Modeling Dynamic Transcriptional Circuits with CRISPRi

Samuel E. Clamons, Richard M. Murray

California Institute of Technology, Pasadena, CA, United States

Abstract

Targeted transcriptional repression with catalytically inactive Cas9 (CRISPRi) can be used to build gene regulatory nets similar in principle to those made with traditional transcription factors, and promises to do so with better orthogonality, programmability, and extensibility. We use a simple dynamical model of CRISPRi to understand its behavior and requirements, and to show that CRISPRi can recapitulate several classic gene regulatory circuits, including the repressilator, a toggle switch, and an incoherent feed-forward loop pulse generator. Our model also predicts that these circuits are highly sensitive to promoter leak, but that promoter leak can be offset with active degradation of dCas. We provide specifications for required fold-repression and dCas degradation rates for several dynamic circuits. Our modeling reveals key engineering requirements and considerations for the construction of dynamic CRISPRi circuits, and provides a roadmap for building those circuits.

1. Introduction

A central challenge of modern bioengineering is that of "programming" cells, and in particular encoding complex, dynamic behavior in cells. Simple examples of genetically encoded dynamic functions include cell state oscillation [1][2][3], event detection and logging [4], molecular fold-change detection [5, 6] or signal level discrimination [7]. A common challenge when engineering complex behavior is the need for multiple *specific* interactions between components. In general, engineering specific, efficient, non-promiscuous reactions between molecules is difficult for a number of reasons, so we are typically forced to utilize natural systems with built-in specificity and selectivity as building blocks for our own devices.

One such natural molecular system is that of the gene regulatory network. Synthetic gene regulatory networks exploit the ability of transcription factors to specifically control the actions of target promoters to "wire" together transcriptional units, much the same way microchip manufacturers use spatial arrangements to wire together silicon-based components like transistors and logic gates. Although gene regulatory networks have been successfully used to build small circuits of interest [1][8][9], to date, gene regulatory networks have not been used to build systems larger than about a dozen regulators. Major barriers to scaling up genetic regulatory networks include a lack of orthogonal transcription factors (the largest verified-orthogonal library of repressors currently consists of about 16 genes [10]), mismatches in output and input levels between different regulators, and increased metabolic burden on the host cell.

One system that promises to bypass some of the limitations of classical genetic regulators is CRISPRi, a system of repression using a catalytically inactive mutant of the programmable endonuclease Cas9 ("dCas"). The dCas protein is inactive until loaded with a guide RNA (gRNA) containing a roughly 20-bp variable region. Once loaded, dCas will bind to any double-stranded DNA sequence matching the variable region of the gRNA, so long as it is immediately upstream of a short PAM region (different for different variants of dCas, but usually NGG) [11]. Binding of dCas can interfere with prokaryotic transcription, either preventing initiation of transcription (if the gRNA is targeted within or immediately around a promoter) or blocking elongation (if the gRNA is targeted downstream of a promoter) [12].

CRISPRi repressors has several potential advantages over traditional transcription factors. The clearest advantage of CRISPRi is that it provides an almost limitless supply of orthogonal repressors – assuming that guides must differ from each other by at least 5 bases, there are 4^{15} or approximately a billion possible orthogonal guide sequences. Another advantage of CRISPRi is the relative uniformity of CRISPRi repressors. Since many CRISPRi operators can be made using the same core promoter sequence, it might be expected that different CRISPRi repressors should act with similar input/output relationships.

Since it was first proposed [13], CRISPRi has been widely used for biophysical characterization of Cas9 [14][15][16][17][18][19] and for control of host gene expression [20][21][22]. More rarely, CRISPRi has been used as a synthetic tool in eukaryotic systems. Layerable CRISPRi endpoint logic gates have been designed at least twice [10][23], and circuits up to eight

gates deep and utilizing up to a dozen gates in total have been constructed (although signal degradation over multiple layers has been a consistent problem). CRISPRi has not yet been widely adapted to create scalable prokaryotic circuits, or to create circuits with programmed dynamic behavior (as opposed to endpoint behavior).

We explore the possibility of building dynamic circuits using CRISPRi, and use a simple model to characterize the functional requirements of CRISPRi for a few example dynamic circuits.

2. A Model of CRISPRi

Most of our analysis will use a mass-action ODE model, even though there is good reason to believe that at least some components of a CRISPRi network will be present at low concentrations (< 10 molecules/cell), on the grounds that 1) ODE models are easier to write, simulate, understand, and analyze than stochastic models, and 2) ODE models insightful even in systems where the bulk assumptions of a mass-action may not be justified.

2.1. A Description of the Model

The following processes make up a simple CRISPRi model, as shown in Figure 1:

- Production of dCas
- Production of gRNAs
- (Optional) Leak production of gRNAs from dCas-bound promoters
- Active degradation of free gRNAs (by RNAses)
- (Optional) Active degradation of dCas and dCas complexes (by proteases)
- Global dilution
- Maintenance of promoter copy number
- Binding and unbinding of gRNAs from dCas
- Binding and unbinding of gRNA:dCas complexes from target promoters

Production of dCas is taken to be constant; gRNA production is constant from unbound promoters, and constant from bound promoters with a lower rate (possibly zero). Guide RNAs are actively degraded at a rate proportional to their abundance. Binding reactions (and unbinding of gRNA from dCas) follow standard mass action binding and unbinding kinetics.

All CRISPRi promoters are assumed to have identical dynamics (aside from the identity of their repressors). Though not shown here, in practice promoter strengths in most circuits can be changed by as much as 10x without qualitatively affecting predicted circuit dynamics.

One unusual feature of this CRISPRi model involves the unbinding of dCas from its target promoters. In particular, dCas binds extremely tightly to its targets. In bacterial cells, the rate of dCas unbinding from DNA is substantially slower than the rate of bacterial replication, even in non-lab-adapted strains with division times of over 100 minutes [19] (though possibly not [18]). This means that dCas effectively only unbinds as a consequence of DNA replication, with the bacterial replication machinery releasing dCas from its target. Thus, any CRISPRi circuit that shows complex, non-monotonic dynamics will require dilution or some other mechanism of actively unbinding dCas from its DNA targets. Additionally, because DNA replication causes detachment of dCas, any model of *in vivo* CRISPRi should couple dCas:DNA unbinding to DNA replication (or, equivalently, cell dilution).

Accordingly, all models in this report apply global cell dilution to all components, using a series of reactions of the form $X \rightarrow \emptyset$, where X is any non-DNA species in the model. A caveat is that total promoter concentrations will be held constant, as dilution of DNA is assumed to be equally balanced by replication. DNA bound to dCas follows a special combined dilution/unbinding "reaction" of the form dCas:DNA $\rightarrow$ DNA + $\frac{1}{2}$ dCas, which proceeds at a rate equal to dilution.

