Engineered RNA-Interacting CRISPR Guide RNAs for Genetic Sensing and Diagnostics

Roberto Galizi, John N Duncan, William Rostain, Charlotte M Quinn, Marko Storch, Manish Kushwaha, Alfonso Jaramillo

The CRISPR Journal
Volume 3, Number 5, 2020

DOI: 10.1089/crispr.2020.0029

The CRISPR Journal

RESEARCH ARTICLE

Engineered RNA-Interacting CRISPR Guide RNAs for Genetic Sensing and Diagnostics

Roberto Galizi,[1,2,†] John N. Duncan,[3,†] William Rostain,[3] Charlotte M Quinn,[2,4] Marko Storch,[2,5] Manish Kushwaha,[3,6] and Alfonso Jaramillo[3,7,*,‡]

Abstract

CRISPR guide RNAs (gRNAs) can be programmed with relative ease to allow the genetic editing of nearly any DNA or RNA sequence. Here, we propose novel molecular architectures to achieve RNA-dependent modulation of CRISPR activity in response to specific RNA molecules. We designed and tested, in both living *Escherichia coli* cells and cell-free assays for rapid prototyping, *cis*-repressed RNA-interacting guide RNA (igRNA) that switch to their active state only upon interaction with small RNA fragments or long RNA transcripts, including pathogen-derived mRNAs of medical relevance such as the human immunodeficiency virus infectivity factor. The proposed CRISPR-igRNAs are fully customizable and easily adaptable to the majority if not all the available CRISPR-Cas variants to modulate a variety of genetic functions in response to specific cellular conditions, providing orthogonal activation and increased specificity. We thereby foresee a large scope of application for therapeutic, diagnostic, and biotech applications in both prokaryotic and eukaryotic systems.

Introduction

The CRISPR loci and the associated Cas proteins constitute a type of prokaryotic adaptive immune system.[1] According to the type, CRISPR loci include a variable number of gene coding sequences, identified as CRISPR-associated (Cas) genes, linked to a repeat-spacer array by the leader sequence. The array contains short repeats separated by unique spacers acquired from invading viral RNAs or plasmids after the first infection (adaptation). These are transcribed and processed into small RNAs (sRNAs) that associate with one or more Cas proteins to form the active ribonucleoprotein (RNP) complex (expression). The RNA spacers are used to target specific nucleic acid sequences of mobile genetic elements upon subsequent invasions, activating the protein part of the RNP complex to cleave or degrade the invading DNA or RNA sequences (interference). In the most commonly exploited type II CRISPR system found in *Streptococcus pyogenes*, a single Cas9 protein acts as RNA-guided endonuclease by forming a complex with the CRISPR RNA (crRNA) and the trans-activating crRNA (tracrRNA) fragments to bind and cleave the double-stranded DNA target.[2] A single engineered RNA can also be used, by fusing the crRNA and the tracrRNA into one short synthetic guide RNA (identified as sgRNA or just gRNA). The gRNA is formed of three main domains: (1) the crRNA section, including the ~20 nt targeting region (or spacer) and the 22 nt segment base pairing with the tracrRNA complementary region; (2) the tracrRNA region complementary to the crRNA; and (3) the 3′ handle region, which allows termination of transcription of the gRNA fragment as well as stabilization of the Cas9-gRNA RNP. After RNP formation, the targeting region of the crRNA (or sgRNA) remains single stranded, with the 12 nt of the targeting region proximal to the Cas9 handle (identified as ''seed'' sequence) maintained in an A-form conformation to allow its hybridization with the complementary DNA sequence targeted.[3] The Cas9 handle is composed of four

[1]*Centre for Applied Entomology and Parasitology, School of Life Sciences, Keele University, Keele, United Kingdom;* [2]*Department of Life Sciences, Imperial College London, London, United Kingdom;* [3]*School of Life Sciences, University of Warwick, Coventry, United Kingdom;* [4]*Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, United Kingdom;* [5]*London Biofoundry, Imperial College London, London, United Kingdom;* [6]*Université Paris-Saclay, INRAE, AgroParisTech, MIcalis Institute, Paris, France;* [7]*Warwick Integrative Synthetic Biology Centre (WISB), University of Warwick, Coventry, United Kingdom.*
[†]These authors contributed equally to this work and are considered to be co-first authors
[‡]On leave of absence from CNRS.

*Address correspondence to: *Alfonso Jaramillo, PhD, School of Life Sciences, Gibbet Hill Campus, The University of Warwick, Coventry CV4 7AL, United Kingdom,* Email: Alfonso.Jaramillo@synth-bio.org

stem loops, the first being formed by the interaction between crRNA and tracrRNA (linked through a short loop in the sgRNA) followed by three smaller stems, the nexus, and the two 3′ terminal hairpins (Fig. 1A). These secondary structures have proved crucial for the formation of an active Cas9-gRNA RNP, which can be affected if the targeting region contains significant secondary structures[4] or modifications are added to its 5′ end.[5]

Both the protein and the RNA components of CRISPR-Cas systems have been extensively re-engineered. Specific mutations of the respective nuclease sites generate nuclease-deactivated proteins, such as the most popular dCas9 and dCpf1 (dCas12a), which can be used as RNA-guided DNA-binding repressors for genetic inactivation, also referred as CRISPR interference (CRISPRi). A multitude of other Cas protein variants and fusions

