Regulation strategies for two-output biomolecular networks

Emmanouil Alexis[1], Carolin CM Schulte[1,2,#], Luca Cardelli[3], and Antonis Papachristodoulou[1,*]

[1]Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

[2]Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK

[3]Department of Computer Science, University of Oxford, Oxford OX1 3QD, UK

[#]Current affiliation: Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA

[*]Correspondence: antonis@eng.ox.ac.uk

Abstract

Feedback control theory allows the development of self-regulating systems with desired performance which are predictable and insensitive to disturbances. Feedback regulatory topologies are used by many natural systems and have been of key importance in the design of reliable synthetic bio-devices operating in complex biological environments. Here, we study control schemes for biomolecular processes with two outputs of interest, expanding previous traditional concepts describing one-output systems. This is a step forward in building bio-devices capable of sophisticated functions. Regulation of such processes may unlock new design possibilities but it can be challenging due to coupling interactions while potential disturbances applied on one of the outputs may affect both. We therefore propose architectures for robustly manipulating the ratio and linear combinations of the outputs as well as each of the output independently. To demonstrate their characteristics, we apply these architectures to a simple process of two mutually activated biomolecular species. We also highlight the potential for experimental implementation by exploring synthetic realizations both *in vivo* and *in vitro*.

1 Introduction

For more than two decades we have witnessed significant advances in the highly interdisciplinary field of synthetic biology whose goal it is to harness engineering approaches in order to realize genetic networks that produce user-defined cellular outcomes. These advances have the potential to transform several aspects of our life by providing efficient solutions to a long list of critical global issues related to food security, healthcare, energy and the environment [1–6]. A fundamental characteristic of living systems is the presence of multi-scale feedback mechanisms facilitating their functioning and survival [7, 8]. Feedback control enables a self-regulating system to adjust its current and future actions by sensing the state of its outputs. This seems to be the answer to a number of major challenges that prevent successful implementation of synthetic genetic circuits and keep innovative endeavours in the field trapped at a laboratory-stage. Control theory offers a rich toolkit of powerful techniques to design and manipulate biological systems and enable the reliable function of next-generation synthetic biology applications [9–13].

Engineering life aims at constructing modular biomolecular devices which are able to operate in

a controllable and predictable way in constantly changing environments with a high level of burden and cross-talk. It is therefore a requirement for them to be resilient to context-dependent effects and show some kind of adaptation to external environmental perturbations. Several control approaches inspired by both natural and technological systems have recently been proposed allowing for effective and robust regulation of biological networks *in vivo* and/or *in vitro* [14–19]. Despite conceptual differences, these research efforts share a common feature: they focus on biomolecular systems with one output of interest, such as the expression of a single protein.

Building advanced bio-devices capable of performing more sophisticated computations and tasks requires the design of genetic circuits where multiple inputs are applied and multiple outputs are measured. In control engineering these types of systems are also known as multi-input multi-output or MIMO systems [20]. This may be the key for achieving control of the whole cell, which can be regarded as a very complex MIMO bio-device itself. Regulation of processes comprising multiple interacting variables of interest can be challenging since there may be interactions between inputs and outputs. Thus, a change in any input may affect all outputs. At the same time, attempts to apply feedback control by "closing the loop" could be impaired by input - output pairing. Addressing such problems therefore requires alternative, suitably adjusted regulation schemes which take into account the presence of mutual internal interactions in the network to be controlled (open-loop system). The research area of MIMO control bio-systems has up until now remained relatively unexplored. There have been only a few studies towards this direction coming mainly from the field of cybergenetics where a computer is a necessary part of the control feedback loop [21, 22]. In contrast, substantial progress has been made in a closely related area, namely MIMO logic bio-circuits which are able to realize Boolean functions [23, 24] while "multi-layer" control concepts for one-output processes [25, 26] and resource allocation in gene expression [27] have also been proposed .

In this paper, we investigate regulation strategies for biomolecular networks with two outputs of interest which can correspond, for example, to the concentration of two different proteins inside the cell, assuming the presence of mutual interactions. Both the open-loop and the closed-loop system (open-loop system within a feedback control configuration) are represented by chemical reaction networks (CRNs) obeying the law of mass action [8]. Consequently, the entire regulation process takes place in the biological context of interest without the use of computer-aided methods. We exploit "multi-loop" concepts based on two independent feedback loops as well as concepts where the

control action is carried out jointly considering both outputs simultaneously. Moreover, our designs take advantage of the adaptation benefits stemming from integral feedback action realized through molecular sequestration [28].

Specifically, we present regulating architectures, which we refer to as regulators, capable of achieving one of the following control objectives: robustly driving a) the ratio of the outputs; b) a linear combination of the outputs; and c) each of the outputs to a desired value (set point). At steady state, the architectures of a) and b) result in two coupled outputs which can still affect each other, albeit in a specific way dictated by the respective control approach. On the other hand, the architectures for c) achieve steady state decoupling, thus making the two outputs independent of each other. It is important to emphasize that our control schemes can be used for regulation of any arbitrary open-loop process provided that the resulting closed-loop system has a finite, positive steady state and the closed-loop system converges to that steady-state as time goes to infinity (closed-loop (asymptotic) stability). Thus, the present analysis focuses exclusively on such scenarios. Furthermore, we mathematically and computationally demonstrate their special characteristics by applying these schemes to a simple biological process of two mutually activating species. Finally, to highlight their biological relevance and motivate further experimental investigation, we explore potential implementations of our designs.

Results

2 Control schemes with steady state coupling

In Figure 1A we show a general biomolecular process with two outputs of interest for which we first present two bio-controllers aiming to regulate the ratio and an arbitrary linear combination of the outputs, respectively. The different types of biomolecular reactions as well as their graphical representations used in this work are presented in 1B.

2.1 Regulating the ratio of outputs

Figure 1C illustrates a motif which we call R-Regulator and consists of the following reactions:

$$Y_1 \xrightarrow{k_1} Y_1 + Z_1\,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2\,, \quad Y_2 + Z_2 \xrightarrow{k_3} Z_2\,, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \tag{1}$$

This controller consists of two species, Z_1 and Z_2, which annihilate each other. The production of Z_1, Z_2 is catalyzed by the target species Y_1, Y_2, respectively while Y_2 is also inhibited by Z_2.

The dynamics of the R-Regulator are described by the following system of Ordinary Differential Equations (ODEs):

$$\dot{Z}_1 = k_1 Y_1 - \eta Z_1 Z_2 \tag{2a}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta Z_1 Z_2 \tag{2b}$$

Equations (2a)-(2b) give rise to a non physical "memory" variable which enables integration, i.e.:

$$\dot{Z}_1 - \dot{Z}_2 = k_1 Y_1 - k_2 Y_2$$

or

$$(Z_1 - Z_2)(t) = k_1 \int_0^t \left(Y_1(\tau) - \frac{k_2}{k_1} Y_2(\tau) \right) d\tau \tag{3}$$

As a result, assuming closed-loop stability ($\dot{Z}_1$, $\dot{Z}_2 \to 0$ as t $\to \infty$), we get:

$$\frac{Y_1^*}{Y_2^*} = \frac{k_2}{k_1} \tag{4}$$

where the * notation indicates the steady state concentration of a species. As can be seen, the integrand in Equation (3) corresponds to an error quantity which converges to zero over time, thus guaranteeing that the output ratio $\left(\frac{Y_1^*}{Y_2^*}\right)$ will converge to the set point $\left(\frac{k_2}{k_1}\right)$. Moreover, the aforementioned stability depends on the structure of the open-loop process, which is unknown here, as well as the set of the reaction rates/parameter values we select for the closed-loop system.