2.2. Parameterization of the Model

A set of typical parameters used in simulations below is given in Table 1. DNA concentrations will be assumed to be 2nM unless otherwise noted, which roughly corresponds to the concentration of a genomically-integrated CRISPRi system in actively growing *E. coli* cells.

It is worth mentioning that there is still a great deal of uncertainty around the kinetics of dCas binding. The rates given above for binding and unbinding of gRNAs to dCas were taken from Mekler *et al.* [15], who measured dCas/gRNA association rates *in vitro*. It is unclear how closely this estimate

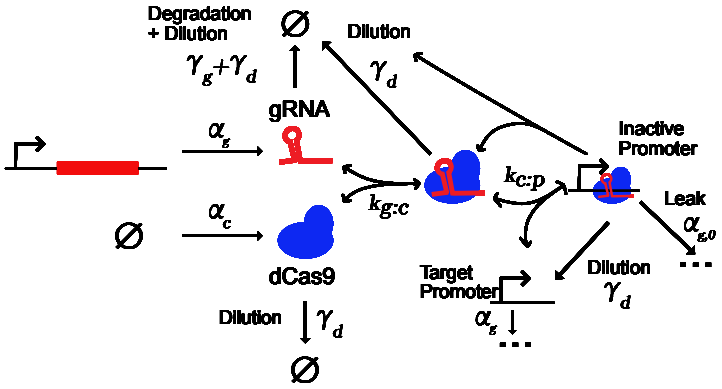

Figure 1: Diagrammatic representation of the CRISPRi model used in this report. Rates for active degradation of dCas and leak activity of dCas-bound promoters are set to zero in some simulations.

Parameter	Typical Value	Units
gRNA Production (α_g)	5	min^{-1}
dCas Production (α_C)	1	min^{-1}
gRNA/dCas Binding ($k^f_{g:c}$)	$\frac{\ln 2}{375}$ ($\approx 1.8e-3$)	$\text{nM}^{-1}\text{sec}^{-1}$
gRNA/dCas Unbinding ($k^r_{g:c}$)	0	sec^{-1}
gRNA:dCas/Promoter Binding ($k^f_{c:p}$)	$\frac{\ln 2}{60}$ ($\approx 1.2e-2$)	$\text{nM}^{-1}\text{sec}^{-1}$
gRNA:dCas/Promoter Unbinding ($k^r_{c:p}$)	0.0003	sec^{-1}
gRNA Degradation	$\frac{\log 2}{100}$ ($\approx 6.9e-3$)	sec^{-1}
Dilution	$\frac{\log 2}{30}$ ($\approx 2.3e-2$)	min^{-1}

Table 1: Typical parameter values used for simulating CRISPRi circuits. Parameters are estimated from the literature, except where noted in the text

follows *in vivo* kinetics. For example, the same authors show that the addition of total human lung RNA to an *in vitro* dCas:gRNA assembly reaction slowed dCas:gRNA binding by at least an order of magnitude. Therefore, the estimate in table 1 may be an optimistic one.

Rates of association between gRNA-loaded dCas and its DNA targets have been more widely studied, but there the literature is still conflicted on their actual values. For example, Gong *et al.* [18] report an unbinding rate of dCas from DNA of about 1/(6.5 min) in a radiolabeled pulse chase assay, but that is incompatible with the observation of Jones *et al.* [19] that dCas dissociation *in vivo* is driven by cell division, or with the real-time, single-molecule measurements of Boyle *et al.* [17], who could not observe

sufficient unbinding events over several hours to estimate an unbinding rate for matched gRNAs (setting an upper bound on unbinding time on th order of hours). The parameters for dCas:DNA interactions chosen in Table 1 use binding rates from Mekler *et al.*, and reflect the canonical understanding in the field that, for all practical purposes, dCas does not unbind from DNA. Still, it should be noted that this is a best-case scenario, and that dCas:DNA binding is likely much slower *in vivo*.

3. Modeling Results

The full CRISPRi model predicts that a variety of dynamic circuits can be constructed from CRISPRi, including a repressilator, a toggle switch, a pulse generating type I incoherent feed forward loop (IFFL), and multiple IFFLs independently driven by a 5-node oscillator (Figure 3. However, these circuits do not operate well under all possible (or even "reasonable") parameter values. We will first attempt to use an approximation of the CRISPRi model, and then simulations under varying parameters, to understand the constraints and requirements of CRISPRi.

3.1. An Approximation

Can we understand CRISPRi dynamics in rational, analytical terms? Should we expect an oscillator made from CRISPRi components to actually oscillate? A toggle switch? An IFFL? According to traditional genetic circuit analysis, the toggle switch [8] and repressilators [1] require a degree of cooperative binding. There is no obvious "cooperative" mechanism in the CRISPRi model, so we might wonder whether we should expect these circuits to function at all.

Unfortunately, the full CRISPRi model outlined in Section 2.2 is not particularly amenable to analysis – even the steady state binding between a single gRNA, dCas, and the gRNA's target is barely analytically tractable (finding it requires the roots of a rather messy 4th order polynomial). To attempt to make some headway, we split the model into those parts making up an "idealized," easy-to-analyze CRISPRi process (informally, "first-order" considerations, though this should not be taken to imply linearity) and kinetic considerations that make CRISPRi difficult to analyze ("second-order" considerations).

What is an "idealized" CRISPRi, then? We propose the following assumptions for a first-order CRISPRi model: dCas is always present in abun-

dance relative to both DNA targets and gRNAs; binding between gRNAs, dCas, and DNA is instantaneous; and binding of dCas to DNA targets is irreversible. Under these assumptions, the binding of gRNA to dCas to target DNA reduces to a simple "linear" model – with increasing concentrations of gRNA, dCas binds 1:1 with DNA until the DNA is completely saturated. obviously neglects some important features of CRISPRi (binding kinetics and loading effects on dCas, to name two), but it is a *tractable* model, and one that can still provide insight into CRISPRi circuits.

Let us consider a CRISPRi toggle switch consisting of two gRNAs repressing each other. The first-order model of the CRISPRi toggle switch requires only two differential equations to model:

$$\frac{dg_1}{dt} = \alpha \max(0, P_1 - g_2) + \alpha_0 \min(P_1, g_2) - \gamma g_1$$
$$\frac{dg_2}{dt} = \alpha \max(0, P_2 - g_1) + \alpha_0 \min(P_2, g_1) - \gamma g_2$$

Here, g_1 and g_2 are concentrations of two mutually-repressing gRNAs, P_1 and P_2 are total concentrations of promoters for those guides, α is the production rate of gRNA from an unbound promoter, α_0 is the production rate of gRNA from a bound promoter (leak), and γ is the division rate of the cell (dilution).