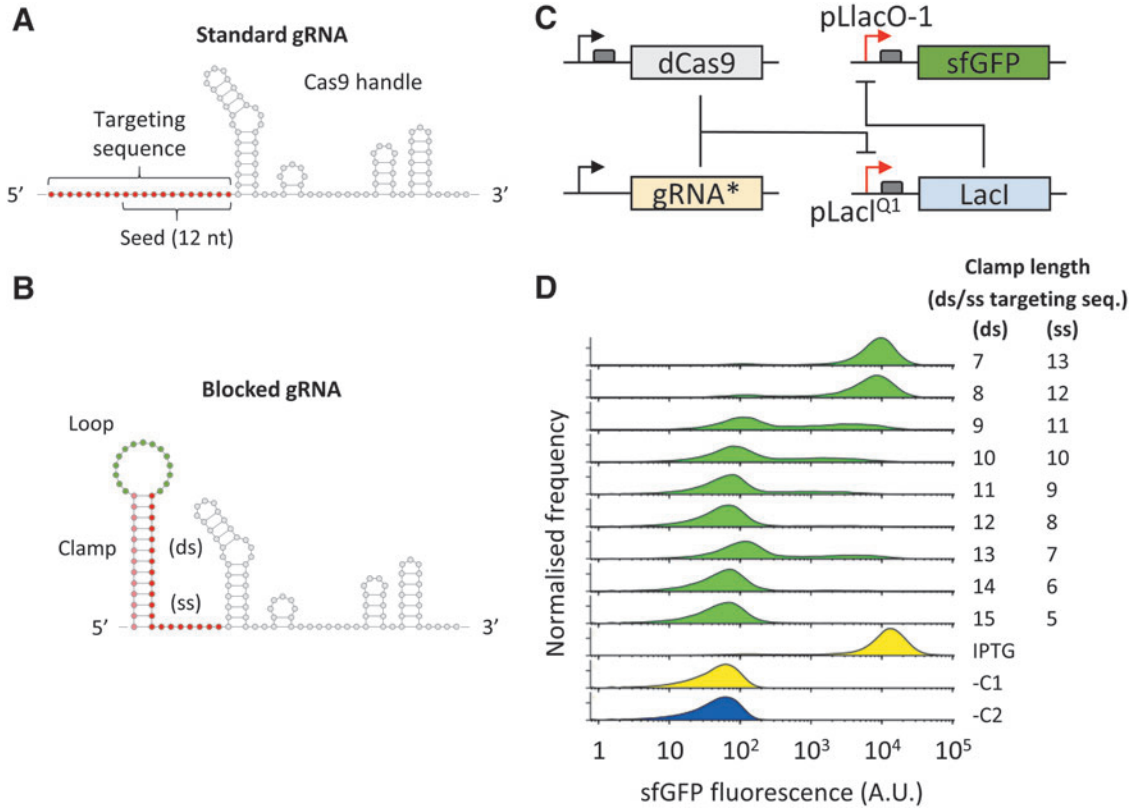

FIG. 1. **(A)** Schematic representation of the unmodified pLacI^{Q1}-targetting sgRNA with some of the major features highlighted: the 20 nt targeting sequence (red) including the 12 nt seed sequence, the Cas9 handle and the *Streptococcus pyogenes* terminator sequences (light gray). **(B)** Schematic representation of the *cis*-repressed gRNA via the 5′ addition of a 15 nt loop region (green) and a 7–15 nt clamp sequence (light red) hybridizing with the targeting sequence (red). Only the 13 nt clamp used in the following experiment is shown for simplicity. **(C)** Schematic representation of the *Escherichia coli* reporter system used for the testing of clamp length required to block CRIPSR-gRNA activity. The unmodified sgRNA complexes with dCas9 targeting the pLacI^{Q1} promoter to block the expression of the chromosomally inserted LacI repressor thereby resulting in expression of the LacI^{Q1}-driven sfGFP reporter. Inactivation of gRNA binding upon the addition of blocking clamp sequences (indicated as gRNA*) results in repression of the sfGFP signal. The ribosomal binding site (RBS) for protein translation is also indicated (dark gray boxes). **(D)** Flow cytometry fluorescence measurements of *E. coli* MG1655Z1 cells, containing plasmids expressing sfGFP reporter, dCas9, and the modified gRNAs with different clamp lengths. The reporter induced with IPTG is used as a positive control ("IPTG"). Cells containing only the sfGFP and LacI plasmids ("-C1") or only the Cas9 plasmid ("-C2") are used as negative controls to measure basal fluorescence levels. The values on the right indicate the length of the clamp sequence hybridizing with the gRNA targeting sequence (ds) and the length of the targeting sequence remaining single stranded (ss).

have been generated and used for programmable repression or activation of translation as well as for genetic detection.[6–9] CRISPR-gRNAs have also been engineered by adding riboswitches[10] or aptazymes[11] to create small molecule-responsive gRNAs, or fused to MS2 or PP7 binding regions, which can be targeted by MS2/PP7 protein fusions.[12–15] They can also be incorporated into computational RNA circuits regulated by antisense RNA inhibitors[16] or used for nucleic acid detection.[17–19]

RNA engineering is greatly facilitated by the possibility of exploiting secondary structure prediction software[20] to create RNA switches that are exclusively activated if in complex with another RNA. This has notably been used for the computational design of riboregulators.[21,22] We have previously shown the use of RNA circuitry to engineer logic operations both *in vitro* and *in vivo*, either with fully artificial parts or interfacing with natural RNAs.[23] On the other hand, the CRISPR system combines the flexibility of the RNA component, for programmed targeting of virtually any genomic sequence, with various genetic functions allowed by the multitude of CRISPR effectors currently available. This combination offers a virtually infinite number of options for genetic editing and RNA circuit designs.

The rational design of structured sequences in the 5′ region of mRNAs, able to change conformation upon binding of a ligand, has allowed the regulation of gene expression to be engineered entirely within the 5′UTR, without requiring additional protein co-factors. Ligands based on sRNA molecules have been used to design riboregulators and anti-terminators to modulate gene expression.[19,21,24–27] We previously proposed a new molecular architecture to engineer RNA-interacting guide RNAs (igRNAs) that change their conformation to the active state exclusively upon hybridization with single-stranded trigger RNAs (trRNAs).[19] Compared to other similar approaches (Supplementary Table S1),[28–33] in this work, we aimed to generate flexible igRNA architectures that can be designed with relative ease and virtually no targeting or sensing constraints to activate CRISPR editing upon sequence-specific interaction with synthetic or natural RNA molecules, including full-length and actively transcribed and translating mRNAs in living cells. The absence of interaction between the sensing segment (complementary to the trRNA) and the DNA-targeting sequence of the igRNA allows complete allosteric regulation of CRISPR activity. This means that any trRNA could be used to activate an igRNA targeting a specific genomic sequence and, vice versa, the same trRNA could be used to activate multiple igRNAs targeting different sequences.

We tested our novel RNA-responsive igRNA designs, in both genetic activation and repression settings, upon sensing of sRNA fragments or entire mRNA transcripts with high structural complexity such as those encoding for the fluorescent protein mKate2 and the human immunodeficiency virus (HIV) viral infectivity factor (VIF). We also show proof of principle of adaptability and scalability of igRNA design and testing using fully automatable *in vitro* cell-free methods for rapid prototyping.

Methods

Plasmid design

For the *Escherichia coli* assay, the pLlacO-1 promoter[34] was combined with sfGFP[35] within a pSC101-E93K plasmid[36] containing a kanamycin resistance marker. This was further modified using a BioBrick[37] pJ23119-igRNA-HDV ribosome[38] gBlock to create one of the parental plasmids used in this study. The second parental plasmid (Addgene, 46569),[39] containing the dCas9 coding sequence (p15A, chloramphenicol resistance marker). All plasmids used in this study were generated from these two plasmids. For the cell-free transcription-translation (TXTL) assay, the P70a-deGFP plasmid[40] was modified to generate the reporter constructs expressing the deGFP fluorescent protein with the addition of the same CRISPR target sequence used in the *E. coli* assay at the N-terminus of the deGFP coding sequence. The TX1 reporter carries the target sequence within the deGFP template strand, while the TX4 carries the target in the same position but reverted (non-template strand). The constructs expressing igRNAs, trRNA, or targeting and non-targeting gRNA controls (AJ1–AJ8) were constructed by Golden Gate cloning of each fragment into the receiving vector PAJ486 containing two copies of the J23119 promoter for the co-expression of igRNA and corresponding trRNA (AJ1 and AJ5), targeting gRNA (AJ2 and AJ6) as positive control, non-targeting gRNA (AJ4 and AJ8), and igRNA with the non-corresponding trRNA as negative controls. A detailed description of plasmids, primers, and methods used for their generation is provided in the Supplementary Material (Supplementary Methods and Supplementary Tables S2 and S3).

