# Engineering RNA-interacting CRISPR guide RNA for genetic sensing and diagnostics

Roberto Galizi1,2\*, John N. Duncan3\*, Charlotte M Quinn2,4, Marko Storch2,5,6, William Rostain3, Manish Kushwaha3,7, Alfonso Jaramillo3,8,9†#

1Centre for Applied Entomology and Parasitology, School of Life Sciences, Keele University, Keele, UK

2Department of Life Sciences, Imperial College London, London SW7 2AZ, UK

3School of Life Sciences, University of Warwick, Coventry, UK.

4Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK

5Department of Bioengineering, Imperial College London, London SW7 2AZ, UK

6Imperial College Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, UK

7Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

8Warwick Integrative Synthetic Biology Centre (WISB), University of Warwick, CV4 7AL, Coventry, UK.

9Institute for Integrative Systems Biology (I2SysBio), University of Valencia-CSIC, 46980 Paterna, Spain

\*These authors contributed equally to this work and are considered to be co-first authors.

†On leave of absence from CNRS

#To whom correspondence should be addressed. Tel: +44 (0) 24 765 73432; Email; alfonso.Jaramillo@warwick.ac.uk

## Abstract

CRISPR guide RNAs can be manipulated 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 ed and tested, both in living *E. coli* cells andeasily scalable cell-free assays, *cis*-repressed RNA-interacting guide RNA (igRNA) that switch to their active state upon interaction with small RNA fragments as well as long RNA transcripts, including pathogen-derived mRNAs of medical relevance such as the HIV infectivity factor. The proposed CRISPR-igRNAs are fully customisable and easily adaptable to the majority if not all the available CRISPR-Cas variants and therefore can be used to modulate a variety of genetic functions in response to virtually any RNA sequence. We thereby see a large scope of application for therapeutic, diagnostic and biotech applications in both prokaryotic and eukaryotic systems.

## Introduction

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) loci and the associated Cas (CRISPR-associated system) 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, whilst the protein part of the ribonucleoprotein (RNP) complex cleaves or degrades the DNA or RNA target causing its inactivation. In the most exploited type-II CRISPR system found in *Streptococcus pyogenes*, a single Cas9 protein acts as RNA-guided endonuclease by forming a complex with the two RNA fragments (crRNA and tracrRNA) to target and cleave double-stranded DNA.2 A single 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 ~20nt targeting region (or spacer) of the crRNA fragment; 2) the Cas9 handle, which allows its complexing with the Cas9 protein; and 3) the twohairpins at the RNA 3’ end for termination of transcription. The handle and harpin sequences constitute the tracrRNA sequence which partially hybridises to the crRNA to form the targeting component of the CRISPR-Cas system. After forming the ribonucleoprotein (RNP) complex with the Cas protein, 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 hybridisation with the complementary DNA sequence targeted.3 The Cas9 handle is composed of three stem loops, the first being formed through the interaction of the crRNA and tracrRNA RNAs (linked through a short loop in the sgRNA) followed by three smaller stems, the nexus and the two hairpins (Fig. 1A). These secondary structures have proved crucial for the formation of an active RNP complex and can be sensible to modification if the targeting region contains significant secondary structure4 or modification 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 enzymatically inactive or “dead” proteins, such as the most popular dCas9 and dCpf1 (dCas12a), which can be used as RNA-guided DNA-binding protein for genetic inactivation, also referred as CRISPR interfering (CRISPRi). A multitude of other Cas protein variants and fusions have been generated and used for programmable repression or activation of translation as well as for genetic detection6–9. CRISPR gRNAs can also be engineered by adding riboswitches10 or aptazymes11 to create small molecule-responsive gRNAs, or fused to MS2 or PP7 binding regions which are then targeted by MS2/PP7 protein fusions.12 They can also be incorporated into computational RNA circuits regulated by antisense RNA inhibitors13 or used for nucleic acid detection.14–16