Under what conditions does this system admit two stable steady states? To answer this, we should consider the intermediate steady state of the system, far from the bounds set by 0, P_1, and P_2. In general, toggle-switch-like circuits undergo a supercritical pitchfork bifurcation at this point – when it is stable, the system admits only one state (Figure 2A), but when it is unstable, the system will have two steady states (the "togglable" steady states) (Figure 2B). In particular, the middle steady state will be unstable (and the toggle switch will correctly "toggle") if and only if that system has a single non-trivial steady state that is unstable. This corresponds to the case where at least one of the eigenvalues of the Jacobian of the system has positive real part. To find when this is true, we note that far from any saturating bounds (where we are likely to find the central steady state), the system reduces to

$$\frac{dg_1}{dt} = \alpha P_1 g_2(\alpha - \alpha_0) - \gamma g_1$$
$$\frac{dg_2}{dt} = \alpha P_2 g_1(\alpha - \alpha_0) - \gamma g_2$$

whose Jacobian has eigenvalues $-(\alpha - \alpha_0) - \gamma$ and $(\alpha - \alpha_0) - \gamma$. The first eigenvalue always has negative real part. The first eigenvalue has positive real part (and the system "toggles") when $\alpha - \alpha_0 > \gamma$. In short, a toggle

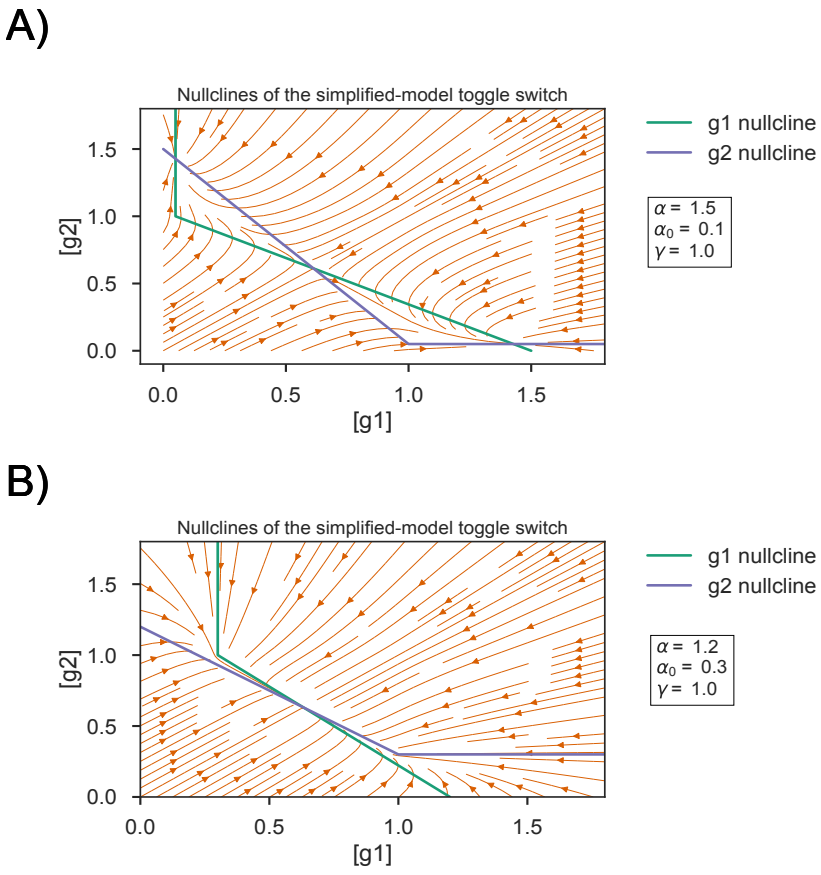

Figure 2: Flow fields for the first-order approximation model of a CRISPRi toggle switch. Depending on the parameters chosen, the toggle could (A) admit two stable steady states separated by an unstable steady state or (B) admit a single stable steady state(B).

should function as long as the difference between production rates of bound and unbound promoters is sufficiently large relative to dilution.

We can apply a similar analysis to a three-node CRISPRi repressilator, which is an oscillator consisting of an odd number of guide RNAs (three, here) in a circular circuit topology, each gRNA repressing the next in the cycle (Figure 4A). Bounded dynamical systems with repressilator-like architecture typically have a single non-trivial steady state. As with the toggle, the desired behavior (oscillations, in this case) occurs when that central steady state is unstable. The Jacobian for a three-node CRISPRi repressilator has eigenvalues $\frac{1}{2}\left(\pm\sqrt{-3(\alpha-\alpha_0)^2}+(\alpha-\alpha_0)-2\gamma\right)$ and $-(\alpha-\alpha_0)-\gamma$. The last eigenvalue always has negative real part. The first pair of eigenvalues each have positive real part (and the system oscillates) when $\frac{\alpha-\alpha_0}{2}>\gamma$. As with the toggle switch, the difference between production rates of bound and unbound promoters must be sufficiently great for the CRISPRi repressilator to oscillate.

Both cases are somewhat surprising given the lack of cooperative behavior in CRISPRi, as some degree of cooperativity is generally thought to be required for both the toggle switch [8] and the repressilator [24]. What the simplified CRISPRi model shows is that cooperativity is not strictly necessary in either of these circuits – rather, cooperativity is necessary in either of these circuits only when genes are expected to bind in a Hill-like fashion.

Now that we have conditions on the parameters required to make functional toggle switches and 3-node repressilators, we will see if those conditions still hold under the full CRISPRi model.

3.2. Repressilators

Consider, again, the 3-node CRISPRi repressilator (Figure 4A). It is quite slow, with a period of about five hours, or about twice as long as the original protein-based repressilator [1]. This is partly because the binding events in CRISPRi operate on roughly the same time scale as translation (minutes), but mostly because the time scale of the repressilator is limited by degradation and dilution rates, not production rates, as we will see in later experiments.

Perhaps more importantly, the CRISPRi repressilator does not oscillate under the default parameters outlined in Table 1, instead converging quickly to a uniform steady state. It would seem we can't simply throw three gRNAs together in a repressing ring and expect it to oscillate.