E. coli assay

E. coli MG1655Z1 cells[41] were grown overnight to early stationary phase at 37°C with shaking in M9 minimal media[21] with the addition of appropriate antibiotics (anhydrotetracycline at a concentration of 100 ng/mL). Cells were then diluted 1:200 in 2 mL of M9 minimal media with the addition of antibiotics and anhydrotetracycline and grown at 37°C with shaking for 2 h before being added to a black-wall clear-bottom 96-well plate. The plates were incubated in a Tecan F500 fluorescent plate reader at 37°C with shaking. A time series of optical

density (600 nm) and fluorescence (480/20 nm excitation and 530/25 nm emission for sfGFP, 580/20 nm excitation 635/35 nm emission for mKate2[42]) readings were taken at 15 min intervals. An average of measurements taken from three biological repeats, each with three technical repeats, was calculated and reported together with the relative standard deviations. All measurements were blanked against M9 media (containing antibiotics and supplements) and against MG1655Z1 cells containing empty plasmids with the appropriate antibiotic markers without igRNA/trRNA/dCas9/sfGFP components to acquire basal fluorescence levels.

Flow cytometry measurements were performed using a BD LSRFortessa (BD Biosciences). Single colonies were grown overnight in LB supplemented with 17.5 μg/mL kanamycin, 50 μg/mL spectinomycin, and 17.5 μg/mL chloramphenicol. The next day, cultures were diluted on a 96-well plate (5 μL in 195 μL), grown (4 h, 37°C, 200 × *g*), and then diluted for a second time following the same procedure. Once a cell density corresponding to $OD_{600} = 0.3$ had been reached, 10 μL of the culture was diluted in 190 μL phosphate-buffered saline (with 2 mM kanamycin to block further protein synthesis). Samples were kept on ice and measured the same day. Flow cytometer voltage settings were: FSC, 632; SSC, 237; B485-530/30, 509. For all the samples tested, FCS data were gated to remove cell debris, and fluorescence histograms were plotted with the Flow Cytometry GUI for Matlab.

Quantitative polymerase chain reaction analysis

Total RNA was extracted from 1.5 mL bacterial cells ($OD_{600nm} = 0.6$) using a standard Trizol (Thermo Fisher Scientific) protocol. Extracted RNA was precipitated with isopropanol, washed in 70% ethanol, and stored in DEPC-water. The RNA was then treated with DNase I for 1 h at 37°C, re-extracted with Trizol and resuspended in DEPC-water. RNA was tested for DNA contamination using the primers used for quantitative polymerase chain reaction (qPCR). No PCR products were found, indicating the absence of DNA contamination. cDNA (20 μL final volume) was produced from 1 μg of RNA (ReadyScript cDNA synthesis mix; Sigma–Aldrich) and diluted 1:10 with DEPC-water. Diluted cDNA (0.5 μL) was used to confirm cDNA integrity. A qPCR master mix was created according to the manufacturer's specifications (KiCqStart SYBR Green qPCR ReadyMix Low ROX; Sigma–Aldrich) by adding the premixed primer pairs and cDNA (or samples containing no cDNA, and non-retrotranscribed RNA as controls). The qPCR reactions were performed on a Stratagene MX3005p qPCR machine with the following protocol: 1 min at 94°C followed by 40 cycles (15 s at 94°C, 10 s at 57°C, and 20 s at 72°C), melt analysis, and cooling to 30°C. Results were normalized to the expression of the chloramphenicol resistance gene present in the plasmid.

Gel-shift assay

The 3A igRNA RNA fragment was *in vitro* transcribed (TranscriptAid T7 High-Yield; Thermo Fisher Scientific) from a PCR product containing the T7 promoter upstream of the 3A igRNA. The resulting RNA was extracted with Trizol, precipitated with isopropanol and ethanol, washed and re-suspended in DEPC-water. This RNA (1 μg) was mixed with 1 μg of synthesized 3A or 2A trRNA oligonucleotides (Integrated DNA Technologies), incubated at room temperature for 10 min, and run on a 12% polyacrylamide gel for 2.5 h. The RNA was visualized on a UV transilluminator after staining with ethidium bromide for 20 min.

Cell-free automated assay

Cell-free reactions were performed using 5 nM of each construct in a total volume of 4 μL per reaction. Each reaction contains a combination of three constructs: (1) TX1 or TX4 deGFP reporter; (2) Cas9 or dCas9 plasmid; and (3) plasmid carrying both igRNA (replaced with targeting or non-targeting gRNAs as positive or negative control) and trRNA (2A or 3A). After purification with magnetic beads (AMPure XP; Beckman Coulter), plasmids at standard concentrations were added to 3 μL myTXTL master mix (507024; Arbor Biosciences) and distilled water to a final volume of 4 μL. Each component was transferred from 384-Echo microplates (PPL-0200; Labcyte) into Greiner black small-volume low-base 384-well plates (788876; Greiner) using the acoustic liquid handler (Echo 525; Labcyte) as per the manufacturer's instructions. The Synergy Mx microplate reader (BioTek) was used to analyze deGFP fluorescence (485/20 nm excitation and 528/20 nm emission) every 3 min for 234 acquisitions (~12 h) at 29°C. A total of 12 replicates were performed for each combination of plasmids.

Results and discussion

5′ added cis-regulatory sequence dictates mechanism of CRISPR-gRNA regulation

With the aim of blocking gRNA annealing to the complementary DNA target, we initially introduced a stem-loop structure to the 5′ end of the 20 nt targeting region of a LacI promoter-targeting gRNA. The structure includes a predesigned 15 nt variable sequence loop[22] preceded by a 7–15 nt long ''clamp'' sequence complementary to the seed region of the gRNA (Fig. 1B). To test the repression of gRNA binding activity in *E. coli* cells, we used a reporter system where the gRNA targets the LacI[Q1]

