from distinct positions compared to wild-type dCas9, thus expanding the number of targetable promoter elements. Taken together, our CRISPRa platform offers a high degree of modularity, in which plasmids encoding different dCas9 variants can be synergistically combined with different ADs to create unique combinations optimized for different gene targets and applications.

MATERIALS AND METHODS

Plasmid assembly

All plasmids used in this study can be found in Supplementary Table S1 with key sequences provided in Supplementary Table S2–S7. A schematic of representative plas-

mids used is shown in Supplementary Figure S1. dCas9 variants were expressed with a P_{tet} , P_{bad} or a constitutive J23150 promoter in a p15a chloramphenicol resistant vector. sgRNAs were expressed with a J23119 promoter in a ColE1 ampicillin resistant vector. Constructs containing ADs fused to SYNZIP were expressed with J23106 promoter in a CloDF spectinomycin resistant vector. Reporter plasmids were cloned from pJF076Sa plasmid (addgene #113322) (31,32)—an RFP reporter with J23117 promoter and upstream NGG-rich sequence—into a kanamycin resistant pSC101 vector. Constitutive promoter sequences are available at the iGEM Registry of Standard Biological Parts (parts.igem.org). Plasmids were assembled through a combination of PCR ligations, Gibson assembly (41) or Golden Gate assembly (42).

Fluorescence measurements

Fluorescence characterization was performed using *Escherichia coli* strain MG1655, except for data shown in Figure 1B where *E. coli* K-12 BW25113 and *E. coli* K-12 BW25113 $\Delta rpoZ$ (JW3624-1, Keio collection) was used (43). Each experiment was performed with four biological replicates. For each condition, plasmid combinations were transformed into chemically competent *E. coli* cells and plated on LB + Agar (Difco) plates containing combinations of 100 $\mu\text{g ml}^{-1}$ carbenicillin, 34 $\mu\text{g ml}^{-1}$ chloramphenicol, 25 $\mu\text{g ml}^{-1}$ kanamycin and 50 $\mu\text{g ml}^{-1}$ spectinomycin, depending on the plasmids used, and incubated ~ 17 h overnight at 37°C. In some experiments involving transformation of ≥ 3 plasmids, a subset of plasmids was first transformed, competent cells prepared from resulting colonies and the remaining plasmid transformed. Plates were taken out of the incubator and left at room temperature for ~ 7 h. Single colonies were used to inoculate four cultures of 300 μl of LB containing antibiotics at the concentrations described above in a 2 ml 96-well block (Costar). 96-well blocks were incubated overnight in a Vortemp 56 (Labnet) incubator bench top shaker at 37°C and 1000 RPM. 10 μl of the overnight culture were added to 290 μl of LB, with antibiotics and 1 ng ml^{-1} of anhydrotetracycline (aTc) inducer or 50 mM of arabinose when required. For Supplementary Figure S2, acyl-homoserine lactone (AHL) [*N*-(β -ketocaproyl)-L-homoserine lactone, Cayman chemical] solution was also added with a final concentration of 10 and 100 μM as indicated. After 8–10 h, 50 μl of culture were transferred to a 96-well plate containing 50 μl of phosphate buffered saline (PBS) solution. An Infinite m1000 Pro plate reader (Tecan) was used to measure optical density (OD) at 600 nm, and to measure RFP fluorescence (FL) (540 nm excitation and 600 nm emission). For time course experiments shown in Supplementary Figure S5, cells were diluted 1:30 in LB and 100 μl were added to a 96-well plate, which was transferred to the plate reader with constant shaking, and OD and RFP measurements were taken every 10 min for 10 h.

Bulk fluorescence data analysis

For each experiment there were two sets of controls: a media blank and *E. coli* MG1655 cells transformed with a

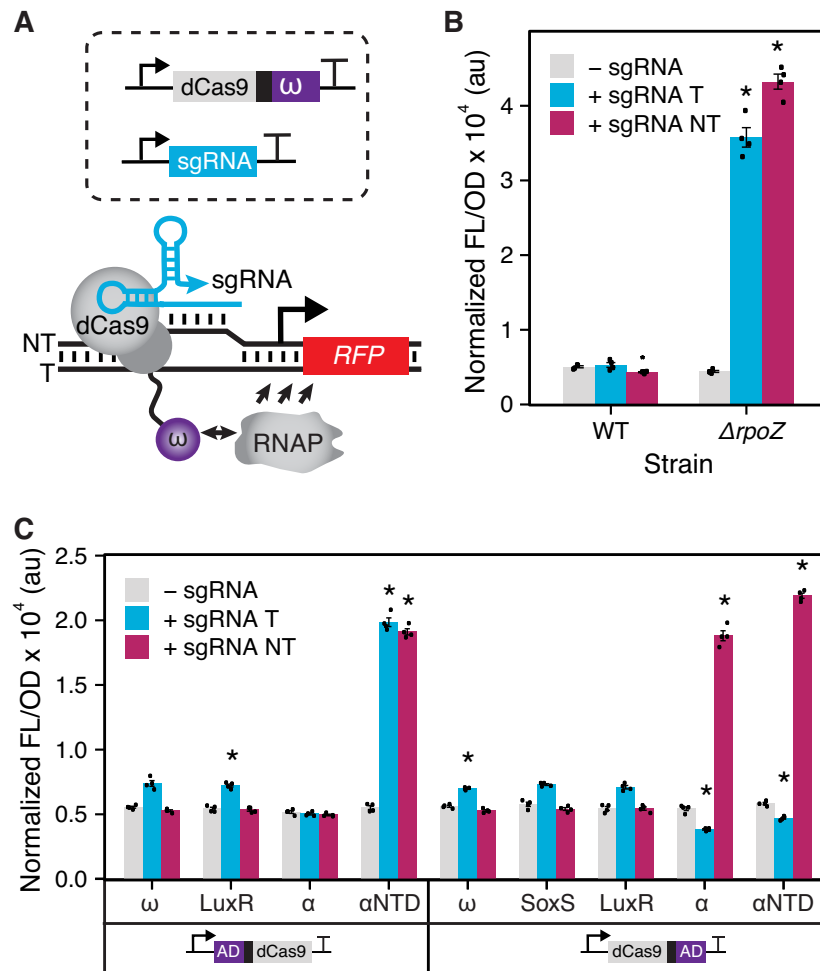


Figure 1. Synthetic transcription activation with dCas9 fusions to ADs. (A) Schematic of a CRISPRa system composed of dCas9 fused to the ω subunit of RNAP (dCas9- ω). For characterization, dCas9- ω in complex with the sgRNA was targeted upstream of a promoter driving RFP expression. Activation is achieved by local recruitment of RNAP to promoter elements. (B) Fluorescence characterization of a dCas9- ω system in wild-type strain *E. coli* K-12 BW25113 (WT) and a modified strain that lacks the gene encoding the ω subunit ($\Delta\rho Z$). (C) Fluorescence characterization of dCas9 with N- or C-terminal fusions to different ADs through a two-alanine linker in *E. coli* MG1655 strain. ADs are derived from RNAP subunits: ω , α and α NTD; or transcription factors: SoxS and LuxR. Fluorescence measurements (measured in units of fluorescence [FL]/optical density [OD] at 600 nm) were performed with *E. coli* cells transformed with an RFP reporter plasmid, a plasmid encoding the dCas9 fusions, and a sgRNA-encoding plasmid or a no-sgRNA control plasmid. sgRNA variants used targeted PAMs located at 80 bp upstream of the promoter TSS on the template strand (+ sgRNA T) or 81 bp upstream on the non-template strand (+ sgRNA NT). Data represent mean values and error bars represent s.d. of $n = 4$ biological replicates. A two tailed Student's *t* test was used to calculate *P* value comparing against the no-sgRNA control. * $P < 0.0001$; $P > 0.0001$ has no asterisk.