RNA engineering is greatly facilitated by the possibility to exploit secondary structure prediction softwares17 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.18,19 We previously shown the use of RNA circuitry to engineer logic operations both *in vitro* and *in vivo*, with either fully artificial parts or interfacing with natural RNAs.20 The CRISPR system combines the flexibility of the RNA component, for programmed targeting of nucleotide sequences, and the highly versatile protein for effecting the output function, in the form a multitude of genetic 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 that are able to change conformation under the binding of a ligand, has allowed the regulation of gene expression to be engineered entirely within the 5’UTR in the absence of any protein co-factors. Ligands based on sRNA molecules have been used to computationally design gene-expression altering riboregulators and anti-terminators.16,18,21–24 We previously proposed a new molecular architecture to engineer RNA-interacting guide RNA (named igRNA) that change their conformation to the active state uniquely upon hybridisation with single-stranded RNAs16. Compared to other similar approaches,25–27 our designs aim to generate fully flexible CRISPR-based genetic editors that can be engineered to target any CRISPR-targetable sequence, according to the CRISPR nuclease used, and can be specifically activated by virtually any RNA molecule according to the ultimate application.16

In this work we used *E.coli­*-based assays to demonstrate, for the first time, coupling of CRISPR activation and RNA environment in a living cell. We tested our novel RNA-responsive igRNA designs in both genetic activation and repression settings upon sensing of sRNA fragments or entire mRNA transcript with high structural complexity such as those encoding for the fluorescent protein mKate2 and the 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.

## Materials and Methods

### Plasmid design

For the *E. coli* assay, the pLlacO-1 promoter28 was combined with sfGFP29 on a pSC101-E93K plasmid30 containing a kanamycin resistance marker. This was further modified using a BioBrick31 pJ23119-igRNA-HDV ribosome32 gBlock to create one of the parental plasmids used in this study. The second parental plasmid was purchased from Addgene (46569)33 and contains dCas9 (p15A, chloramphenicol resistance marker). All plasmids used in this study were generated from these two plasmids (see Supplementary materials and methods).

For the cell-freeassay, the P70a-deGFP plasmid34 was modified to generate the TXTL 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 whilst the TXT4 carries the target in the same position but reverted (nontemplate strand). The constructs expressing igRNAs, RNA triggers or targeting and nontargeting gRNAs controls (AJ1 to 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, nontargeting gRNA (AJ4 and AJ8) and igRNA with the noncorresponding trRNA as negative controls. Plasmids and strains used in this study are listed in Supplementary Tables S1, S2 and S3.

### *E. coli* assay

Cells were grown overnight to early stationary phase at 37 °C with shaking in M9 minimal media18 plus appropriate antibiotics (anhydrotetracycline was always present at a concentration of 100 ng/ml). Cells were then refreshed 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 hours before being added to a black-sided 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 mKate235) readings were taken at 15 minute intervals. Average of measurements taken from 3 biological repeats, each with 3 technical repeats, was calculated and reported together with the relative standard deviations (SD). All measurements were blanked against M9 media (containing antibiotics and supplements) and against MG1655Z1 cells containing empty plasmids with the appropriate antibiotic markers but no igRNA/trigger/dCas9/sfGFP components to generate 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 refreshed in a 96 well plate (5 µl in 195 µl), grown (4h, 37°C, 200 rpm), then refreshed for a second time the same way. Once a cell density of OD600 = 0.3 had been reached, 10 µl of the culture was diluted in 190 µl PBS (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. FCS data was gated to remove cell debris and fluorescence histograms were plotted with the Flow Cytometry GUI for Matlab.

### qPCR analysis

Total RNA was extracted from 1.5 ml of bacterial cells (OD600 nm = 0.6) using a standard Trizol (Thermo Fisher) protocol, the RNA was precipitated with isopropanol, washed in 70 % ethanol and stored in DEPC-water. The RNA was then further treated with DNase I for 1 hr at 37°C, re-extracted with Trizol, re-purified and cleaned. RNA was tested for DNA contamination using the primers used for qPCR, no PCR products were found indicating absence of DNA contamination. 20 µl of cDNA was produced from 1 µg of RNA (ReadyScript cDNA synthesis mix, Sigma-Aldrich) and diluted 1:10 with DEPC-water. 0.5 µl of the diluted cDNA was used for a PCR to positively confirm cDNA integrity. A qPCR master-mix was created according to the manufacturer specifications (KiCqStart SYBR Green qPCR ReadyMix Low ROX, Sigma-Aldrich), by adding the pre-mixed 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 (15s 94°C, 10s 57°C, 20s 72°C), melt analysis and cooling to 30°C. Results were normalised to the expression of the chloramphenicol resistance gene present on the plasmid.