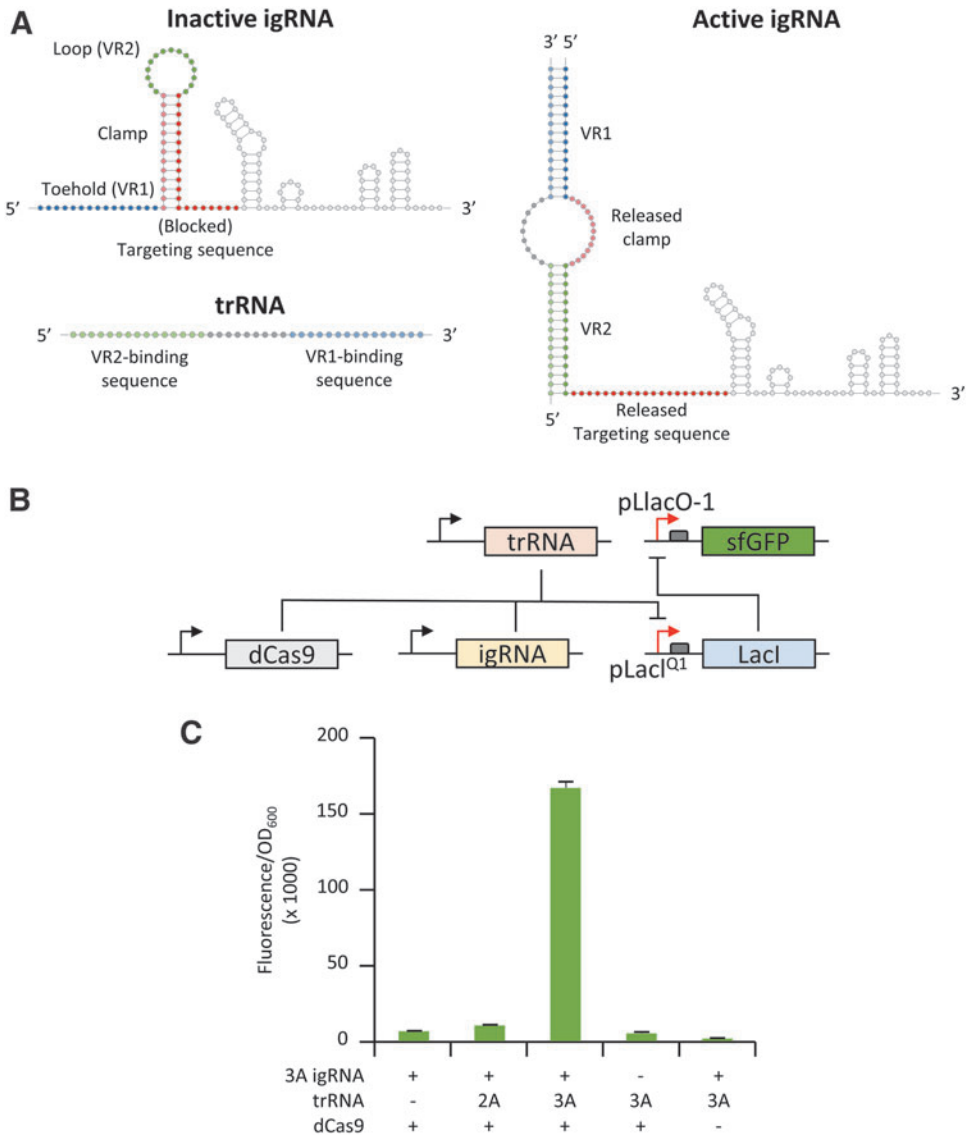

FIG. 2. **(A)** Schematic representation of igRNA in its inactive state, trigger RNA (trRNA) and activated igRNA after hybridization with the trRNA. The igRNA differs from the unmodified sgRNA (represented in Fig. 1A) for the addition of the toehold, clamp, and variable loop sequences. The toehold variable region 1 (VR1, blue) and variable loop region 2 (VR2, green) are complementary to the trRNA 3′ (light blue) and 5′ sequences (light green), respectively. The igRNA design envision no binding between the sRNA and the constrained clamp sequence (complementary to the targeting sequence). Therefore, the toehold, clamp, and targeting sequences of the igRNA are fully modifiable. Hybridization between the sRNA 5′ and 3′ ends with the igRNA toehold and loop sequences expose the targeting region and reactivates binding to its DNA target. **(B)** Schematic representation of the *E. coli* genetic activation reporter. The reporter represented in Figure 1C was modified by adding the trRNA component that regulates CRISPR-igRNA guided derepression of sfGFP. **(C)** Average ratio ± standard deviation (*SD*) of fluorescence intensity to OD_{600} measured from *E. coli* MG1655Z1 cells containing the plasmids carrying the LacI-repressed sfGFP reporter in all samples, the igRNA carrying the 3A switch (lane 1, 2, 3, and 5), dCas9 (1, 2, 3, and 4), and the 3A trRNA (3, 4, and 5). The 3A trRNA was replaced with the 2A (line 2) to test the specificity of igRNA switching in lane 2. Error bars indicate the *SD* value from all the biological and technical replicates measured.

promoter controlling the expression of the LacI repressor. This was coupled with a sfGFP fluorescent reporter under the control of the IPTG-inducible pLlacO-1 promoter and the dCas9 protein under the endogenous *S. pyogenes* promoter. In the resulting sensor circuit, the fluorescence signal, normally repressed, is reactivated as a result of CRISPR activity (Fig. 1C).

Upon assay in *E. coli* cells, we found that the addition of a 7–8 nt clamp, predicted to leave the 12 nt seed region of the gRNA fully exposed, resulted in strong sfGFP signal comparable to the positive control (cells induced with IPTG), indicating successful inhibition of the LacI repressor by the gRNA–dCas9 complex. These results indicate that the 5′ end of the gRNA can accommodate additional nucleotides and secondary structures without affecting its functional activity. On the other hand, clamp lengths hybridizing with ≥9 nt of the targeting sequence resulted in a marked reduction of fluorescence with an increasing proportion of non-fluorescent cells up to levels comparable to the control measurements. These results indicate that a folded RNA secondary structure immediately upstream of the targeting region can generate an active or inactive gRNA, depending on the length of the hybridizing sequence (Fig. 1D).

CRISPR-igRNAs allow allosteric sRNA-responsive modulation of genetic activation or repression in *E. coli*

In order to transform the 5′ blocked gRNA into a robust allosteric RNA sensor, we added an additional ''toehold'' structure to the 5′ end, which includes: two variable regions (VR) complementary to the 5′ and 3′ end of the trRNA; the 15 nt toehold sequence (VR1); and the 15 nt loop (VR2), flanking a segment (clamp) that is complementary to 13 nt of the 20 nt targeting region. The resulting igRNA incorporates two discrete parts that are fully independent of each other: the first encompassing two variable regions complementary to the trRNA sequence, including toehold (V1) and loop (V2); and the second constituted by the clamp sequence blocking the 20 nt targeting region in the igRNA inactive state (Fig. 2A). We used the publicly available NUPACK webserver[43] to analyze 30–40 nt segments of the mKate2 sequence and to select two 39 nt long trRNA sequences (identified as 2A and 3A) based on free and co-folding energy with their cognate igRNA molecules as well as reduced self-complementarity at 37°C. Simultaneously, the corresponding igRNA sensors (including toehold, clamp, and loop region) were analyzed to verify minimal complementarity with the downstream secondary structure of the gRNA. The 3A igRNA was initially tested in *E. coli* using the same LacI-repressed sfGFP reporter used for the clamp testing, with the conditional addition of the trRNA to activate fluorescence expression as a result of CRISPR-igRNA-guided inhibition of the LacI repressor (Fig. 2B). The 3A igRNA showed minimal levels of activity in the absence of the trRNA comparable to the negative controls where the igRNA or the dCas9 constructs were not added. Interestingly, the addition of the corresponding 3A trRNA resulted in high levels of sfGFP fluorescence, corresponding to a 26-fold increase compared to the control experiment where the trRNA was not added. To test the specificity of igRNA activation in blocking the LacI repressor, we also tested the 3A igRNA in the presence of the non-complementary 2A trRNA, which showed only a small increase of fluorescence (1.6-fold) compared to the control without trigger (Fig. 2C). Specificity of hybridization between sensor and trigger sequences was also confirmed by *in vitro* gel-shift assay where the 3A igRNA showed reduced