combination of the control plasmids pJEC101, pJEC102, pJEC103 and pJEC598 (blank cells), which only contained the antibiotic resistance cassette and thus not expressing the reporter gene (Supplementary Figure S1). Each experiment contained four biological replicates of each control. OD and FL values for each colony were first corrected by subtracting the corresponding mean values of the media blank. The ratio of FL to OD (FL/OD) was then calculated for each well (grown from a single colony) and Normalized FL/OD values were obtained subtracting FL/OD of blank cells from the four colonies characterized. Means of FL/OD were calculated over replicates and error bars represent standard deviations (s.d.). To calculate fold activation, a no-sgRNA control was used in which cells were transformed with a control plasmid (pJEC102). Mean of FL/OD of the no-sgRNA control was calculated and used as a measurement of the basal expression state. Fold acti-

vation was calculated from this by dividing the FL/OD of the targeting sgRNA (experimental condition) by the no-sgRNA control. Error propagation was performed to obtain the s.d. values.

Measuring distance-dependent effects

To evaluate distance-dependent effects, a library of seven sgRNAs (Supplementary Table S7) was programmed to target protospacer adjacent motifs (PAMs) located at positions 70, 80, 90 on the template strand and 61, 71, 81, 91 on the non-template strand upstream of the TSS of the RFP reporter plasmid (pJEC581). These same sgRNAs were used to target 9 additional reporter plasmids derived from pJEC581 (Reporter +1 to Reporter +9) that each contained additional nucleotides (1–9 nt) to extend the distance between the promoter TSS and each PAM (Supple-

mentary Figure S6). As the basal RFP expression from each reporter varied slightly, the values are reported as fold activation, normalized against the no-sgRNA control for each reporter.

Construction of phylogenetic tree

An expanded multifurcated phylogenetic tree, based on the NCBI taxonomy database, was generated using phyloT (<https://phylot.biobyte.de/>) to display the bacterial species from which α NTD sequences were obtained for data shown in Figure 2.

RESULTS

Identifying activation domains (ADs) for a bacterial CRISPRa system using dCas9 recruitment

As a starting point, our goal was to establish a basic CRISPRa system in *E. coli*. One of the most established routes to do this, is to fuse onto the N- or C-terminus of dCas9 the omega (ω) subunit of RNAP (28), which has served as a versatile AD in a variety of regulatory contexts (35,44) (Figure 1A). This is due to the ability of the ω subunit to structurally stabilize and recruit the β' subunit of RNAP (45), functionalities that can activate transcription when localized close to promoter elements. To characterize this basic design, plasmids were constructed in which the ω subunit was fused to the C-terminus of the dCas9 using a two-alanine linker. To measure transcription activation, we constructed a reporter plasmid containing a constitutively expressed RFP gene with a J23117 promoter and multiple NGG protospacer adjacent motifs (PAMs) upstream of the promoter element, which was derived from a previously described plasmid (31,32). We designed corresponding sgRNA expression plasmids to target two PAMs located 80 bp upstream of the TSS in the template strand (80T) and 81 bp upstream of the TSS in the non-template strand (81NT). A plasmid containing only the antibiotic resistance cassette was constructed to be used as the no-sgRNA control. *E. coli* cells were transformed with the dCas9- ω plasmid, the RFP reporter plasmid, and either a sgRNA-encoding plasmid or the no-sgRNA control. Experiments were performed in two strains: *E. coli* K-12 BW25113 or an engineered *E. coli* K-12 BW25113 $\Delta rpoZ$ strain (43), in which the endogenous *rpoZ* gene encoding the ω subunit had been knocked out, which had previously shown to allow for greater activation from ω -based ADs (28,46). RFP fluorescence (540 nm excitation and 600 nm emission) and optical density at 600 nm were measured for each culture. From these experiments, we observed no significant activation when this CRISPRa system was targeted to either the template or non-template strand in the parental strain *E. coli* K-12 BW25113. However, in the knockout strain, targeting either DNA strand resulted in \sim 8-fold activation (Figure 1B). These results confirmed that while the ω subunit of RNAP can serve as an AD, a significant limitation is its dependence on the use of a knockout strain. This is problematic because although the ω subunit is considered non-essential for many bacterial functions, its absence is known to reduce growth rate (45,47), inhibit antibiotic production (48), and inhibit morphological

differentiation (49), potentially limiting the utility of these CRISPRa systems.

To optimize the activation of this CRISPRa system and remove the dependence on knockout strains, we next performed screening to identify alternative ADs. To do this, we identified a panel of ADs derived from the subunits of *E. coli* RNAP and bacterial transcription factors (Supplementary Table S3). Specifically, we evaluated the alpha subunit (α) and the α subunit N-terminal domain (α NTD). Our interest in the α subunit was due to previous demonstrations using it as an AD in two-hybrid screens (50), a function most likely explained by the role of the α subunit in initiating RNAP assembly (51). Additionally we used transcription factors from the AraC and LuxR family of regulators known to act as transcription activators through interactions with RNAP (52,53). For each AD, fusions to the N- and C-terminus of dCas9 were constructed, with the exception of SoxS for which only the C-terminal fusion could be obtained. These different fusions were then targeted to the RFP reporter plasmid at positions 80T and 81NT, and fluorescence was measured (Figure 1C). From these experiments, we observed that the α NTD fused to the N-terminus of the dCas9 was the strongest variant when targeted to both the non-template and template strand, achieving up to 5-fold activation. For experiments using LuxR as an AD, we also performed measurements in the presence of its inducer, *N*-acyl homoserine lactone (AHL) (Supplementary Figure S2), but no change in fluorescence was observed. In addition to this panel, we tested a recently described strong AD based on an evolved AsiA, an anti-sigma factor protein derived from T4 bacteriophage (33), targeting several upstream positions (Supplementary Figure S3). However, we did not observe significant activation, likely due to differences in our experimental system, including the use of different reporters and different sgRNA sequences. Taken together, these results demonstrate that the α NTD is an attractive AD for construction of a bacterial CRISPRa system.