### Gel-shift assay

The 3A igRNA RNA fragment was *in-vitro* transcribed (TranscriptAid T7 High-Yield, Thermo Fisher) 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 before being re-suspended in DEPC-water. 1 µg of this RNA was mixed with 1 µg of synthesised 3A or 2A trigger sRNA (Integrated DNA Technologies), incubated at room temperature for 10 minutes run on a 12 % polyacrylamide gel for 2.5 hours. The RNA was visualised on a UV transilluminator after staining with ethidium bromide for 20 minutes.

### Cell-free automated assay

Cell-free reactions were performed using 5 nm of each construct in a 4 μl final volume per reaction. Each reaction contains a combination of three constructs: 1) TX1 or TX4 deGFP reporter; 2) Cas9 or dCas9 plasmid; 3) plasmid carrying both igRNA (or targeting or nontargeting gRNAs as positive or negative control) and trRNA (2A or 3A). After purification with magnetic beads (AMPure XP, Beckman Coulter), plasmids were resuspended in 1 μl of distilled water and added to 3 μl of myTXTL master mix (Arbor Biosciences, 507024) in 384-well polypropylene microplates (Labcyte, PPL-0200), to obtain the final concentration, using the acoustic liquid handler (Echo 550, Labcyte) as per manufacturer’s instructions. The Synergy Mx microplate reader (BioTek) was used to analyse deGFP fluorescence (485/20 nm excitation and 528/20 nm emission) every 3 minutes for 234 acquisitions at 29 °C. A total of 12 replicates were performed for each combination of plasmids.

## Results and discussion

### **5’ added *cis*-regulatory sequences 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 loop19 preceded by a 7 to 15 nt long “clamp” sequence complementary to the seed region (Fig. 1B). To test the repression of gRNA binding activity in *E. coli* cells, we used a reporter system where the gRNA targets the LacIQ1 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 result of CRISPR activity (Fig. 1C).

Upon assay in *E. coli* cells, we found that the addition of a 7 to 8 nt clamp, predicted to leave the 12 nt seed region of the gRNA fully exposed, resulted in high amount of sfGFP fluorescence, comparable to the positive control (cells induced with IPTG), indicating successful inhibition activation 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 structure without affecting its functional activity. On the other hand, clamp lengths ≥ 9 nt resulted in a marked reduction of fluorescence with an increasing proportion of nonfluorescent cells up to levels comparable to the control measurements lacking dCas9. 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 and relative energy of the hybridising sequence (Fig. 1D).



**FIG. 1.** **(A)** Schematic representation of the unmodified pLacIQ1-targetting gRNA sgRNA with some of the major features highlighted: the 20 nt targeting region including the 12 nt “seed” region, the Cas9 handle and the *S. pyogenes* terminator. **(B)** Schematic representation of the *cis*-repressed gRNA via the 5’ addition of a 15 nt loop region and a 7 to 15 nt clamp sequence. **(C)** Schematic representation of the *E. coli* reporter system used. The unmodified sgRNA complexes with dCas9 targeting the pLacIQ1 promoter to block the expression of the chromosomally inserted LacI repressor thereby resulting in expression of the LacIQ1-driven sfGFP reporter. Inactivation of gRNA binding (upon the addition of blocking clamp sequences) results in repression of the sfGFP signal. **(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 (“GFP + IPTG”). Cells containing only the sfGFP reporter plasmid (“GFP”) or only the Cas9 plasmid (“Cas9 only”) are used as a negative control to show basal fluorescence levels.