There are a few different knobs we can turn to recover oscillations from a CRISPRi repressilator. We can decrease the rate of production of either dCas or gRNAs; we can speed the binding between dCas:gRNA complexes

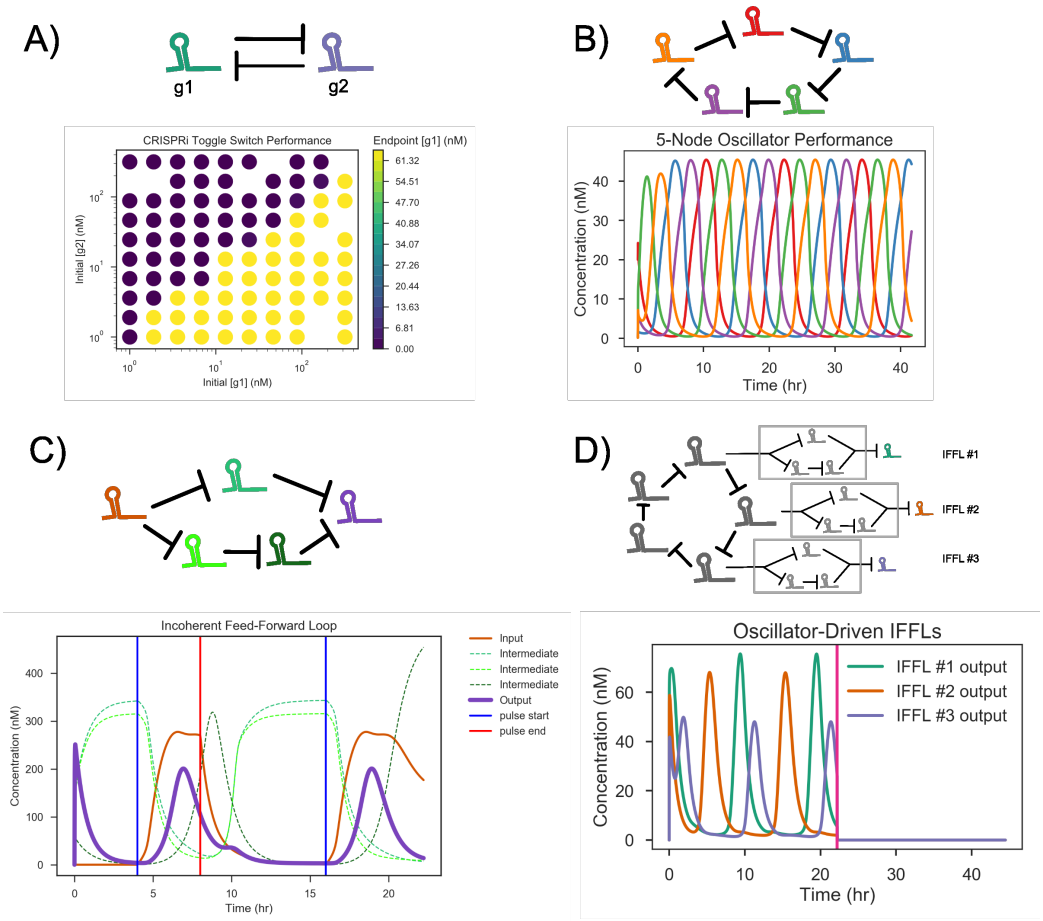

Figure 3: Four examples of dynamic circuits made from CRISPRi components, simulated with the full CRISPRi model. Nodes in circuit diagrams represent gRNA expression units; blunted arrows represent dCas-mediated repression. A) A toggle switch, with endpoint steady states shown for a variety of initial conditions of gRNA concentration. B) A 5-node oscillator. C) A type-I IFFL (pulse generator). The purple trace tracks the IFFL output. Vertical blue and red lines mark activation and return to baseline of the input gRNA promoter, respectively. D) Outputs of two IFFLs driven by a 5-node oscillator. At the half-way time, the node of the oscillator driving the IFFLs is removed, demonstrating that the IFFLs require driving to pulse.

and DNA (in contrast, speeding binding between dCas and gRNAs appears to have little effect); we can add active degradation of dCas (also increasing the *speed* of the oscillator considerably); and we can grow the repressilator from three nodes to five nodes (Figure 4C-F).

Notably, the condition for oscillation given by the first-order approximate CRISPRi model is satisfied in all of the simulations shown in Figure 4. According to the first-order model, all of those systems should oscillate. That some of them do not indicates that the first-order model is, unfortunately, far from sufficient for explaining CRISPRi circuit behavior. The first-order model might be useful for screening out architectures that will not work under *any* parameters, but for now we are stuck with simulation to gain intuition about CRISPRi circuits (Figure 5C).

Unfortunately, the CRISPRi repressilator is not, in general, robust against leak. The first-order model predicts that an increase in leak should stabilize the system towards a steady state, eventually driving it to equilibrium with no oscillations. This is qualitatively accurate, though not quantitatively so – addition of as little as 1% leak destroys oscillations in all but one of the simulations shown in Figure 4 (the exception is that of Figure 4D, which breaks between 2 and 4% leak). This is a serious practical concern – dCas repression has been reported in cell-free extract with fold-repression between 7 and 100, which puts the best CRISPRi repression in a leak range that is unlikely to allow a repressilator to work, and as we will show later, CRISPRi transcriptional units designed for modularity (as described in section 2.1) are even more leak-prone. Fortunately, loss of oscillations by leak can be offset by adding active degradation to dCas, for example as might occur if the protein were fused to a degradation tag.

The tradeoff between degradation and leak rate is shown in Figure 5A. This figure is, in effect, a specification sheet for a CRISPRi repressilator, in terms of two of the engineerable features of CRISPRi (dCas degradation rate and leak). The system will not oscillate at all if there is not active degradation of dCas, or if the degradation rate is less than about twice the speed of dilution. However, the system *also* will not oscillate if dCas degradation is faster than about 5.5 times the speed of dilution. Furthermore, the more closely dCas degradation can be tuned to its "optimal" value, the more leak the system will be able to tolerate – but no matter what the degradation rate of dCas, leak cannot be more than a few percent of unbound expression. Those findings strongly constrain the dCas variants and CRISPRi-repressable promoters that can be used to make this particular dynamic circuit. Inter-

estingly, the toggle switch has similar parameter requirements on these two axes, suggesting that there may be some requirements shared by some class of dynamic CRISPRi circuits.