Exploring modular protein-protein interaction domains for AD recruitment

We next decided to investigate alternative protein fusion strategies to recruit the AD by dCas9. In particular, we were interested in utilizing noncovalent protein-protein interactions that would allow for the dCas9 and AD to be encoded on separate and easily interchangeable plasmids and thus allow for a high degree of modularity within our CRISPRa system. We envisaged this would be particularly valuable to eliminate the laborious cloning of the large dCas9-encoding plasmid and to allow for rapid screening of different synergistic combinations of the dCas9 and AD components. To this end, we decided to explore the use of synthetic coiled-coil SYNZIP interaction domains (39,40), which have provided a versatile protein interaction domain for molecular engineering applications. To test this, we constructed two new plasmids, one containing dCas9 fused to SYNZIP18 and the other containing the AD fused to SYNZIP17 (Supplementary Table S4), which form highly selective antiparallel heterodimers (Figure 2B, Supplementary Figure S4) (40,54). In order to benchmark the effectiveness of this ap-

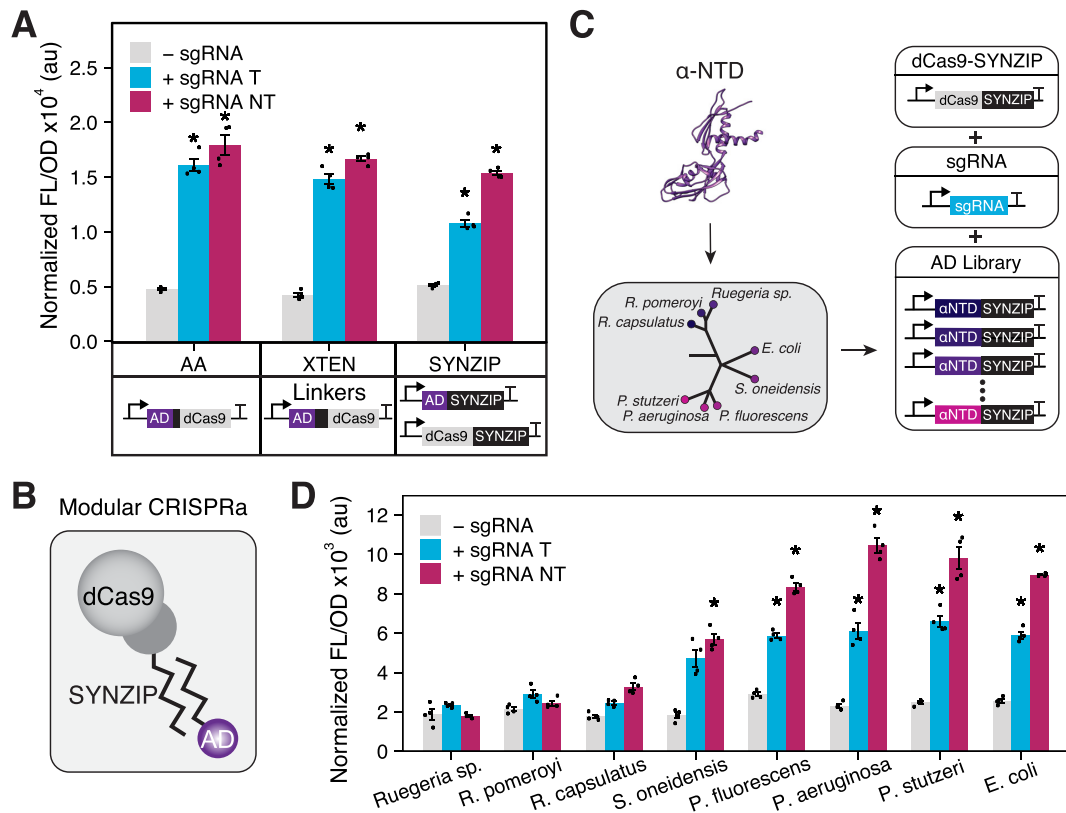


Figure 2. Evaluation of different ADs and fusion strategies. **(A)** Characterization of different protein fusion strategies to recruit the α NTD AD through dCas9 by a covalent two-alanine (AA) linker, covalent 16 amino acid XTEN linker, or a noncovalent SYNZIP interaction domain. Schematics of DNA constructs shown under each data set. **(B)** Schematic of the modular CRISPRa design. The dCas9 and AD are independently fused to SYNZIP domains that form heterodimers. **(C)** A schematic illustrating creation and screening of an AD library using our modular CRISPRa platform. Different ADs were identified from α NTDs derived from 8 different bacterial species and translationally fused to a SYNZIP domain. These AD-encoding plasmids were then co-transformed with dCas9-SYNZIP and sgRNA-encoding plasmids to create new CRISPRa systems. **(D)** Characterization of the AD library. Fluorescence characterization (measured in units of fluorescence [FL]/optical density [OD] at 600 nm) was performed with MG1655 *E. coli* cells transformed with an RFP reporter plasmid, a dCas9-AD plasmid or separate dCas9- and AD-encoding plasmids, and a sgRNA-encoding plasmid or a no-sgRNA control plasmid. sgRNA variants used targeted PAMs located at 80 bp upstream of the promoter TSS on the template strand (+ sgRNA T) and 81 bp upstream on the non-template strand (+ sgRNA NT). Data represent mean values and error bars represent s.d. of $n = 4$ biological replicates. A two tailed Student's t test was used to calculate p value comparing against the no-sgRNA control. * $P < 0.0001$; $P > 0.0001$ has no asterisk.

proach, we compared it against two designs that used covalent linkers between the AD and the dCas9. These designs used linkers composed of either two alanines or a 16 amino acid sequence known as XTEN (SGSETPGTSESATPES), which is an artificial extended and unstructured linker that has seen broad utility for creating chimeric fusions with Cas proteins (55,56). Plasmids encoding these variants using the α NTD as the AD were constructed and characterized by targeting the RFP reporter plasmid at positions 80T and 81NT. This characterization revealed that in comparison to covalent linkers, the noncovalent linker maintained a similar level of activation fold, while offering increased modularity (Figure 2A). To demonstrate the inducibility of the modular CRISPRa system, we expressed SYNZIP-dCas9 under the control of a P_{tet} and P_{bad} promoters and the corresponding inducer was added. We observed that with P_{tet} , probably due to the leakiness of the promoter, more than 2-fold activation was obtained in the absence of the inducer and no further activation was achieved in the presence of the inducer. While P_{bad} promoter evidenced no activation in the absence of the inducer and >2 -fold activation when the inducer was added. (Supplementary Figure

S5A). We evaluated alternative designs of dCas9 fused to SYNZIP18, including fusions to the C- and N-terminus (dCas9-SYNZIP and SYNZIP-dCas9), and fusion of two tandem SYNZIP domains (dCas9-2x[SYNZIP]) (Supplementary Figure S5B). Similar levels of fold activation were achieved across these different design variants. In addition, we showed that growth (Supplementary Figure S5C) and FL levels (Supplementary Figure S5D) were not affected when the modular CRISPRa system was expressed without a sgRNA compared to cells expressing only the reporter.