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

In order to transform the 5’ clamped gRNA into a robust RNA sensor we added an additional 15 nt “toehold” structure to the 5’ end. The resulting gRNA sequence is thereby formed by two segments complementary to the 5’ and 3’ end of the trigger RNA (the 15 nt variable loop and the 15 nt toehold sequence respectively) and a segment that is complementary to 13 nt of the 20 nt targeting region (clamp sequence). This creates a modified sgRNA with two discrete, separate parts that are fully independent of each other, the first encompass two binding regions for the trigger RNA (toehold and loop), the second is constituted by the clamp and the 20 nt targeting region (Fig. 2A and B). We have termed this device as RNA-interacting guide RNA (igRNA). For our initial igRNA designs, we used the publicly available NUPACK webserver.36 We selected two versions of the igRNA sensor (identified as 2A and 3A) including toehold, clamp and variable region showing minimal complementarity with the downstream gRNA secondary structure. Simultaneously the respective 35-45 nt sRNA trigger sequences were selected based on free and co-folding energy with their cognate igRNA molecules (Fig. 2A). The 3A igRNA was tested in *E. coli* using the same LacI-repressed sfGFP reporter used for the clamp testing with the conditional addition of the trigger RNA (Figure 2B). The 3A igRNA showed minimal levels of activity in the absence of the RNA trigger, comparable to the negative controls where the igRNA or the dCas9 constructs were not added. Interestingly, the addition of the corresponding 3A trigger showed high levels of sfGFP fluorescence, expected as a result of CRISPR activity in blocking the LacI repressor, corresponding to a 26-fold increase compared to the control experiment where the triggering RNA 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 noncomplementary 2A trigger, which showed only a small increase of fluorescence (1.6-fold) compared to the control without trigger (Figure 2C). Specificity of hybridisation between sensor and trigger sequences was also confirmed by the *in vitro* gel-shift assay were the 3A igRNA showed reduced mobility only in the presence of the cognate 3A RNA trigger (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 20nt pLacIQ1 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. This configuration allows the binding of the CRISPR RNP complex to either side of the promoter.37 The hybrid promoter was inserted upstream of a strong RBS to drive sfGFP expression and the resulting reporter plasmid (pHyb-sfGFP) was used to replace the LacI repressor and LacIQ1-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 sRNA trigger was added, whilst combining the previously tested 3A igRNA and cognate 3A sRNA trigger generated a strong reduction (~35-fold) of sfGFP expression (Fig. 3B).

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



**FIG. 2.** **(A)** Schematic representation of igRNA in its inactive state, sRNA trigger and activated igRNA after hybridisation with the sRNA trigger. The igRNA differs from the unmodified sgRNA (represented in Fig. 1A) for the addition of the toehold, clamp and variable loop sequences. The toehold (dark green) and variable loop (purple) regions are respectively complementary to the 3’ (light green) and 5’ sequences (pink) of the sRNA trigger. The igRNA design envision no binding between the sRNA and the clamp sequence therefore trigger, clamp and targeting sequence of the igRNA are fully modifiable. Hybridisation 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 activationreporter. The reporter represented in Figure 1C was modified by adding the RNA trigger component that regulates CRISPR-igRNA guided derepression of sfGFP. **(C)** Average ratio ± SD of fluorescence intensity to OD600 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 trigger (3, 4 and 5). The 3A trigger 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.



**FIG. 3.** **(A)** Schematic representation of the *E. coli­*-basedgenetic 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 olasmid). **(B)** Average ratio ± SD of fluorescence intensity to OD600 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).

### CRISPR-igRNA exhibit selective activation by full-length and translating mRNA

After the successful demonstration of sRNA-specific activation of igRNA, we tested whether full-length and actively translating mRNAs could function as trigger 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 protein. The mKate2 mRNA transcript was interrogated by using the NUPACK tool to select a 39 nt sequence and cognate igRNA switch (3AV9). The same global architecture used in the previously tested 3A igRNA was maintained for the design of the new mRNA-responsive igRNA. *E. coli* testing using thegenetic repressionreporter, previously overviewed in Figure 2B, showed that the addition of the mKate2 mRNA to the 3AV9 igRNA led to a ~2.1-fold increase in sfGFP production 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 hybridisation and consequent activation of the igRNA, we generated a second version of the mKate2 trigger construct by deleting the respective ribosomal binding site (RBS). As expected, the addition of the translationally inactive mKate2 mRNA trigger showed only basal levels of mKate2 fluorescence. However, sfGFP fluorescence was maintained at similar levels compared to the translationally active trigger, indicating that active translation did not affect the ability of the mRNA trigger to activate CRIPSR-igRNA activity (Figure 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 virions38,39. 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 mRNA (named HIV3-igRNA). The subsequent testing for genetic repressionin *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 quantitative PCR analysis supporting the HIV3-igRNA ability to switch to its active form upon interaction with the actively translating viral mRNA (Supplementary Fig. S2).

****

**FIG. 4.** **(A)** Average ratio ± SD of fluorescence intensity to OD600 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 trigger construct where the mKate2 RBS was omitted to obliterate concomitant translation of the trigger mRNA (lane 2). **(B)** Same analysis of cells carrying the HIV-responsive igRNA (HIV3-igRNA) without (lane 1) or with the addition (lane 2) of the actively transcribing HIV VIF mRNA trigger.