2.2 Regulating a linear combination of the outputs

In Figure 1D a second motif, which we call LC-Regulator, is depicted. The only difference to the R-Regulator is that species Z_1, Z_2 are also produced through two independent processes with constant rates θ_1, θ_2, respectively. More analytically, the corresponding reaction network is:

$$\begin{gathered} \varnothing \xrightarrow{\theta_1} Z_1\,, \quad \varnothing \xrightarrow{\theta_2} Z_2\,, \quad Y_1 \xrightarrow{k_1} Y_1 + Z_1\,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2\,, \\ Y_2 + Z_2 \xrightarrow{k_3} Z_2\,, \quad Z_1 + Z_2 \xrightarrow{\eta} \varnothing \end{gathered} \tag{5}$$

The dynamics of LC-Regulator is given by the set of ODEs:

$$\dot{Z}_1 = \theta_1 + k_1 Y_1 - \eta Z_1 Z_2 \tag{6a}$$

$$\dot{Z}_2 = \theta_2 + k_2 Y_2 - \eta Z_1 Z_2 \tag{6b}$$

Similar to before, in order to see the memory function involved, we subtract Equations (6a) - (6b) and integrate to get:

$$(Z_1 - Z_2)(t) = \int_0^t \left(\Big(k_1 Y_1(\tau) - k_2 Y_2(\tau) \Big) - \Big(\theta_2 - \theta_1 \Big) \right) d\tau$$

Under the assumption of closed-loop stability ($\dot{Z}_1, \dot{Z}_2 \to 0$ as t $\to \infty$), we have at steady state:

$$k_1 Y_1^* - k_2 Y_2^* = \theta_2 - \theta_1 \tag{7}$$

3 Control schemes with steady state decoupling

We now present three alternative bio-controllers, which we call D-Regulator I, II and III, capable of achieving independent control of each output in the arbitrary biomolecular process (Figure 1A). In particular, D-Regulators are able to drive each output species to a desired steady state concentration unaffected by the behaviour of the other species.

3.1 D-Regulator I

The set of reactions describing D-Regulator I (Figure 2A) is:

$$\begin{aligned} &Y_1 \xrightarrow{k_1} Y_1+Z_1\,, \quad Y_2 \xrightarrow{k_2} Y_2+Z_2\,, \quad Y_1+Z_1 \xrightarrow{k_3} Z_1\,, \quad Y_2+Z_2 \xrightarrow{k_4} Z_2 \\ &\varnothing \xrightarrow{\theta_1} Z_3\,, \quad \varnothing \xrightarrow{\theta_2} Z_4\,, \quad Z_1+Z_3 \xrightarrow{\eta_1} \varnothing\,, \quad Z_2+Z_4 \xrightarrow{\eta_2} \varnothing \end{aligned} \tag{8}$$

Here there are four controller species. The target species Y_1, Y_2 catalyze the formation of two of them, Z_1, Z_2, which, in turn, inhibit the former. In addition, Z_3, Z_4, which are produced independently at a constant rate, participate in annihilation reactions with Z_1 and Z_2, respectively.

The dynamics of D-Regulator I can be modelled using the following set of ODEs:

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{9a}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 Z_4 \tag{9b}$$

$$\dot{Z}_3 = \theta_1 - \eta_1 Z_1 Z_3 \tag{9c}$$

$$\dot{Z}_4 = \theta_2 - \eta_2 Z_2 Z_4 \tag{9d}$$

In contrast to the regulation strategies presented in the preceding section, D-Regulator I includes two memory variables which carry out integral action independently. Indeed, combining Equations (9a), (9c) results in:

$$(Z_3 - Z_1)(t) = k_1 \int_0^t \left(\frac{\theta_1}{k_1} - Y_1 \right) d\tau \tag{10}$$

while combining Equations 9b, 9d gives:

$$(Z_4 - Z_2)(t) = k_2 \int_0^t \left(\frac{\theta_2}{k_2} - Y_2 \right) d\tau$$

Consequently, the steady state output concentrations under the assumption of closed-loop stability ($\dot{Z}_1, \dot{Z}_2 \to 0$ as t $\to \infty$) are:

$$Y_1^* = \frac{\theta_1}{k_1}\,, \quad Y_2^* = \frac{\theta_2}{k_2} \tag{11}$$

3.2 D-Regulator II

By using four controller species as before and exploiting the control concept introduced in [28], we construct D-Regulator II (Figure 2B) consisting of the following reactions:

$$\begin{gathered} Y_1 \xrightarrow{k_1} Y_1 + Z_1\,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2\,, \quad \varnothing \xrightarrow{\theta_1} Z_3\,, \quad \varnothing \xrightarrow{\theta_2} Z_4\,, \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1 \\ Z_4 \xrightarrow{k_4} Z_4 + Y_2\,, \quad Z_1 + Z_3 \xrightarrow{\eta_1} \varnothing\,, \quad Z_2 + Z_4 \xrightarrow{\eta_2} \varnothing \end{gathered} \tag{12}$$

In this case, species Z_3, Z_4 catalyze the formation of the target species Y_1, Y_2, respectively, and Z_3, Z_4 are produced at a constant rate. Furthermore, species Z_1, Z_2 are catalytically produced by Y_1, Y_2, respectively, while the pairs Z_1-Z_3 and Z_2-Z_4 participate in an annihilation reaction.

Note that the species of D-Regulator II are described by the same ODE model as D-Regulator I (Equations (9a)-(9d)). Thus, the memory variables involved as well as the steady state output behaviour (Equation (18)) are identical in these two motifs (provided that close-loop stability is guaranteed). Nonetheless, in general, regulating the same open-loop process via the aforementioned controllers results in different output behaviour until an equilibrium is reached or, in other words, we have different transient responses. This is because of the different topological characteristics of the two motifs which cannot be captured by focusing only on the controller dynamics: considering closed-loop dynamics is required, which is addressed in a later section.

3.3 D-Regulator III

The last bio-controller presented in this study is D-Regulator III (Figure 2C) whose structure is composed of the following reactions:

$$\begin{gathered} Y_1 \xrightarrow{k_1} Y_1 + Z_1\,, \quad Y_2 \xrightarrow{k_2} Y_2 + Z_2\,, \quad \varnothing \xrightarrow{\theta_1} Z_3\,, \quad Z_3 \xrightarrow{k_3} Z_3 + Y_1 \\ Y_2 + Z_2 \xrightarrow{k_4} Z_2\,, \quad Z_1 + Z_3 \xrightarrow{\eta_1} C\,, \quad Z_2 + C \xrightarrow{\eta_2} \varnothing \end{gathered} \tag{13}$$

Here there are three controller species. Z_1, Z_3 interact with the target species Y_1 as well as with each other in the same way as in D-Regulator II. The complex C, which is formed by the binding of Z_1, Z_3, and the third controller species, Z_2, can annihilate each other. Finally, the target species Y_2 catalyzes the production of Z_2 which, in turn, inhibits Y_2 analogous to D-Regulator I.