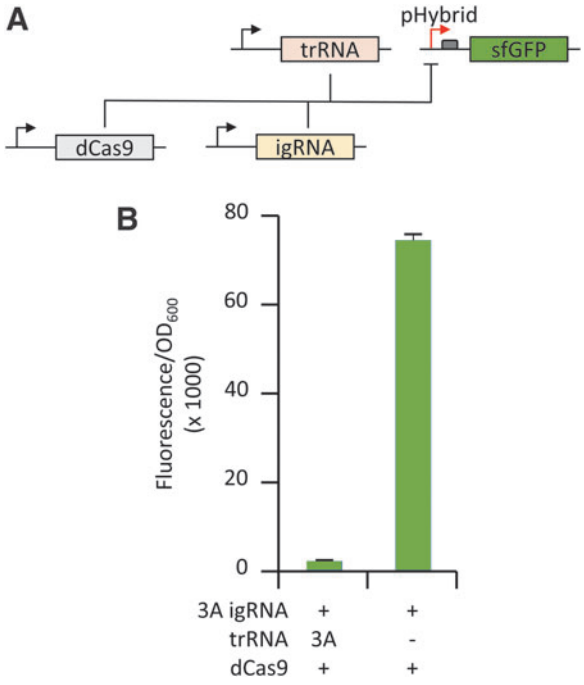

FIG. 3. **(A)** Schematic representation of the *E. coli*-based genetic repression reporter where the RNP complex formed by dCas9, igRNA, and trRNA binds and therefore represses the newly designed hybrid promoter used to drive expression of the sfGFP reporter (pHyb-sfGFP plasmid). **(B)** Average ratio ± *SD* of fluorescence intensity to OD_{600} measured from the *E. coli* MG1655Z1 cells carrying the pHyb-sfGFP reporter plasmid, the 3A igRNA, and a plasmid expressing both the dCas9 and 3A trRNA (lane 1) or only the dCas9 expression cassette (lane 2).

mobility only in the presence of the cognate 3A trRNA (Supplementary Fig. S1).

To further explore the adaptability of our igRNA designs, we converted the *E. coli*-based gene-activating reporter into a genetic repression system. A new hybrid promoter was constructed by fusing a copy of the 20 nt pLacI^{Q1} targeting region at either side of a computationally designed constitutive promoter devoid of known repressor binding sites (i.e., tetO and lacO) and predicted to promote high levels of expression.[44] This configuration allows the binding of the CRISPR RNP complex to either side of the promoter to improve transcriptional repression when the dCas9 is combined with an actively binding gRNA or trigger-activated igRNA.[44] The hybrid promoter was inserted upstream of a strong ribosomal binding site (RBS) to drive sfGFP expression, and the resulting reporter plasmid (pHyb-sfGFP) was used to replace the LacI repressor and LacI^{Q1}-driven sfGFP constructs (Fig. 3A). *E. coli* testing of CRISPR-igRNA activity in the new gene-repression architecture showed high levels of sfGFP expression from the hybrid promoter when no trRNA was added, while combining the previously tested 3A igRNA and cognate 3A trRNA generated a strong reduction (~35-fold) of sfGFP expression (Fig. 3B).

Altogether, these results show that CRISPR-igRNAs can be interchangeably utilized for RNA-dependent genetic repression or activation in live cellular environments.

CRISPR-igRNAs exhibit selective activation by full-length actively transcribed and translating mRNA

After the successful demonstration of sRNA-specific activation of igRNA, we tested whether full-length actively transcribed and translating mRNAs could function as triggers for CRISPR activation in *E. coli* cells. We initially used the mRNA of the red-fluorescent protein mKate2 to allow visual confirmation of parallel translation of the mRNA trigger via production of the red fluorescent signal. The secondary structure of the mKate2 mRNA transcript was interrogated by using the NUPACK tool to select single-stranded RNA regions. A 39 nt sequence and the cognate igRNA switch (3AV9) were designed following the same architecture used for the 3A igRNA. *E. coli* testing using the genetic activation reporter (described in Fig. 2B) showed a ~2.1-fold increase in sfGFP production upon addition of the mKate2 mRNA compared to the control where the mKate2-expressing construct was not added. In order to test whether interactions between the mRNA trigger and cellular translational machinery may interfere with the hybridization and consequent activation of the igRNA, we generated a second version of the mKate2 trigger construct by deleting the respective RBS. As expected, the addition of the translationally inactive mKate2 mRNA showed only basal levels of mKate2 fluorescence. However, sfGFP fluorescence was maintained

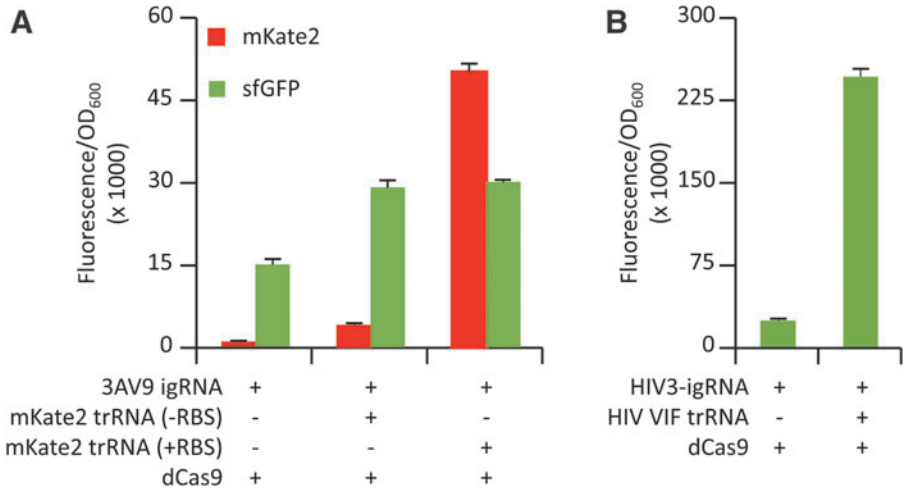

FIG. 4. **(A)** Average ratio ± *SD* of fluorescence intensity to OD_{600} measured from the *E. coli* genetic activation reporter (described in Figure 2B) carrying the LacI-repressed sfGFP reporter, the dCas9 plasmid, and the 3AV9-igRNA in the absence (lane 1) or presence (lane 3) of the mKate2 mRNA trigger. The latter condition was tested also with a variant of the trRNA-expressing construct where the mKate2 RBS was omitted to obliterate concomitant translation of the trRNA transcript (lane 2). **(B)** The same analysis was performed on *E. coli* cells carrying the HIV-responsive igRNA (HIV3-igRNA) without (lane 1) or with the addition (lane 2) of the actively translating HIV VIF mRNA trigger.

at similar levels compared to the translationally active trigger, indicating that active translation did not affect binding and activation of the cognate 3AV9 igRNA (Fig. 4A).