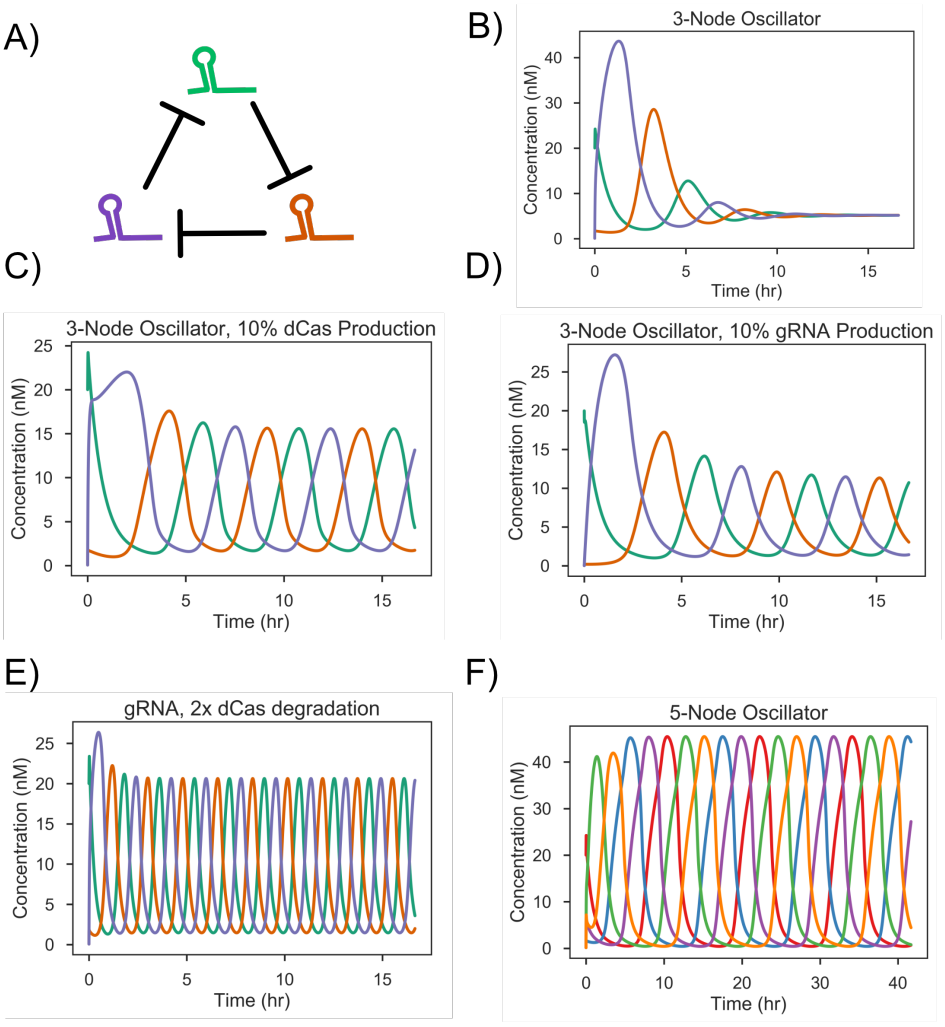

Figure 4: Exploration of the repressilator under a model of CRISPRi. A) A CRISPRi repressilator. B) Simulated dynamics of a 3-node repressilator with parameters taken from Table 1 with 2 nM of each gRNA expression unit, no promoter leak, and no dCas degradation. C-F) Simulated dynamics of CRISPRi repressilators as in (B), but with (C) 1/10 the production rate of dCas or (D) gRNAs, (E) active dCas degradation at twice the rate of dilution, and (F) five nodes instead of three in the ring.

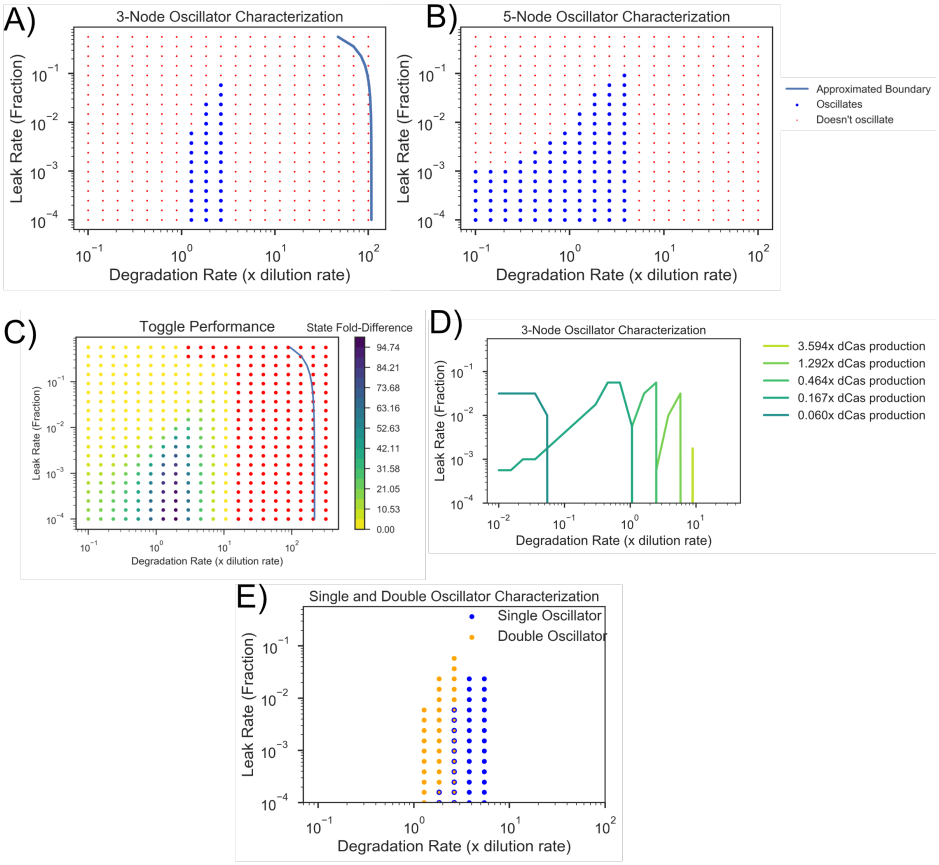

Figure 5: Discrete parameter exploration of CRISPRi repressilators. A-B) Leak vs. degradation rate. Leak is given in units of the rate of gRNA production from an unbound promoter. Degradation rate is in units of cell division rate. Blue dots mark parameter combinations that admit oscillations; red dots mark parameter combinations that damp to a steady state. The solid blue line marks the boundary of oscillation-admitting parameter space according to the first-order model. Results are shown for (A) a 3-node oscillator and (B) a 5-node oscillator. C) Leak vs. degradation in the toggle switch. Dot color corresponds to the fold-difference between high and low states for one guide when simulated. Red dots represent parameters under which the difference between toggle states is less than one part in a thousand. D) Leak vs. degradation vs. dCas production rate. Each curve marks the upper edge of the parameter space that admits oscillations for a single dCas production rate. E) Leak vs. degradation as in (A) (blue), plus the same for a cell with two identical, orthogonal repressilators (orange). Production rates are given in units of min^{-1} (the "default" dCas speed used in (A)).