The dynamics of D-Regulator III can be described by the following set of ODEs:

$$\dot{Z}_1 = k_1 Y_1 - \eta_1 Z_1 Z_3 \tag{14a}$$

$$\dot{Z}_2 = k_2 Y_2 - \eta_2 Z_2 C \tag{14b}$$

$$\dot{Z}_3 = \theta_1 - \eta_1 Z_1 Z_3 \tag{14c}$$

$$\dot{C} = \eta_1 Z_1 Z_3 - \eta_2 Z_2 C \tag{14d}$$

Similarly to the other D-Regulators, the memory function responsible for the regulation of the output Y_1 is carried out by the (non-physical) quantity $Z_3 - Z_1$ (Equation (10)). However, the memory variable related to the output Y_2 is realized in a different way than before. More specifically, combining Equations (14b)-(14d) yields:

$$\dot{Z}_3 + \dot{C} - \dot{Z}_2 = \theta_1 - k_2 Y_2$$

or

$$(Z_3 + C - Z_2)(t) = k_2 \int_0^t \left(\frac{\theta_1}{k_2} - Y_2 \right) d\tau$$

Therefore, assuming closed loop stability, i.e. $\dot{Z}_1$, $\dot{Z}_2 \to 0$ as t $\to \infty$, the steady state output behaviour is:

$$Y_1^* = \frac{\theta_1}{k_1}\ , \quad Y_2^* = \frac{\theta_1}{k_2} \tag{15}$$

4 Specifying the biological network to be controlled

We now turn our focus to a specific two-output open-loop network which will henceforward take the place of the abstract "cloud" process of the preceding sections. This will allow us to implement *in silico* the proposed control motifs and demonstrate the properties discussed above (see **Implementing the proposed regulation strategies**). In parallel, we will be able to explore potential experimental realizations of the resulting closed-loop networks (see **Experimental realization**).

Figure 3A illustrates a simple biological network comprised of two general birth-death processes regarding two target species, Y_1, Y_2. These species are coupled in the sense that each of them is able to catalyze the formation of the other. Such motifs of positive feedback action are ubiquitous in

biological systems [29–31]. In particular, we have the reactions:

$$\begin{gathered} \varnothing \xrightarrow{b_1} Y_1\,, \quad \varnothing \xrightarrow{b_2} Y_2\,, \quad Y_1 \xrightarrow{d_1} \varnothing\,, \quad Y_2 \xrightarrow{d_2} \varnothing\,, \\ Y_1 \xrightarrow{\alpha_2} Y_1 + Y_2\,, \quad Y_2 \xrightarrow{\alpha_1} Y_1 + Y_2 \end{gathered} \tag{16}$$

which can be modelled as:

$$\dot{Y}_1 = b_1 - d_1 Y_1 + \alpha_1 Y_2 \tag{17a}$$

$$\dot{Y}_2 = b_2 - d_2 Y_2 + \alpha_2 Y_1 \tag{17b}$$

For any $d_1 d_2 > \alpha_1 \alpha_2$, ODE system (17a)-(17b) has the following unique positive steady state:

$$Y_1^* = \frac{\alpha_1 b_2 + b_1 d_2}{d_1 d_2 - \alpha_1 \alpha_2}\,, \quad Y_2^* = \frac{\alpha_2 b_1 + b_2 d_1}{d_1 d_2 - \alpha_1 \alpha_2} \tag{18}$$

which is globally exponentially stable (see Section S2 of the supplementary material).

Note that for this system, a change in any of the reaction rates of network (16) due to, for instance, undesired disturbances, will affect the behaviour of both species Y_1 and Y_2 (Figure 3B).

5 Implementing the proposed regulation strategies

We now demonstrate the efficiency of the bio-controllers introduced in **Control schemes with steady state coupling** and **Control schemes with steady state decoupling** by regulating the open-loop network (16) presented in **Specifying the biological network to be controlled**. A detailed analysis of the steady state behaviour of the resulting closed-loop processes can be found in section S3 of the supplementary material.

We show in Figure 4 that R-Regulator and LC-Regulator are capable of driving the ratio and a desired linear combination of the output species to the set point of our choice in the presence of constant disturbances, respectively. Similarly, we illustrate in Figure 5 the ability of D-Regulators to robustly steer each of the output species towards a desired value independently, thus cancelling the steady state coupling. Note that the sets of parameter values used here guarantee closed-loop stability which is, as already discussed, a requirement for successful implementation of the control schemes in question.

Finally, in the topology shown in Figure 5B there are two actuation reactions realized though Z_3 and Z_4. Due to the existence of coupling interactions in the network that we aim to control, it is evident that these actuating species act on both Y_1 and Y_2 simultaneously. Consequently, one could argue that an alternative way of closing the loop would be through a different species pairing (Figure 6). In particular, an annihilation (comparison) reaction between Z_1, Z_4 and Z_2, Z_3 could be used instead (Z_1, Z_2 can be considered as sensing species measuring the outputs Y_1, Y_2, respectively). However, it can be demonstrated (see section S4 of the supplementary material) that this control strategy is not feasible since there is no realistic parameter set that can ensure closed-loop stability.

6 Experimental realization

To highlight the feasibility of experimentally realizing the proposed control schemes, this section describes both *in vivo* and *in vitro* implementations of the open-loop and closed-loop circuits introduced earlier. We first focus on implementations using biological parts that have been characterized in *Escherichia coli* and then discuss a molecular programming approach.

Following the description in **Specifying the biological network to be controlled**, the biological network to be controlled can be realized as shown in Figure 7. In this implementation, Y_1 and Y_2 are heterologous sigma factors [32], which are fused to fluorescent proteins (GFP and mCherry) to facilitate tracking of the output. Through a suitable choice of promoters, Y_1 mediates the expression of Y_2 and *vice versa*. Low levels of Y_1 and Y_2 are continuously produced from constitutive promoters, such as promoters from the BioBrick collection [33]. In all following figures, the biological parts underlying these interactions are not explicitly shown.

6.1 R-Regulator and LC-Regulator

For the proposed implementation of the R-Regulator (Figure 8), Y_2 mediates expression of the hepatitis C virus protease NS3 fused to maltose-binding protein (MBP) (Z_2). Y_1 facilitates expression of a MBP-single-chain antibody (scFv) fusion (Z_1) that specifically binds to and thus inhibits NS3 protease. Inhibition of NS3 protease activity through coexpression with single-chain antibodies in the cytoplasm of *E. coli* has been demonstrated previously [34]. Adding a suitable recognition sequence to Y_2 will further allow for its degradation by NS3. An additional requirement for the LC-Regulator

would be constitutive expression of *malE-scFv* and *malE-scNS3* as indicated in the dashed boxes in Figure 8.