### 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 decide to test their activity with both nuclease-deactivated dCas9 and 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 customisable 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 nontemplate DNA strand. We used an acoustic liquid handler (Echo 550, Labcyte) to rapidly setup reactions in volumes as little as 4 μl in 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). Fluorescent signal acquired showed a significant reduction of signal compared to the negative controls when both 3A igRNA and 3A sRNA trigger were added to the reaction with either dCas9 (3.7-fold reduction) or Cas9 (3.4-fold reduction) to target the nontemplate 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 noncognate 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 nontemplate 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 whilst the *in vitro* cell-free assay were performed at the recommended temperature of 29 °C.40 As expected, the standard gRNA (without the additional sensor sequence included to the 5’ to generate igRNAs), used a a positive control, gave a strong reduction of fluorescence signal when used to target either nontemplate or template strand with Cas9 as well as when targeting the nontemplate 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. 5A and Supplementary Fig. S3). These results confirm previous studies showing that nontemplate strand binding of the dCas9-gRNA RNP complex is usually required to prevent inhibition or removal from the running RNA polymerase and efficiently block elongation of transcription via CRIPSRi.7,33 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. 5B).



**FIG. 5.** **(A)** Cell-free testing of 3A igRNAs with the nuclease-deactivated S. pyogenes dCas9 (left) and the active Cas9 (right). Bar graphs showing the average fluorescence intensity as arbitrary unit ± SD emitted by the fluorescent reporter construct carrying the LacIQ1 gRNA target site used for the *E. coli* experimets in the nontemplate strand of the deGPF coding sequence. Repression activity of the 3A igRNA with the corresponding 3A trigger (lane 1) were compared to the 3A trigger with the non-complementary 2A trigger (lane 2), a standard gRNA targeting the same sequence (+C gRNA) as a positive control (lane 3) and a nontargeting gRNA as a negative control. A totoal of 12 replicates were produced for each reaction. **(B)** Schematic representation of our platform for high throughput testing of igRNA designs involving (left to rigt): 1) computational design, 2) automated cloning of DNA plasmids and assembly of cell-free reactions, 3) acquisition and analysis of CRISPR activity using fully customisable reporter constructs, 4) selection of igRNA and trRNA designs for *in vivo* applications.

## Conclusion

The work presented here demonstrates that CRISPR-gRNAs can be converted into *cis*-repressed RNA-interacting gRNA, named igRNAs, to transform standard CRISPR-Cas enzymes into a totally programmable genetic editor that are precisely activated by specific RNA sequences. The added switch sequence blocks the igRNAs ability to bind the targeted DNA sequence, maintaining the Cas9-igRNA complex functionally inactive in the absence of the trigger RNA. 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 with relative ease to engineer igRNAs showing up to 35-fold differential activation of reporter genes in living cells combined with very low baseline activity. Because they primarily rely on RNA folding and RNA-RNA interactions, the proposed igRNA designs are easily adaptable to the many CRISPR-Cas variants currently identified or engineered in the laboratory expanding enormously their range of application. Importantly, our igRNA architecture offers complete flexibility and decoupling in the design of the gRNA spacer, for sequence-specific targeting of DNA or RNA (if combined with appropriate Cas enzymes), and riboswitch sensor, for sequence-specific activation by virtually any desirable RNA molecule. Our results reveal that igRNAs maintain high degree of sensitivity in response to small RNA fragments as well as full-length and actively transcribing 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-*basedreporter. We also described an innovative platform, combining fully customisable and automatable *in vitro* reporter assays, for high-throughput screening and optimisation of CRISPR-igRNA components supporting future development of our technology. We expect 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 analyser, the Crisanti Lab and the London Biofoundry at Imperial College London for support with the cell-free assay and use of the Echo 550 acoustic liquid handler. We would also like to thank M.B. Elowitz for the MG1655Z1 strain.

## Authorship Confirmation Statement

R.G., J.D., M.K. and A.J. designed the research; J.D., C.Q., M.S., W.R. performed the research; R.G., J.D. and A.J. analysed the data and wrote the paper with input from all authors. All co-authors have reviewed and approved of the manuscript prior to submission.

## 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 number BB/M017982/1 (WISB centre)]; and the School of Life Sciences (U. Warwick) [start-up allocation] to A.J. Funding for open access charge: EPSRC/BBSRC [BB/M017982/1 to A.J.].