To demonstrate the flexibility that our modular SYNZIP fusion approach allows, we screened a panel of seven additional ADs fused to SYNZIP17. As a proof of principle, we decided to screen a panel of α NTD derived from different species across the bacteria domain of life, to begin to understand the design rules governing this class of AD (Supplementary Table S6). Specifically, we obtained α NTD sequences from the genome of *Alphaproteobacteria*: *Ruegeria* sp. TM 1040 (*Ruegeria* sp.), *Ruegeria pomeroyi* and *Rhodobacter capsulatus*, and *Gammaproteobacteria*: *Shewanella oneidensis*, *Pseudomonas fluorescens*, *Pseudomonas aeruginosa* and *Pseudomonas stutzeri* (Figure 2C). Charac-

terizing these ADs in *E. coli* cells, we observed activation when targeting both the template and non-template strands with all of the α NTDs derived from *Gammaproteobacteria*, ranging from 2- to 5-fold activation (Figure 2D). Not surprisingly, the highest activation was obtained from *Pseudomonas* species, which are phylogenetically closer to *E. coli* and whose α NTD protein sequences share 69% sequence identity. In comparison, no activation was observed with the α NTD from *Ruegeria* species, which are more distantly related and share 44% protein sequence identity with *E. coli*. The successful activation using different α NTDs, despite different degrees of activation, suggests that α NTDs could be engineered to be used as an AD for CRISPRa systems across bacterial species.

Distance-dependent activation patterns of our CRISPRa platform

It has previously been shown that bacterial CRISPRa systems that utilize sgRNA-based recruitment of ADs (Supplementary Figure S4B) have periodical activation patterns in which activation is only observed when the system is targeted to a 2–4 bp window that repeats every ~ 10 –11 bp (32). Importantly, this means that for a given gene, only a fraction of target binding sites are functional, which significantly limits gene targeting range. For example, a previous analysis suggested only $\sim 10\%$ of endogenous promoters in *E. coli* contained PAMs at the optimal targetable positions (32). To understand if these effects were observed for our modular CRISPRa design, we next characterized the distance-dependent activation pattern. To do this, we adapted a previously reported approach in which a panel of sgRNAs and base-shifted reporter plasmids are used to measure activation patterns at nucleotide resolution (Supplementary Figure 6, and Materials and Methods) (32). From these experiments, we observed periodic activation windows of 2–4 bp, which repeated every ~ 10 bp (Figure 3). In order to understand if this was an inherent property of the SYNZIP design, or more generally, CRISPRa systems that recruit via dCas9, we also performed these experiments on a CRISPRa system that used the covalent XTEN linker (Figure 3). Interestingly, we observed that regardless of length or even type of linker used (i.e. covalent versus noncovalent linkers) activation patterns are conserved, with the highest activation being obtained in the same positions. These results showed for the first time that in addition to CRISPRa systems that utilize sgRNA-based recruitment of ADs, dCas9-based recruitment has similar distance-dependent periodic activation patterns. While the exact cause for this is unknown, we hypothesized that this was a result of the spatial localization of the AD by dCas9 relative to the bound DNA (32), which may only result in successful positioning of the AD to the nearby promoter when targeted on specific surfaces of the DNA double helix, which rotates every ~ 10.5 bp.

Expanding activation patterns using circularly permuted dCas9 variants

In order to address this limitation, we decided to explore alternative dCas9 variants that could uniquely position ADs

relative to the bound DNA, which we reasoned would create shifted activation patterns. To do this, we decided to use circularly permuted variants of dCas9 (cpdCas9) in which the original N- and C-termini were fused with a 20 amino acid linker and the order of the amino acid sequence changed to create new N- and C-termini (57). Our goal was to identify a library of different cpdCas9 variants that could be easily interchanged with the AD expression plasmids to create new CRISPRa systems capable of activating from distinct positions (Figure 4A). To initially test this idea, we performed a preliminary evaluation with a dCas9 variant circularly permuted at the 1029th residue (cpdCas9¹⁰²⁹). A plasmid encoding an N-terminal SYNZIP fusion was created and combined with a sgRNA plasmid library to target this CRISPRa system to different PAMs located upstream of the RFP reporter gene on the non-template strand (Supplementary Figure S6 and S7). Comparison of the activation pattern of this CRISPRa system, compared to those obtained with wild-type dCas9, showed a ~ 2 bp shifted activation pattern, resulting in activation from positions non-targetable by the wild-type dCas9 system (Figure 4B). Following this initial success, we decided to perform a more extensive screen to identify other potentially useful cpdCas9 variants. In particular, we wanted to identify variants able to activate from positions that appeared to be non-targetable by wild-type dCas9 or cpdCas9¹⁰²⁹. To do this, we created N- and C-terminal fusions of the SYNZIP18 to 11 previously identified cpdCas9 variants (Supplementary Figure S8) and characterized their ability to activate from four PAMs located 65, 66, 67 and 68 bp upstream of the TSS in the non-template strand (Supplementary Figure S9). From this, we identified cpdCas9¹⁹⁹ as a potential candidate, which was then characterized for its distance-dependent activation pattern (Figure 4B). From this, we confirmed that like cpdCas9¹⁰²⁹, a CRISPRa system that utilized cpdCas9¹⁹⁹ resulted in a shifted activation pattern. Interestingly, structural analysis of the new N-termini, to which the SYNZIP was fused in the dCas9, and their angles relative to the axis of the DNA double helix (Figure 4C) showed remarkable similarities to the relative angles of binding sites on the DNA double helix from which activation was observed (Figure 4D). This suggests that the shifted activation pattern could be explained as a result of distinct localization of ADs relative to the DNA double helix.