6.2 D-Regulators

Similar to R- and LC-Regulator, the implementation for D-Regulator I makes use of the interaction between NS3 protease and a suitable single-chain antibody (Figure 9A). However, the antibody is solely expressed from a constitutive promoter in this case. As a second protease-protease inhibitor pair, we suggest use of the *E. coli* Lon protease and the phage T4 protease inhibitor PinA as discussed in our previous work [35]. For this purpose, a suitable degradation tag should be added to Y_1.

To realize the two annihilation reactions in D-Regulator II (Figure 9B), we propose the use of σ-factors and anti-σ-factors as described previously [36, 37]. Specifically, Z_3 could be the σ-factor SigW, which is constitutively expressed and mediates expression of SigF (Y_1). SigF mediates expression on the anti-σ-factor RsiW (Z_1), which binds to SigW. Analogous reactions are realized using SigM (Y_2), SigB (Z_4) and RsbW (Z_2).

The design for D-Regulator III may be more difficult to implement experimentally due to the requirement of a two-stage complex formation by three biomolecules (Z_1, Z_2 and Z_3) in addition to the requirement of Z_3 catalysing the production of Y_1 and Z_2 inhibiting Y_2. While it may be possible to achieve the desired behaviour of biomolecules using protein fusions and/or protein engineering, an alternative method to implement this design (as well as all the others) would be via molecular programming as discussed in the following section.

6.3 Molecular programming implementation

In molecular programming, an abstract reaction network is realized by designing a concrete chemical reaction network using engineered molecules, so that the latter network emulates the kinetics of the former. At the edges of the abstract network, appropriate chemical transducers must be introduced to interface the abstract network with the environment. While such transducers are specific to each application, the core network is generic, and DNA (natural or synthetic) is commonly used to construct it. These systems are typically tested *in vitro* in controlled environments, with the eventual aim of embedding them in living cells, or in other deployable physical media.

We focus here on a molecular programming approach based on *toehold mediated DNA strand dis-*

placement [38], which is a kind of reaction between relatively short DNA strands that is not thought to occur frequently in nature. The species of our abstract reaction networks are each represented by an arbitrary (but carefully chosen) DNA strand; they interact with mediating DNA structures that represent the reactions. No other chemicals are used, except suitable buffer solutions, and no external energy source is provided: the reactions run down thermodynamically from the initial molecule populations.

It has been shown that any chemical reaction network (any finite set of abstract chemical reactions with mass action kinetics, up to time rescaling) can be compiled to such DNA molecules [39]. Each abstract reaction is implemented by a sequence of DNA strand displacement operations, but the scheme can readily approximate to an arbitrary degree mass action kinetics [39]. Because of uniform architecture, the reaction rates are naturally equal for all reactions with the same number or reagents. It has also been demonstrated experimentally that the reaction rates can be tuned across multiple orders of magnitude [38], both in large exponential steps by modifying toehold lengths, and in small tuning steps by choosing particular strand sequences. The reaction rates are largely predictable by models of DNA structure [40], although in practice they are tuned experimentally. Implementations of this approach include systems where three abstract reactions must have the same experimental rates to a good approximation [41, 42], and systems with hundreds of distinct interacting sequences [43]. Within this framework, a number of compilation schemes have been proposed. In Figure 10 we illustrate a particular representation of two-input (i.e., bimolecular) two-output reactions, which covers all the reactions used in this paper (when using dummy species for zero-input, one-input etc. reactions). This representation extends uniformly to n-input m-output reactions (where n, m are non-negative integers). Moreover, two-input two-output reactions are themselves sufficient to approximate any chemical reaction network.

7 Discussion

In this paper, we address the challenge of regulating biomolecular processes with two outputs of interest which are, in the general case, co-dependent due to coupling interactions. This co-dependence means that disturbances applied to one of the outputs will also affect the other - each of the output species may be part of an separate, independent networks and, by extension, be subject to different perturbations . Thus, we propose control schemes for efficient and robust manipulation of such pro-

cesses adopting concepts based on both output steady state coupling and decoupling. The proposed regulators describe biomolecular configurations with appropriate feedback interconnections which, under some assumptions, result in closed-loop systems where different types of output regulation can be achieved.

In particular, we present bio-controllers for regulating the ratio and a linear combination of the outputs referred to as R-Regulator and LC-Regulator, respectively, and three bio-controllers for regulating each of the outputs independently, namely D-Regulators I, II, III. At the core of their functioning lies a "hidden" integral feedback action realized in suitable ways in order to meet the control objectives for each case. Integral control is one of the most widely used strategies in traditional control engineering since it guarantees zero control error and constant disturbance rejection at the steady state. This comes from the fact that with this type of control, the existence of a positive/negative error, regardless of its magnitude, always generates an increasing/decreasing control signal. Essential structural components of these designs are production-inhibition loops [35] and/or annihilation reactions [28]. Moreover, to get a more practical insight, we consider a two-output biomolecular network with positive feedback coupling interactions. Treating the network as an open-loop system, we use our control designs to successfully manipulate its outputs in the presence of constant parameter perturbations. At the same time, we discuss an alternative way of "closing the loop" in D-Regulator-II via a different controller species "pairing". Although it may seem reasonable, we show that this feedback configuration leads to an unstable closed-loop system.

The proposed designs can be used to regulate arbitrary biological processes provided that the closed-loop topologies have an asymptotically stable and biologically meaningful equilibrium. We therefore anticipate that they will be useful for building complex pathways that robustly respond to environmental perturbations in synthetic biology applications. To this end, we describe possible experimental implementations of our regulators using either biomolecular species in *E. coli* or molecular programming.

Biological networks are inherently stochastic due to the probabilistic nature of biomolecular interactions [8, 44–46]. In the present study, we use deterministic mathematical analysis and simulations which offer a good approximation of the CRN dynamics when the biomolecular counts are high. Thus, an interesting future endeavour would be to investigate the behaviour of our topologies within a stochastic mathematical framework examining, for instance, both the stationary mean and variance

[47–50]. Implementation of our regulatory architectures in living cells may involve an additional challenge: a decay mechanism related to cell growth, known as dilution [8], (among other factors) needs to be accounted for since it can affect the species concentrations of the controllers. Future work will therefore focus on quantifying this impact in terms of, for example, the steady state error, and explore ways to minimize it [51].

Data availability

The programming codes supporting this work can be found at: `https://github.com/emgalox/MIMO-bio-controllers`.

Author contributions

Conceptualization and methodology, E.A., C.C.M.S., A.P., L.C.; Formal analysis and Software: E.A., Writing, E.A., C.C.M.S., A.P., L.C.; Supervision: A.P., L.C.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by funding from the Engineering and Physical Sciences Research Council (EPSRC) [grant numbers EP/M002454/1 and EP/L016494/1] and the Biotechnology and Biological Sciences Research Council (BBSRC) [grant number BB/M011224/1]. C.C.M.S. was supported by the Clarendon Fund (Oxford University Press) and the Keble College De Breyne Scholarship. L.C. is supported by a Royal Society Research Professorship.