Following on from these results, we decided to test the ability of other medically relevant mRNA transcripts to activate our igRNA designs. For this purpose, we selected the highly conserved VIF mRNA, which is known to be essential for HIV infectivity by playing a critical role in the production of infectious HIV virions.[45,46] The computational design of the HIV VIF-responsive igRNA was performed following the same method used for the mKate2 3AV9 switch to select a 34 nt region of the corresponding

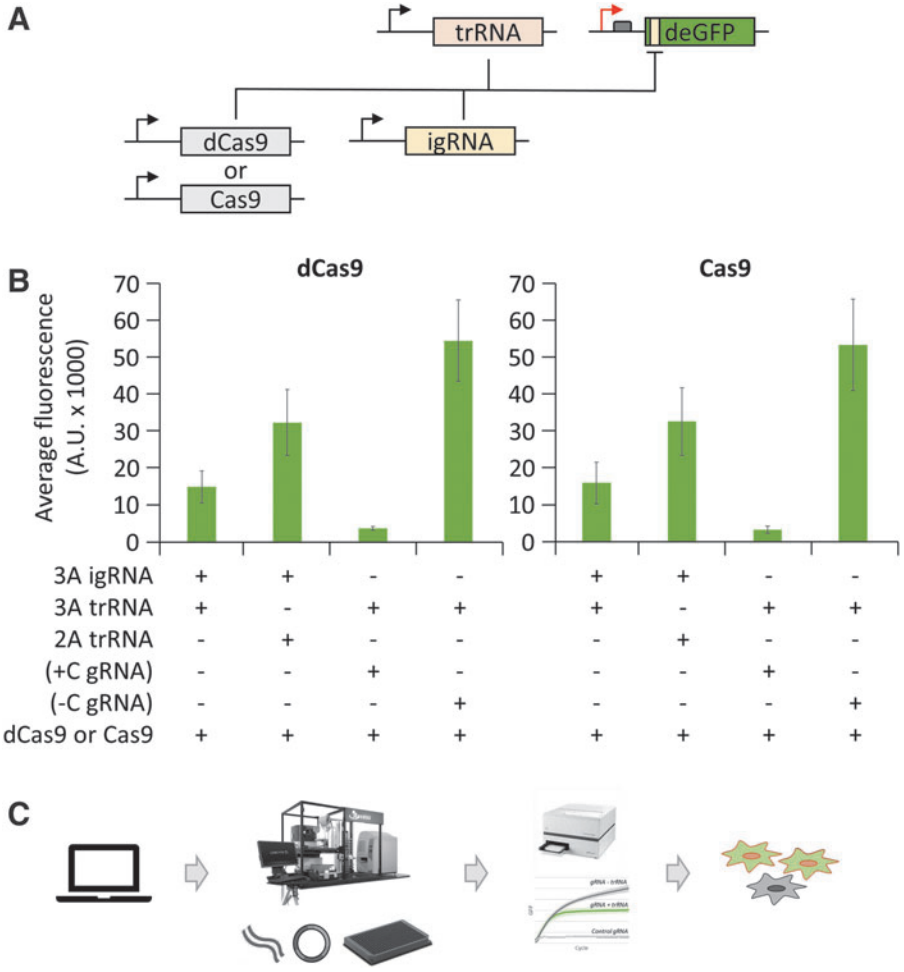

FIG. 5. **(A)** Schematic representation of the cell-free genetic repression reporter where the deGFP fluorescent marker carries the same target site used for the *E. coli* experiments inserted in either the template or non-template strand of the deGFP coding sequence. **(B)** Cell-free testing of 3A igRNAs with the nuclease-deactivated *S. pyogenes* dCas9 (left) and the active Cas9 (right) targeting the non-template reporter. Bar graphs show the average fluorescence intensity (as arbitrary unit ± *SD*) emitted by the fluorescent reporter construct. Repression activity of the 3A igRNA with the corresponding 3A trRNA (first lane) were compared to the 3A trRNA with the non-complementary 2A trRNA (second lane), a standard gRNA targeting the same sequence (+C gRNA) as a positive control (third lane), and a non-targeting gRNA (-C gRNA) as a negative control (fourth lane). A total of 12 replicates were produced for each reaction. The full set of data of 3A and 2A igRNAs targeting the template or non-template strand with dCas9 or Cas9 are summarized in Supplementary Figure S3. **(C)** Schematic representation of our platform for high-throughput testing of igRNA designs involving (left to right): (1) computational design, (2) automated cloning of DNA plasmids and assembly of cell-free reactions, (3) acquisition and analysis of CRISPR activity using fully customizable reporter constructs, and (4) selection of igRNA and trRNA designs for *in vivo* applications.

mRNA (named HIV3-igRNA). The subsequent testing for genetic repression in *E. coli* reporter showed a consistent increase in sfGFP production (~10-fold) in the presence of the HIV VIF mRNA compared to the negative control where the viral mRNA trigger was not added, indicating the successful activation of CRISPR-igRNA activity. The concomitant transcription of the HIV VIF messenger transcript in the tested samples was confirmed by qPCR analysis, supporting the ability of HIV3-igRNA to switch to its active form upon interaction with the transcribed viral mRNA (Supplementary Fig. S2).