We can produce a similar "specification sheet" for dCas degradation and leak for a 5-node CRISPRi repressilator, as shown in Figure 5B. The 5-node repressilator is more robust than the 3-node oscillator. Indeed, the 5-node repressilator can operate with as much as 10% leak or as little as no dCas degradation at all. In the case of CRISPRi repressilators, bigger is not only better, but potentially easier.

On the other hand, the parameter requirements of the toggle switch appear to be quite similar to those of the 3-node repressilator (Figure 5C). Admittedly, the toggle switch and 3-node repressilator have very similar architecture, but the fact that both circuits require similar degradation rates and minimum promoter leak suggests that the regime of functional repressilators may have not-yet-understood underlying properties that are broadly useful for constructing CRISPRi circuits.

There are more than two tunable knobs in the CRISPRi system. One that we have already seen to be important is the *production* rate of dCas. Figure 5D shows how the target parameter set changes with different levels of dCas production. The good news is that with low enough dCas expression there is no need for dCas degradation (though with dCas steady state levels that low, stochastic fluctuations become a more serious problem). The bad news here is that at least one *engineerable* but not *readily tunable* parameter of CRISPRi (namely, dCas degradation rate) has acceptable value ranges that don't overlap for some choices of dCas production rate. This shouldn't be too much of a problem for making a single repressilator, but it does complicate the design and integration of multiple CRISPRi circuits in the same cell. For example, Figure 5E shows the expected effect of expressing two identical CRISPRi repressilators in parallel with no directly cross-interacting nodes. The increased load on dCas drops the effective steady-state concentration of dCas, which has a similar effect as dropping dCas production rate. Namely, this shifts the required rate of dCas degradation. A repressilator that works on its own can be expected to fail when a second repressilator is added, unless dCas's degradation rate is exquisitely well-tuned. More generally, it seems likely that different circuits may require dCas variants with different degradation rates.

3.3. Stochastic Simulations

Genetic regulatory systems in real cells involve small, finite numbers of molecules acting probabilistically. Discrete, stochastic models of genetic circuits can capture important dynamics of these systems that continuous, de-

terministic models fail to capture. To check the function of several CRISPRi systems with stochastic simulation, We will use BioSCRAPE, a fast, flexible simulator for biological circuits [25]. Typical traces from these simulations are shown in Figure 6

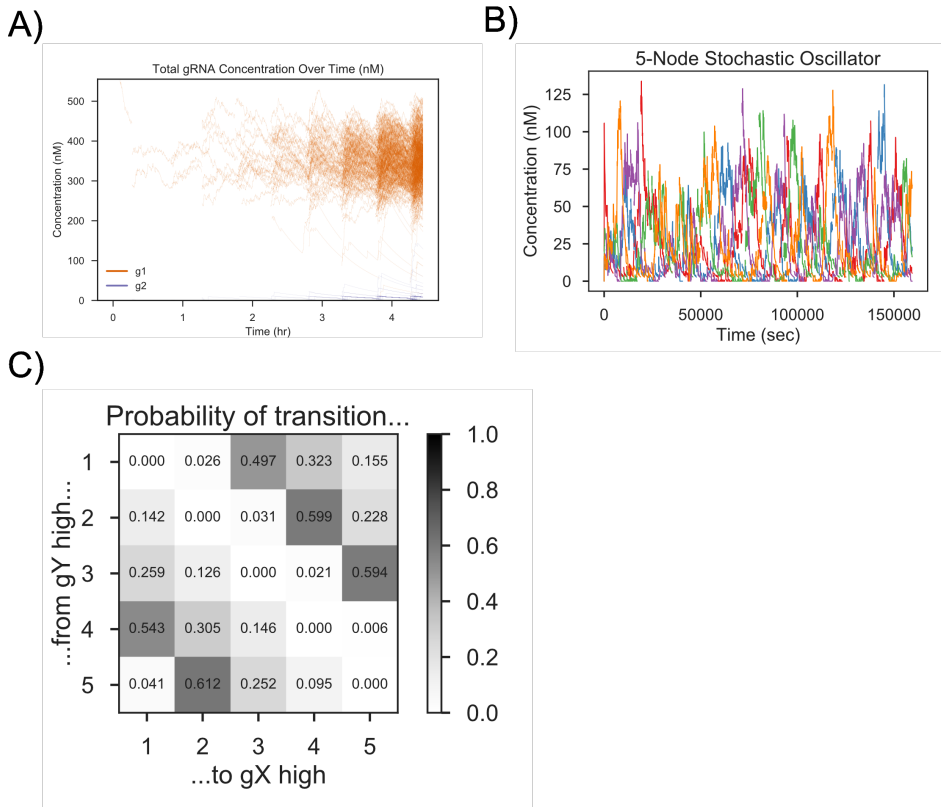

Figure 6: Stochastic simulation of CRISPRi circuits. A-B) Guide RNA dynamics in stochastic models of A) a CRISPRi toggle switch and B) a 5-node CRISPRi repressilator. The "oscillations" observable are caused by dilution and concentration of a constant number of gRNA genes as cells grow and divide. C) Transition matrix for the 5-node repressilator, estimated from stochastic simulation. Expected oscillation order is $1 \to 3 \to 5 \to 2 \to 4 \to 1$.

The toggle switch remains surprisingly robust in stochastic modeling (Figure 6A). Starting with a single dividing cell in a "g_1 high" state, the toggle switch only spontaneously flips in approximately 1-2% of cells over nine generations (approximately 1,000 cells simulated).

In contrast, a five-node repressilator functions rather less regularly in a stochastic model than in a deterministic one (Figure 6B). The stochastic re-

pressilator does still oscillate in a fashion, with each gRNA species cycling with distinguishable peaks and troughs. However, the *order* of gRNA peak expression is not completely regular. To quantify the "accuracy" of the 5-node repressilator, we simulated approximately 600 (in-simulation) hours of growth, division, and repressilation for a single cell. We then estimated a transition matrix for the repressilator, in which each entry represents a probability that a single gRNA will next become most abundant in the cell, given that another gRNA is currently most abundant (so a perfectly-operating five-node repressilator would have a transition matrix with all 0 elements, except elements representing a transition from any gRNA to the RNA two nodes downstream ($1 \rightarrow 3$, $2 \rightarrow 4$,...), which would be 1). The CRISPRi five-node repressilator is far from ideal – although the most likely transition for any particular cell state is still the correct one, the repressilator tends to skip one node too far, skipping most of the cycle. For some applications, this may be acceptable (for instance, if the repressilator is used simply to produce periodic pulses of one species over time, without strict frequency requirements). For any circuit where the order of pulses matters, however, this circuit is unlikely to perform as required.