DISCUSSION

In our work, we apply rational protein engineering to successfully generate and characterize a novel bacterial CRISPRa system presenting a high degree of modularity. Our system encodes the dCas9 and AD as plasmid-independent elements that can be easily exchanged, allowing each to be independently optimized and synergistically combined to create new CRISPRa systems with distinct properties. Through exchange of different ADs, we demonstrate that we can alter regulatory properties, such as fold activation, while exchanging dCas9 variants allows for distinct targeting range. As such, we believe the work described here enhances the current versatility of bacterial CRISPRa

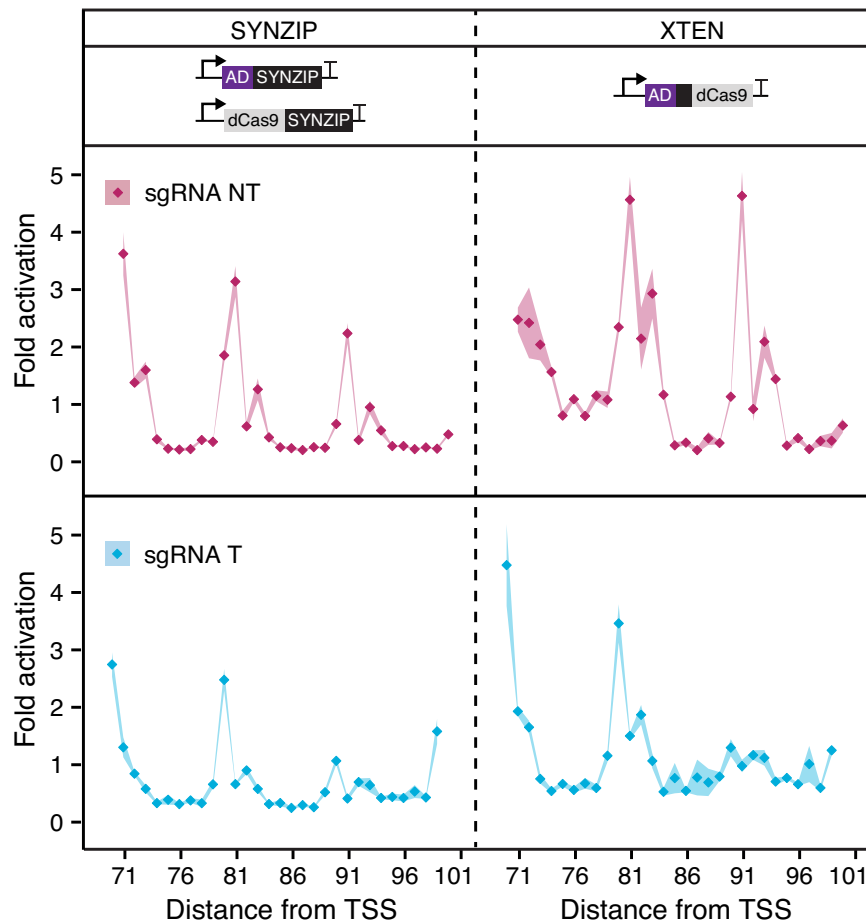


Figure 3. Characterization of distance-dependent activation patterns. Measuring the distance-dependent activation patterns of CRISPRa systems using XTEN and SYNZIP linkers. These systems were targeted to PAMs located between 71 and 101 bp upstream of the promoter TSS on both the template and non-template strands. Fluorescence characterization (measured in units of fluorescence [FL]/optical density [OD] at 600 nm) was performed in MG1655 *E. coli* cells transformed with an RFP reporter plasmid, a dCas9 plasmid, an AD plasmid and a sgRNA-encoding plasmid or a no-sgRNA control plasmid. Fold activation was calculated by dividing the [FL]/[OD] obtained in the presence of a targeting sgRNA against the no-sgRNA control within each reporter plasmid. Data represent mean values and shading represent s.d. of $n = 4$ biological replicates.

systems by creating a highly modular system with expanded target range.

Our screening of different ADs identified the N-terminal domain of the α subunit of RNAP (α NTD) to be a robust AD, able to successfully activate transcription when connected to dCas9 through either covalent or noncovalent linker domains. The α subunit of RNAP is naturally composed of two independently folded domains: α NTD and α CTD. In the process of RNAP assembly, α NTD dimerizes with a separate α NTD to initiate assembly, while α CTD is not required for this process, but instead interacts with transcription factors to recruit the RNAP to specific promoter sequences. Interestingly, we observed that the α NTD appeared to be optimal when fused through its C-terminus, in effect, mimicking its natural configuration in which the α CTD interacts with transcription factors to recruit RNAP. In this case, α CTD was replaced by dCas9 which performs the role of recruitment to a target promoter sequence. This activation by recruitment is analogous to regulatory mechanisms found in natural bacterial transcription activators, which rely on the recruitment of RNAP by protein-

protein interactions. Interestingly, upon screening a library of α NTDs derived from different bacterial species, we observed a surprising level of portability. This is likely due to the high degree of sequence conservation across different species (58,59) and confirms previous observations that α subunits from different species can be functionally interchangeable (60,61). Thus, α NTD potentially provides an attractive AD as researchers attempt to transfer CRISPRa systems into non-model bacterial species. Given that α NTD activates via recruitment of RNAP, there is also potential to synergistically combine it with other classes of ADs that activate transcription through distinct mechanisms, for example, rearrangement of the promoter DNA structure (62).

The presence of distance-dependent periodic activation patterns has been previously reported for bacterial CRISPRa systems that utilize AD recruitment through interactions with the sgRNA (32). Here, we show that the periodic activation pattern is also present for our CRISPRa system that utilizes dCas9-based recruitment of ADs, and that this activation pattern is independent of the length or type (i.e. covalent or noncovalent) of linker connecting