Scalable *in vitro* screening of CRISPR-igRNA repression and cleavage activity

After showing that igRNAs can be used to direct dCas9 activity in *E. coli* cells upon the addition of either short sRNA or long mRNA triggers, we tested the activity of both 3A and 2A igRNAs with either the nuclease-deactivated dCas9 or the nuclease-active Cas9 through an automatable and easily scalable platform for *in vitro* assay. We generated constructs expressing the igRNA (or standard gRNA as control) and trRNA as well as a customizable reporter construct where the target site for CRISPR-directed cleavage (e.g., using Cas9) or repression (e.g., using dCas9) can be easily interchanged. For this assay, we generated two variants of the fluorescent reporter, including the same target site used in the *in vivo* testing positioned in both orientations, in order to test CRISPR-igRNA targeting of both template and non-template DNA strand (Fig. 5A). We used an acoustic liquid handler (Echo 525; Labcyte) to set up reactions rapidly in volumes as small as 4 μL on 384-well plates, combining Cas9, igRNA, and trRNA constructs with a commercially available *E. coli*-based cell-free transcription-translation master mix (myTXTL; Arbor Biosciences). A significant reduction of signal, compared to the negative control, was revealed when both 3A igRNA and 3A trRNA were added to the reaction with either dCas9 (3.7-fold reduction) or Cas9 (3.4-fold reduction) to target the non-template strand of the deGFP reporter. Similar results were also obtained with the 2A igRNA and cognate trigger, giving a 3.6-fold reduction with dCas9 and 4.2 when used with active Cas9. On the other hand, *in vitro* testing of both 3A and 2A igRNAs showed leaky activity when combined with the non-cognate triggers, ranging from a 1.6-fold reduction when the 3A igRNA was combined with the 2A trRNA to 2.8 when the 2A igRNA was combined with the 3A trRNA for non-template targeting with active Cas9. However, we are aware that the 2A and 3A igRNA tested here were designed based on folding prediction at 37°C, while the *in vitro* cell-free assay was performed at the recommended temperature of 29°C.[47] As expected, the standard gRNA (without the additional sensor sequence included to the 5′ to generate igRNAs), used as a positive control, gave a strong reduction of fluorescence signal when used to target either non-template or template strand with Cas9, as well as when targeting the non-template strand with dCas9. On the contrary, targeting of the template strand with dCas9 showed only moderate levels of repression, even when using the control gRNA (Fig. 5B and Supplementary Fig. S3). These results confirm previous studies showing that non-template strand binding of the dCas9–gRNA RNP complex is usually required to prevent inhibition or removal from the progressing RNA polymerase and efficiently block elongation of transcription via CRISPRi.[7,39] Moreover, the approach proposed here offers the opportunity to scale the throughput of igRNA and cognate trRNA screenings by using a fully automatable platform (Fig. 5C).

Conclusion

The work presented here demonstrates that CRISPR-gRNAs can be converted into *cis*-repressed igRNAs to transform standard CRISPR-Cas enzymes into freely programmable genetic editors that are precisely activated by specific RNA sequences. The added switch sequence blocks the igRNA ability to bind the targeted DNA sequence, maintaining the Cas9–igRNA complex functionally inactive in the absence of the trRNA. We used igRNAs to switch CRISPR activity, in both *in vitro* and living *E. coli* settings, to accomplish genetic repression or activation when in combination with either nuclease active or inactive Cas9 proteins. We show that the interaction between the *cis*-repressed riboswitch sequence and the cognate trigger can be computationally predicted to engineer igRNAs, showing up to 35-fold differential activation of reporter genes in living cells while maintaining very low baseline activity. The igRNA mode of action relies primarily on RNA folding and RNA–RNA interactions added to the gRNA-5′ end, thereby simplifying design steps and adaptability to the many CRISPR-Cas variants currently identified or engineered, significantly expanding the range of future applications. For this reason, we believe that our CRISPR-igRNA methodology is generalizable to any prokaryotic and eukaryotic host where the CRISPR system is known to work. Importantly, our igRNA architecture offers virtually complete decoupling in the design of the DNA-targeting sequence and the trigger-complementary riboswitch sensor of the igRNA. This distinctive feature opens to the possibility of targeting any desirable gene upon sequence-specific activation by any desirable RNA molecule, pending adequate prototyping and testing. Additional testing will be required to characterize trRNA/igRNA dose–response dynamics of

CRISPR-igRNA according to specific applications and corresponding parameters (e.g., temperature, salt concentration, and coexisting RNAs). Nonetheless, our results reveal that igRNAs maintain a high degree of sensitivity in response to small RNA fragments, as well as full-length and actively transcribed and translating mRNAs; for example, the one encoding for the viral infectivity factor of HIV, showing up to 10-fold activation in our *in vivo E. coli*–based reporter. In addition, we described an innovative platform, combining fully customizable and automatable *in vitro* reporter assays, for high-throughput screening and optimization of CRISPR-igRNA components supporting future development of our technology. We suggest that the methodology described here could be used for the *de novo* design and testing of various *in vitro* or cellular sensors for a multitude of applications, including the development of new diagnostic and therapeutic tools.

Acknowledgments

The authors would like to thank the Warwick WISB Centre for the use of the BD LSRFortessa cell analyzer; the Crisanti Lab and the London Biofoundry at Imperial College London for support with the cell-free assay and use of the Echo 525 acoustic liquid handler. We would also like to thank M.B. Elowitz for the MG1655Z1 strain.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This work was supported by the 7th Framework Programme (grant numbers 610730 [EVOPROG], 613745 [PROMYS]); the Horizon 2020 Marie Sklodowska-Curie (grant number 642738 [MetaRNA]); the Engineering and Physical Sciences Research Council (EPSRC) and the Biotechnology and Biological Sciences Research Council (BBSRC; grant numbers BB/P020615/1 [EVO-ENGINE] and BB/M017982/1 [WISB Centre]); and the School of Life Sciences (University of Warwick; start-up allocation) to A.J. Funding for open access charge: EPSRC/BBSRC (BB/M017982/1) to A.J. W.R. was supported by a DGA-Dstl fellowship.