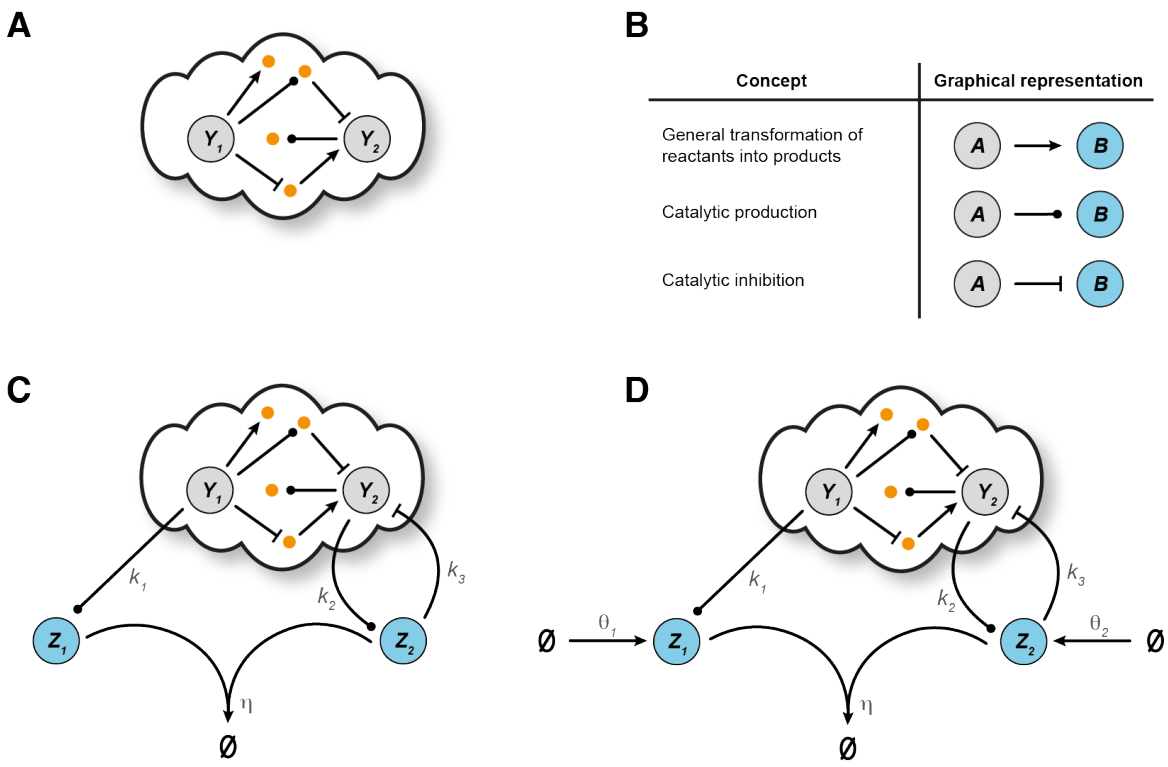

Figure 1: Open-loop biomolecular network and control architectures with steady state coupling.
A Schematic representation of a general biomolecular network with two output species of interest, Y_1, Y_2, and an arbitrary number of other species and/or biomolecular interactions. **B** Graphical representation of the different types of biochemical reactions adopted from our previous work [35]. Schematic representation of a general closed-loop architecture using **C** R-Regulator (CRN (1)) and **D** LC-Regulator (CRN (5)).

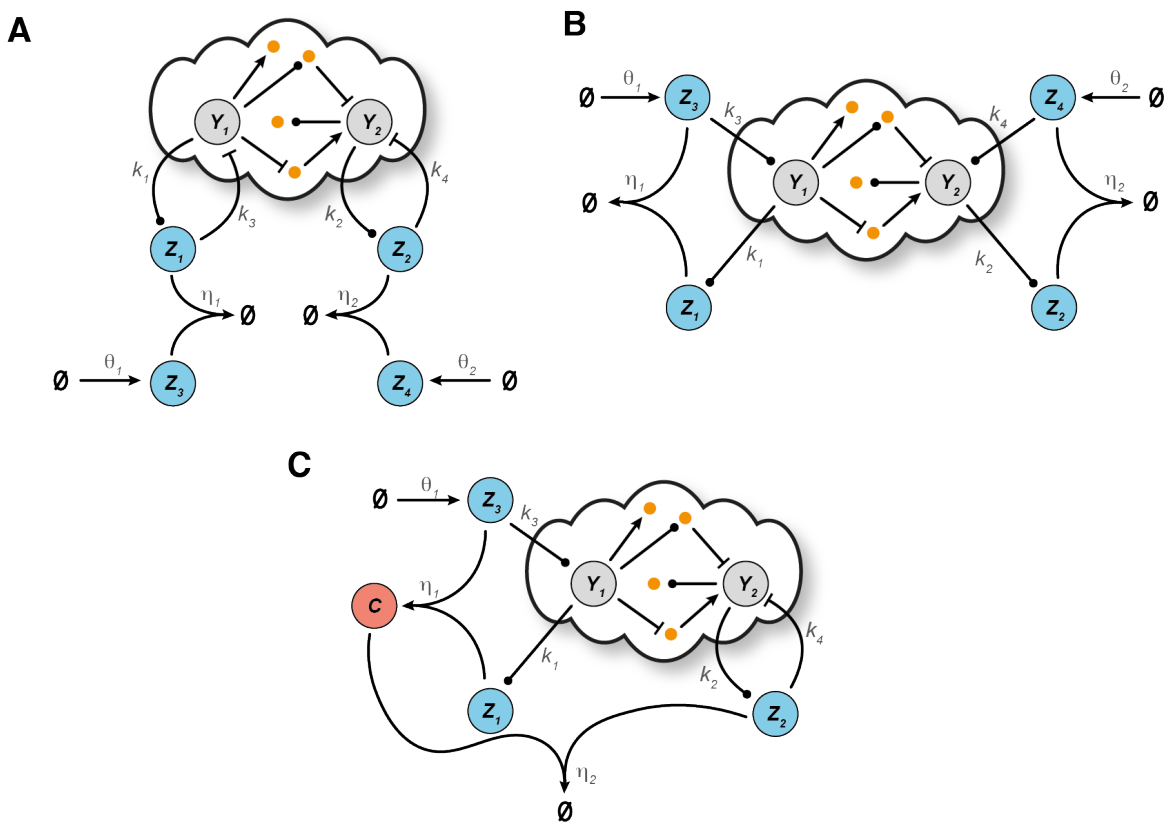

Figure 2: Control architectures with steady state decoupling.
Schematic representation of a general closed-loop architecture using **A** D-Regulator I (CRN (8)), **B** D-Regulator II (CRN (12)) and **C** D-Regulator III (CRN (13)).