4. Engineering Requirements of CRISPRi

Although we do not yet have a complete set of guidelines for engineering any CRISPRi system or combinations of systems, the models and simulations in section 3 do provide a few key lessons:

- Active degradation of dCas can improve circuit performance, and is sometimes necessary.
- Current CRISPRi fold-repression is insufficient for many dynamic circuits.
- The speed at which dCas:gRNA complexes bind to DNA is important for circuit function, and implementing CRISPRi circuits *in vivo* may require increasing this binding rate.

Each of these lessons is accompanied by an engineering requirement. Some possible strategies for fulfilling these requirements are outlined below.

4.1. Degradation of dCas

E. coli can actively degrade proteins using the ClpXP system, which uses the proteases ClpXP and ClpXA selectively degrade proteins bearing ssrA, a small C-terminal peptide tag. The native ClpXP system degrades at least 400 times faster than the rate of cell dilution, which is far too fast for any of the circuits outlined in Section 3 [26]. However, targeted mutations to the ssrA tag have been used to tune degradation rates to anywhere between 2 and 100 times the rate of dilution, which falls solidly into the target degradation range outlined in Figure 5 [27]. Alternatively, dCas could be degraded using the *mf*-Lon protease system, which is similar to the ClpXP system in function and tunability but is orthogonal to any system in *E. coli* [28][29].

4.2. Improving Fold-Repression

The CRISPRi repressors reported in the literature typically repress with strengths between 10x and 100x. Moreover, the strongest CRISPRi repressors are typically targeted inside the target promoter, which limits their design space severely. Simulations suggest that leaks reported for elongation-blocking CRISPRi will break toggle switch and oscillator circuits unless degradation rates are extremely well-tuned. Engineering high fold-change in a promoter may be challenging, but several simple strategies are possible for improving repressor performance.

The effectiveness of dCas repression depends on the location of the gRNA target sequence with respect to the promoter, with more than 10-fold variation in repression strength having been shown for various target locations [30][12][31]. Therefore, a simple way to engineer a better repressor may be to simply "tile" a region between the promoter and 5'-UTR with gRNA binding sites and screen for the promoter with the strongest repression.

Another way to strengthen the repression of CRISPRi promoters is to use a weaker constitutive promoter as a core sequence. Weaker promoters generally bind less strongly to sigma factors, which allows the repressor to more easily displace nascent RNA polymerase complexes, which leads to stronger repression relative to the strength of the promoter. This mechanism should apply to CRISPRi repressors, though only if those repressors are blockers of transcriptional initiation.

4.3. Faster dCas:DNA Binding

As previously mentioned, the results shown in this report assume "best-case" binding rate of dCas:gRNA complexes to their target DNAs, with bind-

ing rates taken from *in vitro* association rate measurement. In live *E. coli*, dCas binding to DNA is much slower, with a single dCas complex estimated to require a few hours to bind to its target [19]. A likely explanation for Cas9's slow binding time is that dCas spends most of its time transiently bound to off-target PAM sites in the genome. The dCas protein can temporarily bind to any double-stranded DNA site with a correct PAM sequence. When it does, it briefly opens the DNA helix to "check" whether its associated guide RNA matches the sequence immediately adjacent to the PAM. Each non-target PAM present in the cell slows the rate of correct binding by acting as a low-affinity "decoy" binding site, which slows binding to the target promoter. The *S. pyogenes* Cas9 (by far the most widely-used Cas variant, and the one used exclusively in this report) uses the PAM "NGG", which can be expected to appear roughly 750,000 times in the (approximately diploid) genome of a growing *E. coli* cell. This represents substantial barrier to correct target identification.

There are several possible solutions to the problem of slow dCas:DNA binding *in vivo*. The most straightforward, at least conceptually, would be to move out of cells entirely and construct circuits exclusively in TX-TL or another cell-free system. Since cell-free systems do not have DNA replication or dilution to remove dCas from DNA, this would have to be done either in a microfluidic device capable of manually diluting a running reaction (see [32]) or using degradation-tagged dCas proteins.

Another simple way to speed up dCas:DNA binding rates would be to simply increase the concentration of either dCas (by increasing the baseline production of dCas) or target (by either genomically integrating multiple copies of the CRISPRi circuit or by expressing the circuit off of a plasmid). Simulations so far suggest that, all other things being equal, increasing the rate of dCas production has the effect of narrowing the window of acceptable dCas degradation rates. More simulation will be required to determine the feasibility of either of these two interventions.

Another possible solution would be to use a dCas (or another programmable binding protein) with a different, more complex PAM sequence. For example, the *Treponema denticola* Cas9 (TD-Cas) uses the PAM NAAAAC. After accounting for nonspecific PAM recognition, TD-Cas PAM sites ought to occur between 64 and 1024 times less frequently than *S. pyogenes* Cas PAM sites, with a corresponding increase in binding rate [11]. If the *other* kinetics of TD-dCas are similar to those of *S. pyogenes* dCas, then we should expect TD-dCas binding to DNA to be only perhaps twice as slow as *in vitro* dCas,

which should be quite manageable.

5. Conclusions and Future Directions

CRISPRi remains an intriguing technology for scaling up genetic regulatory networks. However, building functional CRISPRi circuits is not as simple as sketching a repression net and targeting gRNAs against each other accordingly. In particular, the *general* functional requirements of CRISPRi circuits are still unknown. Our work hints that the an important general requirement of CRISPRi is similarity to a reduced model, but further work is required to identify exactly what constitutes "similarity" in this case.

ODE simulations of CRISPRi do reveal *specific* functional requirements regarding dCas9 protein regulation and repressor characteristics. Some degradation of dCas may be required, but not *too* much; some leak from dCas-repressed promoters is acceptable, but not *too* much. Long timescales of DNA binding make CRISPRi construction more difficult, but not fatally so. Meeting these technical requirements will require part engineering of both dCas and CRISPRi promoter motifs, but using these and other insights, it should be an achievable goal to build and express CRISPRi circuits, which would constitute an important milestone toward engineering cells with complex programmable behavior.

6. Acknowledgments

This work was supported by the Human Frontiers Science Program, and by National Science Foundation award number 1317694.

[1] M. B. Elowitz, S. Leibler, A synthetic oscillatory network of transcriptional regulators, Nature 403 (2000) 335–338.