Supplementary Material

Supplementary Methods
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3

References

1. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007;315:1709–1712. DOI: 10.1126/science.1138140.
2. Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 2012;337:816–821. DOI: 10.1126.science.1225829.
3. Jinek M, Jiang F, Taylor DW, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* 2014;343:1247997. DOI: 10.1126.science.1247997.
4. Thyme SB, Akhmetova L, Montague TG, et al. Internal guide RNA interactions interfere with Cas9-mediated cleavage. *Nat Commun* 2016;7:1–7. DOI: 10.1038/ncomms11750.
5. Mullally G, van Aelst K, Naqvi MM, et al. 5′ modifications to CRISPR-Cas9 gRNA can change the dynamics and size of R-loops and inhibit DNA cleavage. *Nucleic Acids Res* 2020;48:6811–6823. DOI: 10.1093/nar/gkaa477.
6. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 2013;154:442–451. DOI: 10.1016/j.cell.2013.06.044.
7. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 2013;152:1173–1183. DOI: 10.1016/j.cell.2013.02.022.
8. Dahlman JE, Abudayyeh OO, Joung J, et al. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat Biotechnol* 2015;33:1159–1161. DOI: 10.1038/nbt.3390.
9. Chen B, Gilbert LA, Cimini BA, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 2013;155:1479–1491. DOI: 10.1016/j.cell.2013.12.001.
10. Liu Y, Zhan Y, Chen Z, et al. Directing cellular information flow via CRISPR signal conductors. *Nat Methods* 2016;13:938–944. DOI: 10.1038/nmeth.3994.
11. Tang W, Hu JH, Liu DR. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat Commun* 2017;8:15939. DOI: 10.1038/ncomms15939.
12. Mali P, Aach J, Stranges PB, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 2013;31:833–838. DOI: 10.1038/nbt.2675.
13. Zalatan JG, Lee ME, Almeida R, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* 2015;160:339–350. DOI: 10.1016/j.cell.2014.11.052.
14. Konermann S, Brigham MD, Trevino AE, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 2015;517:583–588. DOI: 10.1038/nature14136.
15. Shechner DM, Hacisuleyman E, Younger ST, et al. Multiplexable, locus-specific targeting of long RNAs with CRISPR-Display. *Nat Methods* 2015;12:664–670. DOI: 10.1038/nmeth.3433.
16. Lee YJ, Hoynes-O'Connor A, Leong MC, et al. Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. *Nucleic Acids Res* 2016;44:2462–2473. DOI: 10.1093/nar/gkw056.
17. Gootenberg JS, Abudayyeh OO, Lee JW, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 2017;356:438–442. DOI: 10.1126/science.aam9321.
18. Kellner MJ, Koob JG, Gootenberg JS, et al. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat Protoc* 2019;14:2986–3012. DOI: 10.1038/s41596-019-0210-2.
19. Galizi R, Jaramillo A. Engineering CRISPR guide RNA riboswitches for *in vivo* applications. *Curr Opin Biotechnol* 2019;55:103–113. DOI: 10.1016/j.copbio.2018.08.007.
20. Anderson-Lee J, Fisker E, Kosaraju V, et al. Principles for predicting RNA secondary structure design difficulty. *J Mol Biol* 2016;428:748–757. DOI: 10.1016/j.jmb.2015.11.013.
21. Rodrigo G, Landrain TE, Jaramillo A. *De novo* automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proc Natl Acad Sci U S A* 2012;109:15271–15276. DOI: 10.1073/pnas.1203831109.
22. Green AA, Silver PA, Collins JJ, et al. Toehold switches: *de-novo*-designed regulators of gene expression. *Cell* 2014;159:925–939. DOI: 10.1016/j.cell.2014.10.002.

23. Kushwaha M, Rostain W, Prakash S, et al. Using RNA as molecular code for programming cellular function. *ACS Synth Biol* 2016;5:795–809. DOI: 10.1021/acssynbio.5b00297.
24. Isaacs FJ, Dwyer DJ, Ding C, et al. Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* 2004;22:841–847. DOI: 10.1038/nbt986.
25. Chappell J, Takahashi MK, Lucks JB. Creating small transcription activating RNAs. *Nat Chem Biol* 2015;11:214–220. DOI: 10.1038/nchembio.1737.
26. Kundert K, Lucas JE, Watters KE, et al. Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs. *Nat Commun* 2019;10:1–11. DOI: 10.1038/s41467-019-09985-2.
27. Iwasaki RS, Ozdilek BA, Garst AD, et al. Small molecule regulated sgRNAs enable control of genome editing in *E. coli* by Cas9. *Nat Commun* 2020 Mar 13 [Epub ahead of print]; DOI: 10.1038/s41467-020-15226-8.
28. Hanewich-Hollatz MH, Chen Z, Hochrein LM, et al. Conditional guide RNAs: programmable conditional regulation of CRISPR/Cas function in bacterial and mammalian cells via dynamic RNA nanotechnology. *ACS Cent Sci* 2019 Jun 4 [Epub ahead of print]; DOI: 10.1021/acscentsci.9b00340.
29. Oesinghaus L, Simmel FC. Switching the activity of Cas12a using guide RNA strand displacement circuits. *Nat Commun* 2019 May 7 [Epub ahead of print]; DOI: 10.1038/s41467-019-09953-w.
30. Siu K-H, Chen W. Riboregulated toehold-gated gRNA for programmable CRISPR-Cas9 function. *Nat Chem Biol* 2019;15:217–220. DOI: 10.1038/s41589-018-0186-1.
31. Jin M, Garreau de Loubresse N, Kim Y, et al. Programmable CRISPR-Cas repression, activation, and computation with sequence-independent targets and triggers. *ACS Synth Biol* 2019;8:1583–1589. DOI: 10.1021/acssynbio.9b00141.
32. Ferry QRV, Lyutova R, Fulga TA. Rational design of inducible CRISPR guide RNAs for *de novo* assembly of transcriptional programs. *Nat Commun* 2017;8:14633. DOI: 10.1038/ncomms14633.
33. Li Y, Teng X, Zhang K, et al. RNA strand displacement responsive CRISPR/Cas9 system for mRNA sensing. *Anal Chem* 2019;91:3989–3996. DOI: 10.1021/acs.analchem.8b05238.
34. Lutz R, Bujard H. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* 1997;25:1203–1210. DOI: 10.1093/nar/25.6.1203.
35. Pédelacq J-D, Cabantous S, Tran T, et al. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 2006;24:79–88. DOI: 10.1038/nbt1172.
36. Peterson J, Phillips GJ. New pSC101-derivative cloning vectors with elevated copy numbers. *Plasmid* 2008;59:193–201. DOI: 10.1016/j.plasmid.2008.01.004.
37. Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng* 2008;2:5. DOI: 10.1186/1754-1611-2-5.
38. Been MD, Wickham GS. Self-cleaving ribozymes of hepatitis delta virus RNA. *Eur J Biochem* 1997;247:741–753. DOI: 10.1111/j.1432-1033.1997.00741.x.
39. Bikard D, Jiang W, Samai P, et al. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res* 2013;41:7429–7437. DOI: 10.1093/nar/gkt520.
40. Shin J, Noireaux V. An *E. coli* cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synth Biol* 2012;1:29–41. DOI: 10.1021/sb200016s.
41. Dunlop MJ, Cox RS, Levine JH, et al. Regulatory activity revealed by dynamic correlations in gene expression noise. *Nat Genet* 2008;40:1493–1498. DOI: 10.1038/ng.281.
42. Shcherbo D, Murphy CS, Ermakova GV, et al. Far-red fluorescent tags for protein imaging in living tissues. *Biochem J* 2009;418:567–574. DOI: 10.1042/BJ20081949.
43. Zadeh JN, Steenberg CD, Bois JS, et al. NUPACK: analysis and design of nucleic acid systems. *J Comput Chem* 2011;32:170–173. DOI: 10.1002/jcc.21596.
44. Solovyev V, Salamov A. Automatic annotation of microbial genomes and metagenomic sequences. In Li RW (ed.). Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies. Hauppauge, NY: Nova Science, 2011, 61–78.
45. Fisher AG, Ensoli B, Ivanoff L, et al. The sor gene of HIV-1 is required for efficient virus transmission *in vitro*. *Science* 1987;237:888–893. DOI: 10.1126/science.3497453.
46. Strebel K, Daugherty D, Clouse K, et al. The HIV "A" (sor) gene product is essential for virus infectivity. *Nature* 1987;328:728–730. DOI: 10.1038/328728a0.
47. Shin J, Noireaux V. Efficient cell-free expression with the endogenous *E. coli* RNA polymerase and sigma factor 70. *J Biol Eng* 2010;4:8. DOI: 10.1186/1754-1611-4-8.