## Supplementary Material

Supplementary Methods

Supplementary Figure S1, S2 and S3

Supplementary Table S1, S2 and S3

## References

1. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007;315:1709–1712.

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.

3. Jinek M, Jiang F, Taylor DW, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science 2014;343:1247997.

4. Thyme SB, Akhmetova L, Montague TG, et al. Internal guide RNA interactions interfere with Cas9-mediated cleavage. Nature Communications 2016;7:1–7.

5. Mullally G, Aelst K van, Naqvi MM, et al. 5’ modifications to CRISPR Cas9 gRNA can change the dynamics and size of R-loops and inhibit DNA cleavage. bioRxiv 2020;2020.04.09.033399.

6. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 2013;154:442–451.

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.

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.

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.

10. Liu Y, Zhan Y, Chen Z, et al. Directing cellular information flow via CRISPR signal conductors. Nat Methods 2016;13:938–944.

11. Tang W, Hu JH, Liu DR. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. Nat Commun 2017;8:15939.

12. Zalatan JG, Lee ME, Almeida R, et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 2015;160:339–350.

13. 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.

14. Gootenberg JS, Abudayyeh OO, Lee JW, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. Science 2017;356:438–442.

15. Kellner MJ, Koob JG, Gootenberg JS, et al. SHERLOCK: nucleic acid detection with CRISPR nucleases. Nat Protoc 2019;14:2986–3012.

16. Galizi R, Jaramillo A. Engineering CRISPR guide RNA riboswitches for in vivo applications. Current Opinion in Biotechnology 2019;55:103–113.

17. Anderson-Lee J, Fisker E, Kosaraju V, et al. Principles for Predicting RNA Secondary Structure Design Difficulty. J Mol Biol 2016;428:748–757.

18. 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 USA 2012;109:15271–15276.

19. Green AA, Silver PA, Collins JJ, et al. Toehold switches: de-novo-designed regulators of gene expression. Cell 2014;159:925–939.

20. Kushwaha M, Rostain W, Prakash S, et al. Using RNA as Molecular Code for Programming Cellular Function. ACS Synth Biol 2016;5:795–809.

21. Isaacs FJ, Dwyer DJ, Ding C, et al. Engineered riboregulators enable post-transcriptional control of gene expression. Nat Biotechnol 2004;22:841–847.

22. Chappell J, Takahashi MK, Lucks JB. Creating small transcription activating RNAs. Nat Chem Biol 2015;11:214–220.

23. Kundert K, Lucas JE, Watters KE, et al. Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs. Nature Communications 2019;10:1–11.

24. Iwasaki RS, Ozdilek BA, Garst AD, et al. Small molecule regulated sgRNAs enable control of genome editing in E. coli by Cas9. Nat Commun;11 . Epub ahead of print March 13, 2020. DOI: 10.1038/s41467-020-15226-8.

25. 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 . Epub ahead of print June 4, 2019. DOI: 10.1021/acscentsci.9b00340.

26. Oesinghaus L, Simmel FC. Switching the activity of Cas12a using guide RNA strand displacement circuits. Nat Commun;10 . Epub ahead of print May 7, 2019. DOI: 10.1038/s41467-019-09953-w.

27. Siu K-H, Chen W. Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function. Nature Chemical Biology 2019;15:217–220.

28. 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.

29. Pédelacq J-D, Cabantous S, Tran T, et al. Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 2006;24:79–88.

30. Peterson J, Phillips GJ. New pSC101-derivative cloning vectors with elevated copy numbers. Plasmid 2008;59:193–201.

31. Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. J Biol Eng 2008;2:5.

32. Been MD, Wickham GS. Self-cleaving ribozymes of hepatitis delta virus RNA. Eur J Biochem 1997;247:741–753.

33. 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.

34. 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.

35. Shcherbo D, Murphy CS, Ermakova GV, et al. Far-red fluorescent tags for protein imaging in living tissues. Biochem J 2009;418:567–574.

36. Zadeh JN, Steenberg CD, Bois JS, et al. NUPACK: Analysis and design of nucleic acid systems. J Comput Chem 2011;32:170–173.

37. Automatic Annotation of Microbial Genomes and Metagenomic Sequences. Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies (Ed RW Li), Nova Science Publishers;61–78.

38. 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.

39. Strebel K, Daugherty D, Clouse K, et al. The HIV “A” (sor) gene product is essential for virus infectivity. Nature 1987;328:728–730.

40. 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.