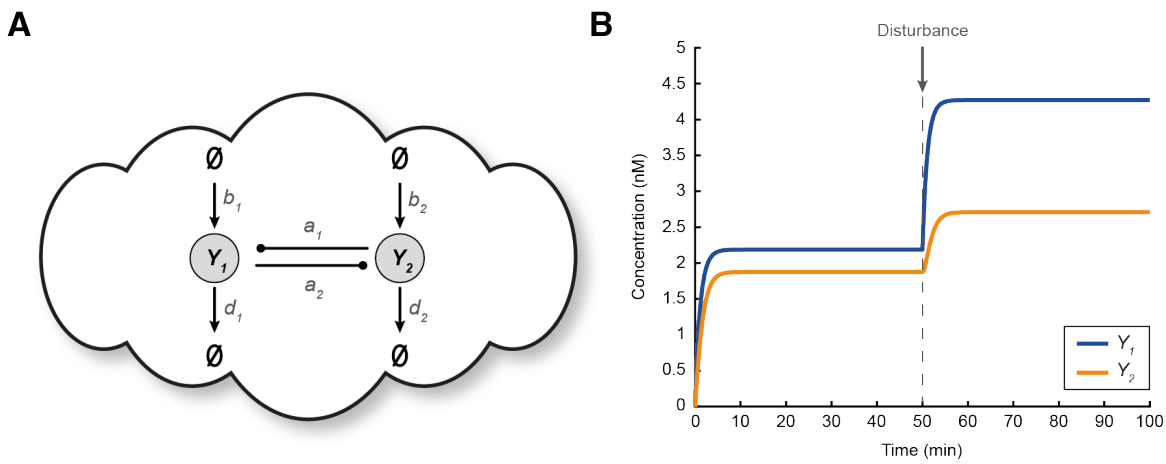

Figure 3: Specifying the open-loop biomolecular network.
A A simple biological process with two mutually activating output species Y_1, Y_2, described by CRN (16). **B** Simulated response of the topology in **A** using the ODE model (17) with the following parameters: $b_1 = 2$ nM min^{-1}, $b_2 = 1$ nM min^{-1}, $d_1 = d_2 = 1$ min^{-1}, $\alpha_1 = 0.1$ min^{-1}, $\alpha_2 = 0.4$ min^{-1}. At time $t = 50$ min, a disturbance on Y_1 is introduced which affects both output species. More specifically, the value of parameter b_1 changes from 2 to 4.

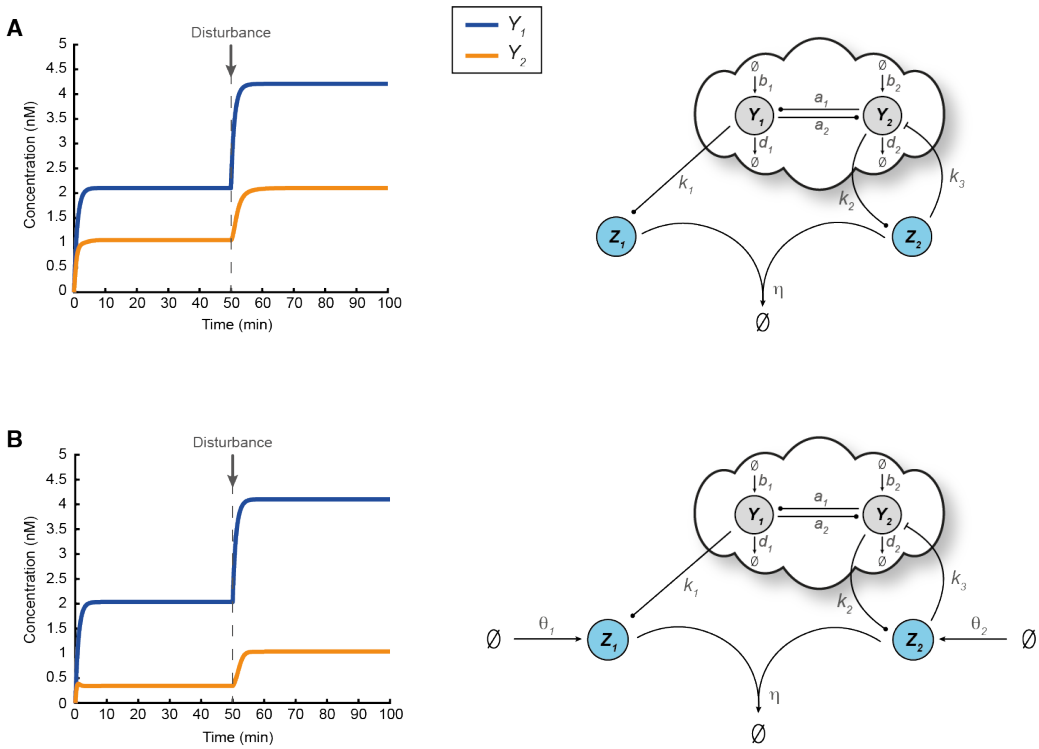

Figure 4: Regulating the ratio and an arbitrary linear combination of the outputs.
A A closed-loop architecture based on the open-loop network shown in Figure 3A and R-Regulator. For the simulated response presented here the following parameters are used: $k_1 = 0.5$ min^{-1}, $k_2 = 1$ min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. At time $t = 50$ min, a disturbance is applied (same as in Figure 3B) which alters the output steady states. Nevertheless, $\frac{Y_1^*}{Y_2^*} = \frac{k_2}{k_1} = 2$ always holds (Equation (4)). **B** A closed-loop architecture based on the open-loop network shown in Figure 3A and LC-Regulator. For the simulated response presented here the following parameters are used: $k_1 = 1$ min^{-1}, $k_2 = 3$ min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $\eta = 10$ nM^{-1} min^{-1}, $\theta_1 = 4$ nM min^{-1}, $\theta_2 = 5$ nM min^{-1}. The rest of the parameters (associated with the open-loop network) as well as the type of the disturbance (including the time of entry) remain the same as in **A**. Although the output steady states change due to the presence of the disturbance, $k_1 Y_1^* - k_2 Y_2^* = \theta_2 - \theta_1$ or $Y_1^* - 3Y_2^* = 1$ always holds (Equation (7)).