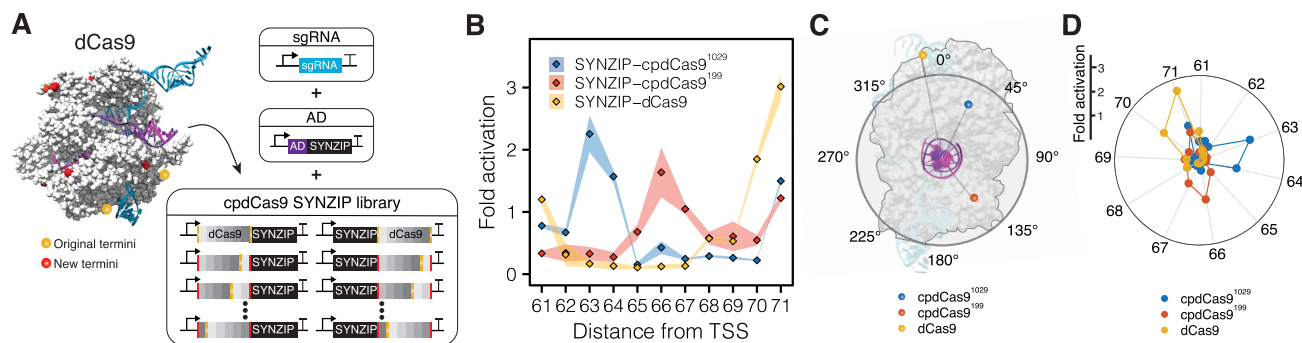


Figure 4. Circularly permuted dCas9 (cpdCas9) variants allow for an expanded target range. (A) cpdCas9 variants were fused to a SYNZIP through the new N- or C-termini to generate a cpdCas9 SYNZIP library that was evaluated as part of the modular CRISPRa system. cpdCas9 variants used in this study are shown on the crystal structure of Cas9 (PDB: 5F9R), DNA is colored purple, sgRNA is colored blue and dCas9 is colored grey. (B) Characterization of the distance-dependent activation patterns of CRISPRa systems using three dCas9 variants: wild-type dCas9 (SYNZIP-dCas9), cpdCas9¹⁰²⁹ (SYNZIP-cpdCas9¹⁰²⁹) and cpdCas9¹⁹⁹ (SYNZIP-cpdCas9¹⁹⁹). Distance-dependent activation patterns were measured by targeting PAMs located between 61 and 71 bp upstream of the promoter TSS on the non-template strand (61NT to 71NT). Fluorescence characterization (measured in units of fluorescence [FL]/optical density [OD] at 600 nm) was performed in MG1655 *E. coli* cells transformed with an RFP reporter plasmid, a plasmid encoding the cpdCas9 or dCas9 variant, a plasmid encoding α NTD-SYNZIP, and a sgRNA-encoding plasmid or a no-sgRNA control plasmid. Fold activation was calculated by dividing the [FL]/[OD] obtained in the presence of a targeting sgRNA against the no-sgRNA control within each reporter plasmid. Data represent mean values and shading represent s.d. of $n = 4$ biological replicates. (C) Structural analysis of the angles corresponding to the different positions of the N-terminal SYNZIP fusions in the dCas9, cpdCas9¹⁰²⁹, and cpdCas9¹⁹⁹ relative to the axis of the DNA double helix (purple). (D) Fold activation values obtained from the different dCas9 variants targeting positions 61NT to 71NT shown as a function of the targeted position in a DNA helical wheel that was set to 10.5 bp per turn.

the AD to the dCas9. Interestingly, the activation pattern we obtained overlaps with the pattern previously reported, which also reported an inability to alter activation patterns through linker engineering (32). Based upon our analysis, it appears that this pattern is a result of spatial localization of the AD by dCas9, which appears to only result in activation when positioned on certain sides of the DNA double helix. This would also explain why the same activation patterns are observed when ADs are recruited through the sgRNA (32) or the N- and C-termini of dCas9, which are all located in close proximity within the tertiary structure of the dCas9-sgRNA complex (63). Exploiting this, we show that circularly permuted variants of dCas9 (cpdCas9) that have been engineered to contain new N- and C-termini at distinct positions within the tertiary structure, can be used to shift activation patterns and allow activation from distinct binding sites compared to wild-type dCas9. This significantly expands the number of binding sites from which CRISPRa systems can achieve activation, which we anticipate can be increased further in the future through more extensive screening of cpdCas9 variants. While in this study we demonstrated expanded targeting range using a synthetic σ^{70} promoter, we anticipate that this approach will allow activation of the large number of endogenous promoters containing PAM binding sites that are non-optimal for other CRISPRa systems (31–34). This approach is also complementary to other efforts to expand the target range of CRISPRa systems using PAM-relaxed dCas9 variants that can utilize alternative PAM sequences (32,64–66). We thus expect that combining cpdCas9 and PAM-relaxed dCas9 variants would provide even greater targeting range, as recently demonstrated for CRISPR–Cas base editor systems (67). Finally, harnessing CRISPR–Cas systems from different types and origins (68–70) could yield further systems with distinct regulatory properties.

This work, along with other recent efforts, have begun to address limitations of bacterial CRISPRa systems; however, challenges remain. First, in general bacterial CRISPRa systems have shown lower fold activation compared to their eukaryotic counterparts and other synthetic RNA-based activating systems (71–73), although recent efforts have identified promising ADs that can allow for high-fold activation (31–33). While low-fold activation is not necessarily a limitation for all applications, we anticipate that it will limit certain applications, such as the construction of complex synthetic gene circuits that depend on high dynamic ranges. Second, a deeper understanding of the rules governing activation patterns is required, particularly for endogenous promoter elements. Such design rules are critical to allow for reliable and even automated design of CRISPRa variants that are required to apply these technologies for large-scale genome-wide perturbation studies. Finally, more work is required to identify optimal ADs and understand the design rules of CRISPRa within non-model bacterial species. Given the remarkable progress made in recent years, we anticipate that these challenges and knowledge gaps will be addressed.

In summary, this work expands the current utility of bacterial CRISPRa systems by enhancing modularity and increasing targeting range. This work creates a novel platform for gene activation, which we anticipate will be of value for applications such as metabolic engineering, and to advance fundamental understanding of gene function and connectivity.

DATA AVAILABILITY

All source data for main and SI figures was deposited in Rice University's Rice Digital Scholarship Archive (<https://doi.org/10.25611/DG85-7Y09>).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

M.C.V.K. and J.C. designed the study, M.C.V.K. and A.J.T. performed experiments. All authors contributed to data analysis and preparation of the manuscript.

FUNDING

Welch Foundation [C-1982–20190330 to J.C.]; Alfred P. Sloan Research Fellowship [FG-2018-10500 to J.C.]; J.C. is an Alfred P. Sloan Research Fellow. Funding for open access charge: discretionary funds.

Conflict of interest statement. None declared.