[2] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, J. Hasty, A fast, robust and tunable synthetic gene oscillator, Nature 456 (2008) 516–519.

[3] A. Swaminathan, M. Gomez, D. Shis, M. R. Bennett, R. M. Murray, Stochastic Gene Expression in Single Gene Oscillator Variants, in: Synthetic Biology: Engineering, Evolution and Design (SEED) Conference.

[4] V. Hsiao, Y. Hori, P. W. Rothemund, R. M. Murray, A populationbased temporal logic gate for timing and recording chemical events, Molecular Systems Biology 12 (2016) 869.

[5] L. Goentoro, O. Shoval, M. W. Kirschner, U. Alon, The Incoherent Feedforward Loop Can Provide Fold-Change Detection in Gene Regulation, Molecular Cell 36 (2009) 894–899.

[6] J. Kim, I. Khetarpal, S. Sen, R. M. Murray, Synthetic circuit for exact adaptation and fold-change detection, Nucleic Acids Research 42 (2014) 6078–6089.

[7] J. R. Rubens, G. Selvaggio, T. K. Lu, Synthetic mixed-signal computation in living cells, Nature Communications 7 (2016) 11658.

[8] T. S. Gardner, C. R. Cantor, J. J. Collins, Construction of a genetic toggle switch in Escherichia coli, Nature 403 (2000) 339–342.

[9] A. a. K. Nielsen, B. S. Der, J. Shin, P. Vaidyanathan, V. Paralanov, E. a. Strychalski, D. Ross, D. Densmore, C. a. Voigt, Genetic circuit design automation., Science (New York, N.Y.) 352 (2016) aac7341.

[10] A. A. Nielsen, C. A. Voigt, Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks, Molecular Systems Biology 10 (2014) 763–763.

[11] K. M. Esvelt, P. Mali, J. L. Braff, M. Moosburner, S. J. Yaung, G. M. Church, Orthogonal Cas9 proteins for RNA-guided gene regulation and editing, Nature Methods 10 (2013) 1116–1121.

[12] L. S. Qi, M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, A. P. Arkin, W. A. Lim, Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression, Cell 152 (2013) 1173–1183.

[13] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A Programmable Dual-RNA Guided, Science 337 (2012) 816–822.

[14] H. Ma, L.-C. Tu, A. Naseri, M. Huisman, S. Zhang, D. Grunwald, T. Pederson, CRISPR-Cas9 nuclear dynamics and target recognition in living cells, The Journal of Cell Biology (2016) jcb.201604115.

[15] V. Mekler, L. Minakhin, E. Semenova, K. Kuznedelov, K. Severinov, Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3'-terminal segment of guide RNA, Nucleic Acids Research 44 (2016) 2837–2845.

[16] D. Singh, S. H. Sternberg, J. Fei, J. A. Doudna, T. Ha, Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9, Nature Communications 7 (2016) 12778.

[17] E. A. Boyle, J. O. L. Andreasson, L. M. Chircus, S. H. Sternberg, M. J. Wu, C. K. Guegler, J. A. Doudna, W. J. Greenleaf, High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding, Proceedings of the National Academy of Sciences 114 (2017) 5461–5466.

[18] S. Gong, H. Hong Yu, K. Johnson, D. Taylor, DNA unwinding is the primary determinant of CRISPR-Cas9 activity, Bioarxiv (2017).

[19] D. Jones, C. Unoson, P. Leroy, V. Curic, J. Elf, Kinetics of dCas9 Target Search in Escherichia Coli, Biophysical Journal 112 (2017) 314a.

[20] J. G. Zalatan, M. E. Lee, R. Almeida, L. A. Gilbert, E. H. Whitehead, M. La Russa, J. C. Tsai, J. S. Weissman, J. E. Dueber, L. S. Qi, W. A. Lim, Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds, Cell 160 (2015) 339–350.

[21] L. Nissim, S. D. Perli, A. Fridkin, P. Perez-Pinera, T. K. Lu, Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells, Molecular Cell 54 (2014) 698–710.

[22] L. a. Gilbert, M. H. Larson, L. Morsut, Z. Liu, A. Gloria, S. E. Torres, N. Stern-ginossar, O. Brandman, H. Whitehead, J. a. Doudna, W. a. Lim, S. Jonathan, CRISPR-Mediated Modular RNA-Guided Regualtion of Transcription in Eukaryotes, Cell 154 (2013) 442–451.

[23] M. W. Gander, J. D. Vrana, W. E. Voje, J. M. Carothers, E. Klavins, Digital logic circuits in yeast with CRISPR-dCas9 NOR gates, Nature Communications 8 (2017) 15459.

[24] H. E. Samad, D. D. Vecchio, M. Khammash, Repressilators and promotilators: loop dynamics in synthetic gene networks, Proceedings of the American Control Conference, 2005. (2005) 4405–4410.

[25] A. Swaminathan, V. Hsiao, R. Murray, Quantitative Modeling of Integrase Dynamics Using a Novel Python Toolbox for Parameter Inference in Synthetic Biology, Bioarxiv (2017).

[26] C. M. Farrell, A. D. Grossman, R. T. Sauer, Cytoplasmic degradation of ssrA-tagged proteins, Molecular Microbiology 57 (2005) 1750–1761.

[27] B. P. Landry, J. Stöckel, H. B. Pakrasi, Use of degradation tags to control protein levels in the cyanobacterium Synechocystis sp. strain PCC 6803, Applied and Environmental Microbiology 79 (2013) 2833–2835.

[28] E. Gur, R. T. Sauer, Evolution of the ssrA degradation tag in Mycoplasma: specificity switch to a different protease., Proceedings of the National Academy of Sciences of the United States of America 105 (2008) 16113–16118.

[29] D. E. Cameron, J. J. Collins, Tunable protein degradation in bacteria, Nature Biotechnology 32 (2014) 1276–1281.

[30] D. Bikard, W. Jiang, P. Samai, A. Hochschild, F. Zhang, L. A. Marraffini, Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system, Nucleic Acids Research 41 (2013) 7429–7437.

[31] R. Marshall, M. Colin, S. Collins, M. Luo, T. Jacobsen, C. Beisel, V. Noireaux, Rapid and scalable characterization of CRISPR technologies using an E. coli cell-free transcription-translation system, Bioarxiv (2017).

[32] H. Niederholtmeyer, Z. Z. Sun, Y. Hori, E. Yeung, A. Verpoorte, R. M. Murray, S. J. Maerkl, Rapid cell-free forward engineering of novel genetic ring oscillators, eLife 4 (2015).