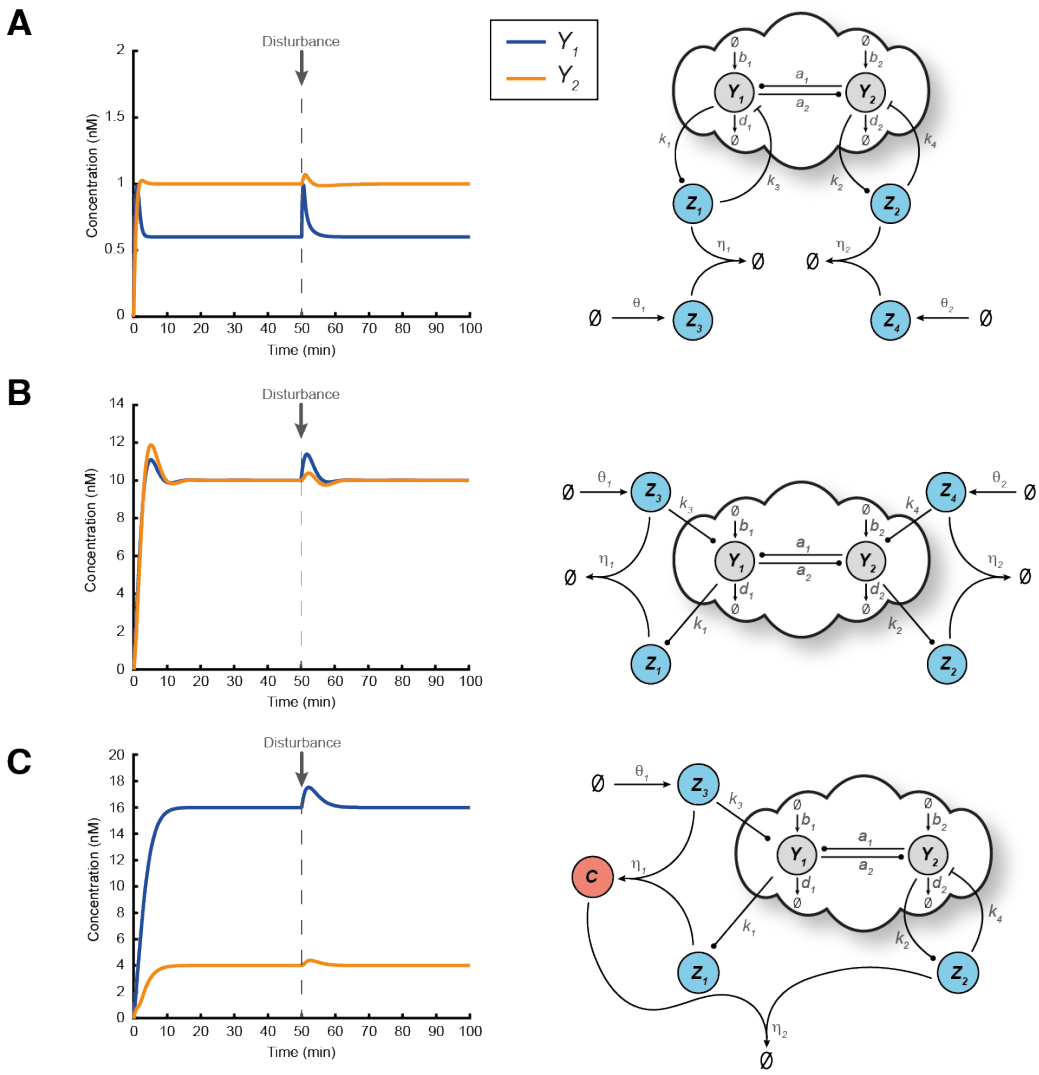

Figure 5: Regulating each output independently.
A A closed-loop architecture based on the open-loop network shown in Figure 3A and D-Regulator I. For the simulated response presented here the following parameters are used: $k_1 = 2.5$ min^{-1}, $k_2 = 0.5$ min^{-1}, $k_3 = 2$ nM^{-1} min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta_1 = \eta_2 = 10$ nM^{-1} min^{-1}, $\theta_1 = 1.5$ nM min^{-1}, $\theta_2 = 0.5$ nM min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. Despite the presence of a disturbance, $Y_1^* = \dfrac{\theta_1}{k_1} = 0.6$ nM, $Y_2^* = \dfrac{\theta_2}{k_2} = 1$ nM always hold (Equation (11)). **B** A closed-loop architecture based on the open-loop network shown in Figure 3A and D-Regulator II. For the simulated response presented here the following parameters are used: $k_1 = 1$ min^{-1}, $k_2 = 0.8$ min^{-1}, $k_3 = k_4 = 0.5$ min^{-1}, $\eta_1 = \eta_2 = 0.5$ nM^{-1} min^{-1}, $\theta_1 = 10$ nM min^{-1}, $\theta_2 = 8$ nM min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. Despite the presence of a disturbance, $Y_1^* = \dfrac{\theta_1}{k_1} = 10$ nM, $Y_2^* = \dfrac{\theta_2}{k_2} = 10$ nM always hold (Equation (11)). **C** A closed-loop architecture based on the open-loop network shown in Figure 3A and D-Regulator III. For the simulated response presented here the following parameters are used: $k_1 = 0.5$ min^{-1}, $k_2 = 2$ min^{-1}, $k_3 = 0.5$ min^{-1}, $k_4 = 2$ nM^{-1} min^{-1}, $\eta_1 = 0.5$ nM^{-1} min^{-1}, $\eta_2 = 10$ nM^{-1} min^{-1}, $\theta_1 = 8$ nM min^{-1} while the rest of the parameters (associated with the open-loop network) are the same as the ones used in Figure 3B. Despite the presence of a disturbance, $Y_1^* = \dfrac{\theta_1}{k_1} = 16$ nM, $Y_2^* = \dfrac{\theta_2}{k_1} = 4$ nM always hold (Equation (15)). The choice of the set points in **A**, **B** and **C** is arbitrary while the type of the disturbance (including the time of entry) is the same as in Figure 3B.

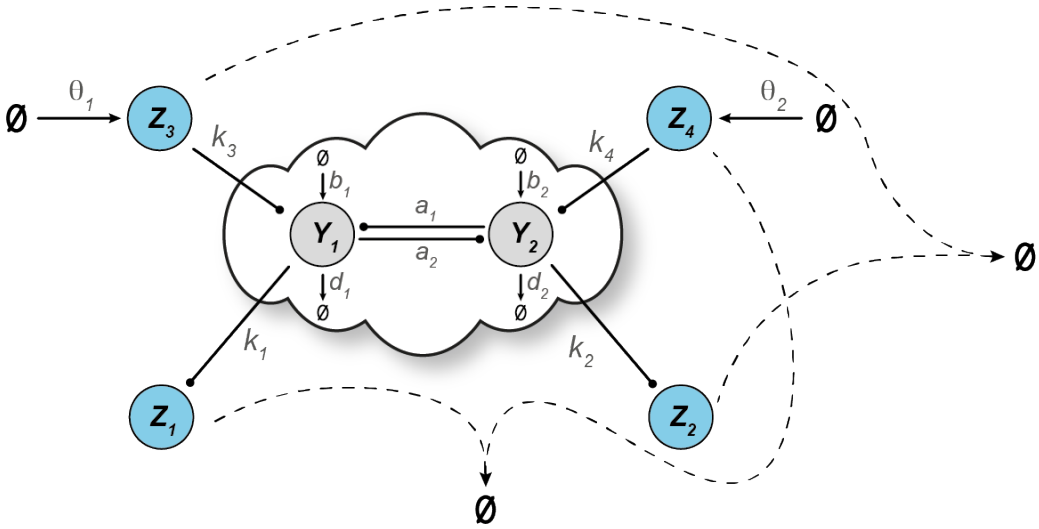

Figure 6: A different feedback configuration regarding the topology shown in Figure 5 b which leads to instability.