REFERENCES

- Adli, M. (2018) The CRISPR tool kit for genome editing and beyond. *Nat. Commun.*, **9**, 1911.
- Ameruso, A., Gambill, L., Liu, B., Villegas Kcam, M.C. and Chappell, J. (2019) Brave new 'RNA' world—advances in RNA tools and their application for understanding and engineering biological systems. *Curr. Opin. Syst. Biol.*, **14**, 32–40.
- Xu, X. and Qi, L.S. (2019) A CRISPR–dCas toolbox for genetic engineering and synthetic biology. *J. Mol. Biol.*, **431**, 34–47.
- Kampmann, M. (2018) CRISPRi and CRISPRa screens in mammalian cells for precision biology and medicine. *ACS Chem. Biol.*, **13**, 406–416.
- Jaitin, D.A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., Salame, T.M., Tanay, A., van Oudenaarden, A. and Amit, I. (2016) Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-Seq. *Cell*, **167**, 1883–1896.
- Adamson, B., Norman, T.M., Jost, M., Cho, M.Y., Nuñez, J.K., Chen, Y., Villalta, J.E., Gilbert, L.A., Horlbeck, M.A., Hein, M.Y. *et al.* (2016) A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell*, **167**, 1867–1882.
- Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G. *et al.* (2014) Genome-scale CRISPR–Cas9 knockout screening in human cells. *Science*, **343**, 84–87.
- Shalem, O., Sanjana, N.E. and Zhang, F. (2015) High-throughput functional genomics using CRISPR–Cas9. *Nat. Rev. Genet.*, **16**, 299–311.
- Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z., Zhang, C. and Xing, X.-H. (2018) Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun.*, **9**, 2475.
- Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H.S., Koo, B.-M., Marta, E. *et al.* (2016) A Comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*, **165**, 1493–1506.
- Boettcher, M., Tian, R., Blau, J.A., Markegard, E., Wagner, R.T., Wu, D., Mo, X., Biton, A., Zaitlen, N., Fu, H. *et al.* (2018) Dual gene activation and knockout screen reveals directional dependencies in genetic networks. *Nat. Biotechnol.*, **36**, 170–178.
- Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C.P., Jerby-Arnon, L., Marjanovic, N.D., Dionne, D., Burks, T., Raychowdhury, R. *et al.* (2016) Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell*, **167**, 1853–1866.
- Li, S., Jendresen, C.B., Landberg, J., Pedersen, L.E., Sonnenschein, N., Jensen, S.I. and Nielsen, A.T. (2020) Genome-wide CRISPRi-based identification of targets for decoupling growth from production. *ACS Synth. Biol.*, **9**, 1030–1040.
- Yao, L., Cengic, I., Anfelt, J. and Hudson, E.P. (2016) Multiple gene repression in cyanobacteria using CRISPRi. *ACS Synth. Biol.*, **5**, 207–212.
- Huang, C.-H., Shen, C.R., Li, H., Sung, L.-Y., Wu, M.-Y. and Hu, Y.-C. (2016) CRISPR interference (CRISPRi) for gene regulation and succinate production in cyanobacterium *S. elongatus* PCC 7942. *Microb. Cell Factories*, **15**, 196.
- Wang, M., Liu, L., Fan, L. and Tan, T. (2017) CRISPRi based system for enhancing 1-butanol production in engineered *Klebsiella pneumoniae*. *Process Biochem.*, **56**, 139–146.
- Wu, J., Zhou, P., Zhang, X. and Dong, M. (2017) Efficient de novo synthesis of resveratrol by metabolically engineered *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.*, **44**, 1083–1095.
- Ly, L., Ren, Y.-L., Chen, J.-C., Wu, Q. and Chen, G.-Q. (2015) Application of CRISPRi for prokaryotic metabolic engineering involving multiple genes, a case study: Controllable P(3HB-co-4HB) biosynthesis. *Metab. Eng.*, **29**, 160–168.
- Kim, S.K., Han, G.H., Seong, W., Kim, H., Kim, S.-W., Lee, D.-H. and Lee, S.-G. (2016) CRISPR interference-guided balancing of a biosynthetic mevalonate pathway increases terpenoid production. *Metab. Eng.*, **38**, 228–240.
- Fontana, J., Voje, W.E., Zalatan, J.G. and Carothers, J.M. (2018) Prospects for engineering dynamic CRISPR–Cas transcriptional circuits to improve bioproduction. *J. Ind. Microbiol. Biotechnol.*, **45**, 481–490.
- Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P. and Lu, T.K. (2014) Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell*, **54**, 698–710.
- Kim, H., Bojar, D. and Fussenegger, M. (2019) A CRISPR/Cas9-based central processing unit to program complex logic computation in human cells. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 7214–7219.
- Nielsen, A.A. and Voigt, C.A. (2014) Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol.*, **10**, 763.
- Kiani, S., Beal, J., Ebrahimkhani, M.R., Huh, J., Hall, R.N., Xie, Z., Li, Y. and Weiss, R. (2014) CRISPR transcriptional repression devices and layered circuits in mammalian cells. *Nat. Methods*, **11**, 723–726.
- Gander, M.W., Vrana, J.D., Voje, W.E., Carothers, J.M. and Klavins, E. (2017) Digital logic circuits in yeast with CRISPR–dCas9 NOR gates. *Nat. Commun.*, **8**, 15459.
- Lee, Y.J., Hoynes-O'Connor, A., Leong, M.C. and Moon, T.S. (2016) Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. *Nucleic Acids Res.*, **44**, 2462–2473.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P. and Lim, W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, **152**, 1173–1183.
- Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F. and Marraffini, L.A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR–Cas system. *Nucleic Acids Res.*, **41**, 7429–7437.
- Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., Iyer, E.P.R., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J. *et al.* (2015) Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods*, **12**, 326–328.
- Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S. *et al.* (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*, **160**, 339–350.
- Dong, C., Fontana, J., Patel, A., Carothers, J.M. and Zalatan, J.G. (2018) Synthetic CRISPR–Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun.*, **9**, 2489.
- Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M. and Zalatan, J.G. (2020) Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.*, **11**, 1618.
- Ho, H., Fang, J.R., Cheung, J. and Wang, H.H. (2020) Programmable CRISPR–Cas transcriptional activation in bacteria. *Mol. Syst. Biol.*, **16**, e9427.
- Liu, Y., Wan, X. and Wang, B. (2019) Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun.*, **10**, 3693.
- Lu, Z., Yang, S., Yuan, X., Shi, Y., Ouyang, L., Jiang, S., Yi, L. and Zhang, G. (2019) CRISPR-assisted multi-dimensional regulation for

- fine-tuning gene expression in *Bacillus subtilis*. *Nucleic Acids Res.*, **47**, e40.
36. Peng, R., Wang, Y., Feng, W.-W., Yue, X.-J., Chen, J.-H., Hu, X.-Z., Li, Z.-F., Sheng, D.-H., Zhang, Y.-M. and Li, Y.-Z. (2018) CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in *Myxococcus xanthus*. *Microb. Cell Factories*, **17**, 15.
 37. Chavez, A., Tuttle, M., Pruiitt, B.W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S.J., Cecchi, R.J., Kowal, E.J.K., Buchthal, J. *et al.* (2016) Comparison of Cas9 activators in multiple species. *Nat. Methods*, **13**, 563–567.
 38. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H. *et al.* (2015) Genome-scale transcriptional activation by an engineered CRISPR–Cas9 complex. *Nature*, **517**, 583–588.
 39. Reinke, A.W., Grant, R.A. and Keating, A.E. (2010) A synthetic coiled-coil interactome provides heterospecific modules for molecular engineering. *J. Am. Chem. Soc.*, **132**, 6025–6031.
 40. Thompson, K.E., Bashor, C.J., Lim, W.A. and Keating, A.E. (2012) SYNZIP protein interaction toolbox: *in vitro* and *in vivo* specifications of heterospecific coiled-coil interaction domains. *ACS Synth. Biol.*, **1**, 118–129.
 41. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A. and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–345.
 42. Engler, C., Kandzia, R. and Marillonnet, S. (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, **3**, e3647.
 43. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006.0008.
 44. Dove, S.L. and Hochschild, A. (1998) Conversion of the omega subunit of *Escherichia coli* RNA polymerase into a transcriptional activator or an activation target. *Genes Dev.*, **12**, 745–754.
 45. Mathew, R. and Chatterji, D. (2006) The evolving story of the omega subunit of bacterial RNA polymerase. *Trends Microbiol.*, **14**, 450–455.
 46. Wu, F., Shim, J., Gong, T. and Tan, C. (2020) Orthogonal tuning of gene expression noise using CRISPR–Cas. *Nucleic Acids Res.*, **48**, e76.
 47. Kurlela, J., Fredman, J., Salminen, T.A. and Tyystjärvi, T. (2020) Revealing secrets of the enigmatic omega subunit of bacterial RNA polymerase. *Mol. Microbiol.*, **115**, 1–11.
 48. Santos-Beneit, F., Barriuso-Iglesias, M., Fernández-Martínez, L.T., Martínez-Castro, M., Sola-Landa, A., Rodríguez-García, A. and Martín, J.F. (2011) The RNA polymerase omega factor RpoZ is regulated by PhoP and has an important role in antibiotic biosynthesis and morphological differentiation in *Streptomyces coelicolor*. *Appl. Environ. Microbiol.*, **77**, 7586–7594.
 49. Kojima, I., Kasuga, K., Kobayashi, M., Fukasawa, A., Mizuno, S., Arisawa, A. and Akagawa, H. (2002) The rpoZ gene, encoding the RNA polymerase omega subunit, is required for antibiotic production and morphological differentiation in *Streptomyces kasugaensis*. *J. Bacteriol.*, **184**, 8.
 50. Dove, S.L., Joung, J.K. and Hochschild, A. (1997) Activation of prokaryotic transcription through arbitrary protein–protein contacts. *Nature*, **386**, 627–630.
 51. Murakami, K., Kimura, M., Owens, J.T., Meares, C.F. and Ishihama, A. (1997) The two α subunits of *Escherichia coli* RNA polymerase are asymmetrically arranged and contact different halves of the DNA upstream element. *Proc. Natl. Acad. Sci.*, **94**, 1709–1714.
 52. Martin, R.G., Gillette, W.K., Martin, N.I. and Rosner, J.L. (2002) 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.
 53. Finney, A.H., Blick, R.J., Murakami, K., Ishihama, A. and Stevens, A.M. (2002) Role of the C-terminal domain of the alpha subunit of RNA polymerase in LuxR-dependent transcriptional activation of the lux operon during quorum sensing. *J. Bacteriol.*, **184**, 4520–4528.
 54. Fulk, E.M., Huh, D., Atkinson, J.T., Lie, M., Masiello, C.A. and Silberg, J.J. (2020) A split methyl halide transferase AND gate that reports by synthesizing an indicator gas. *ACS Synth. Biol.*, **9**, 3104–3113.
 55. Schellenberger, V., Wang, C., Geething, N.C., Spink, B.J., Campbell, A., To, W., Scholle, M.D., Yin, Y., Yao, Y., Bogin, O. *et al.* (2009) A recombinant polypeptide extends the *in vivo* half-life of peptides and proteins in a tunable manner. *Nat. Biotechnol.*, **27**, 1186–1190.
 56. Guilinger, J.P., Thompson, D.B. and Liu, D.R. (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.*, **32**, 577–582.
 57. Oakes, B.L., Fellmann, C., Rishi, H., Taylor, K.L., Ren, S.M., Nadler, D.C., Yokoo, R., Arkin, A.P., Doudna, J.A. and Savage, D.F. (2019) CRISPR–Cas9 circular permuted scaffolds as programmable scaffolds for genome modification. *Cell*, **176**, 254–267.
 58. Ebright, R.H. and Busby, S. (1995) The *Escherichia coli* RNA polymerase alpha subunit: structure and function. *Curr. Opin. Genet. Dev.*, **5**, 197–203.
 59. Murakami, K.S. (2015) Structural biology of bacterial RNA polymerase. *Biomolecules*, **5**, 848–864.
 60. Cho, E.-J., Bae, J.-B., Kang, J.-G. and Roe, J.-H. (1996) Molecular analysis of RNA polymerase alpha subunit gene from *Streptomyces coelicolor* A3(2). *Nucleic Acids Res.*, **24**, 4565–4571.
 61. Kim, J.H., Ham, S.H. and Lee, B.R. (2012) Characterization of the RNA polymerase alpha subunit operon from *Corynebacterium ammoniagenes*. *World J. Microbiol. Biotechnol.*, **28**, 669–676.
 62. Lee, D.J., Minchin, S.D. and Busby, S.J.W. (2012) Activating transcription in bacteria. *Annu. Rev. Microbiol.*, **66**, 125–152.
 63. Jiang, F., Taylor, D.W., Chen, J.S., Kornfeld, J.E., Zhou, K., Thompson, A.J., Nogales, E. and Doudna, J.A. (2016) Structures of a CRISPR–Cas9 R-loop complex primed for DNA cleavage. *Science*, **351**, 867–871.
 64. Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z. *et al.* (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, **556**, 57–63.
 65. Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H. *et al.* (2018) Engineered CRISPR–Cas9 nuclease with expanded targeting space. *Science*, **361**, 1259–1262.
 66. Legut, M., Daniloski, Z., Xue, X., McKenzie, D., Guo, X., Wessels, H.-H. and Sanjana, N.E. (2020) High-throughput screens of PAM-flexible Cas9 variants for gene knockout and transcriptional modulation. *Cell Rep.*, **30**, 2859–2868.
 67. Huang, T.P., Zhao, K.T., Miller, S.M., Gaudelli, N.M., Oakes, B.L., Fellmann, C., Savage, D.F. and Liu, D.R. (2019) Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat. Biotechnol.*, **37**, 626–631.
 68. Gooma, A.A., Klumpe, H.E., Luo, M.L., Selle, K., Barrangou, R. and Beisel, C.L. (2014) Programmable removal of bacterial strains by use of genome-targeting CRISPR–Cas systems. *mBio*, **5**, e00928–13.
 69. Luo, M.L., Mullis, A.S., Leenay, R.T. and Beisel, C.L. (2015) Repurposing endogenous type I CRISPR–Cas systems for programmable gene repression. *Nucleic Acids Res.*, **43**, 674–681.
 70. Leenay, R.T., Maksimchuk, K.R., Slotkowski, R.A., Agrawal, R.N., Gooma, A.A., Briner, A.E., Barrangou, R. and Beisel, C.L. (2016) Identifying and visualizing functional PAM diversity across CRISPR–Cas systems. *Mol. Cell*, **62**, 137–147.
 71. Chappell, J., Takahashi, M.K. and Lucks, J.B. (2015) Creating small transcription activating RNAs. *Nat. Chem. Biol.*, **11**, 214–220.
 72. Chappell, J., Westbrook, A., Verosloff, M. and Lucks, J.B. (2017) Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nat. Commun.*, **8**, 1051.
 73. Green, A.A., Silver, P.A., Collins, J.J. and Yin, P. (2014) Toehold switches: de-novo-designed regulators of gene expression. *Cell*, **159**, 925–939.