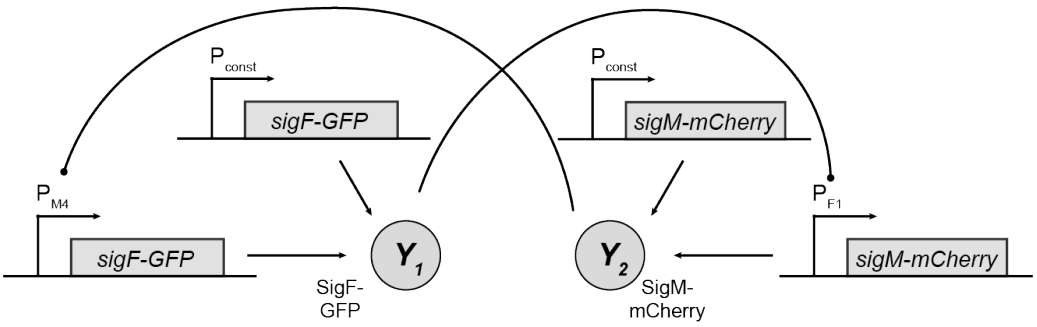

Figure 7: Experimental realization of the network to be controlled described by CRN (16).

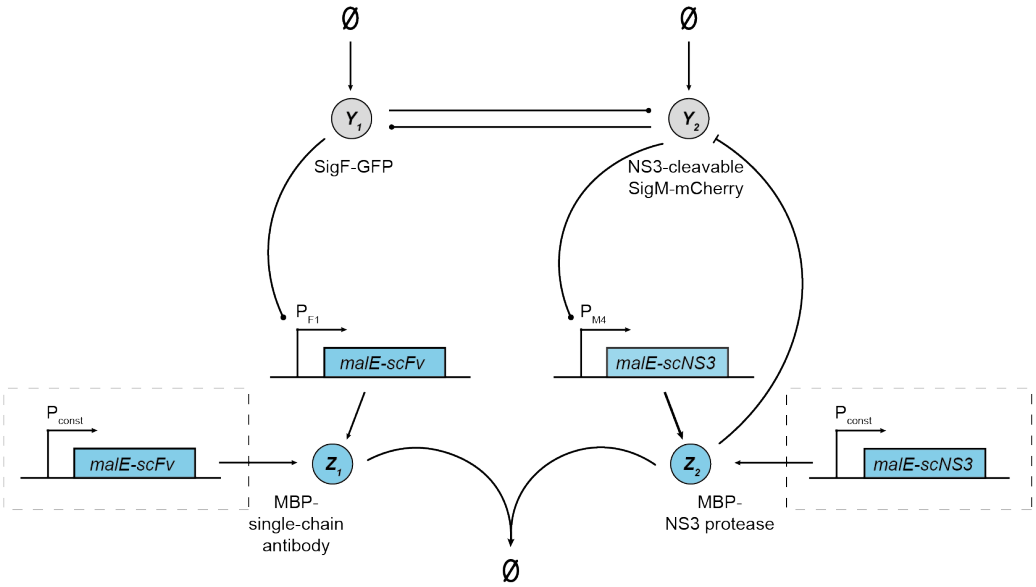

Figure 8: Experimental realization of the closed-loop architecture based on the open-loop network shown in Figure 6 and R-Regulator or LC-Regulator. The biological parts enclosed in dashed boxes are only required for LC-Regulator.

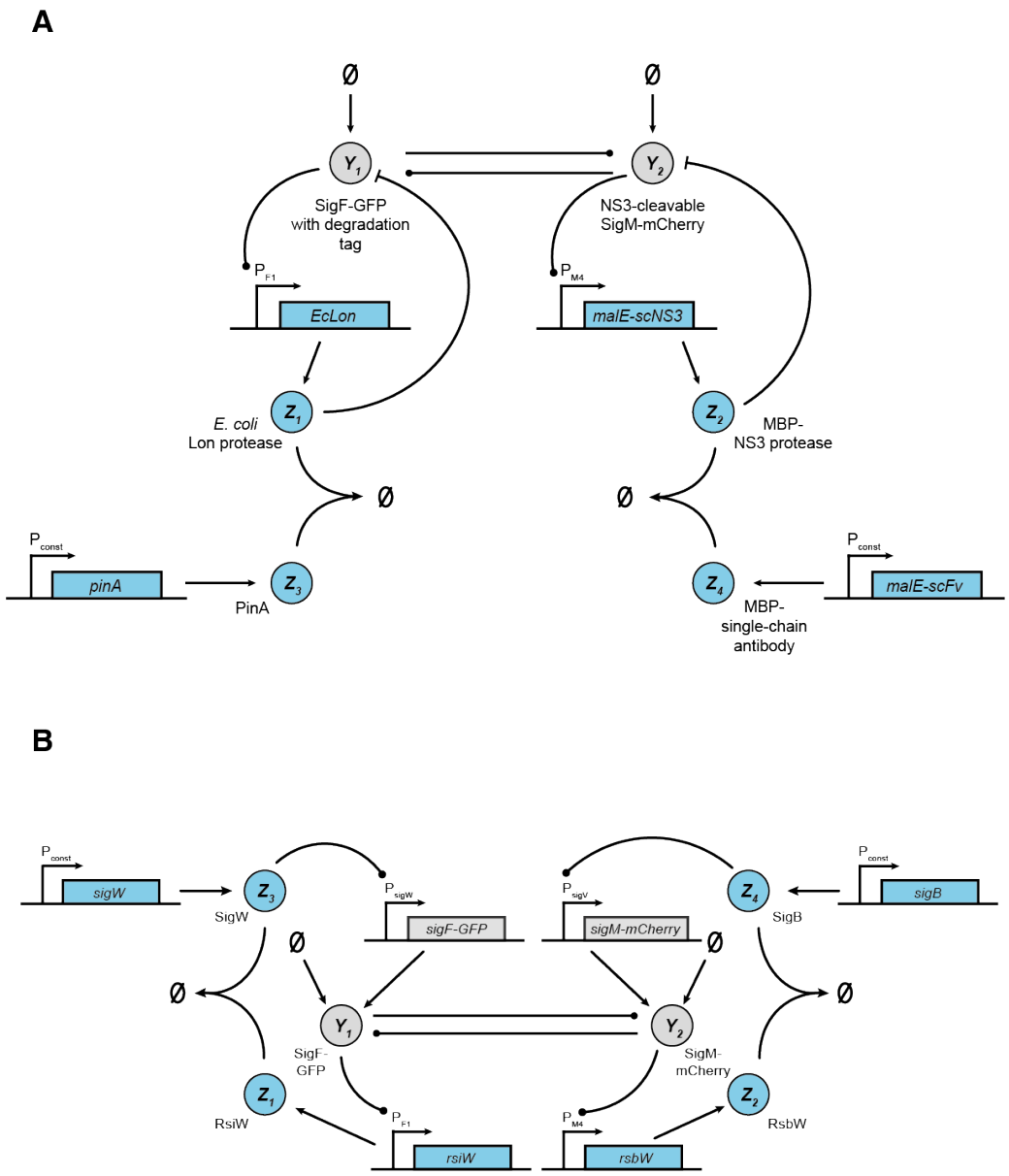

Figure 9: Experimental realization of the closed-loop architecture based on the open-loop network shown in Figure 6 and A D-Regulator I, B D-Regulator II.

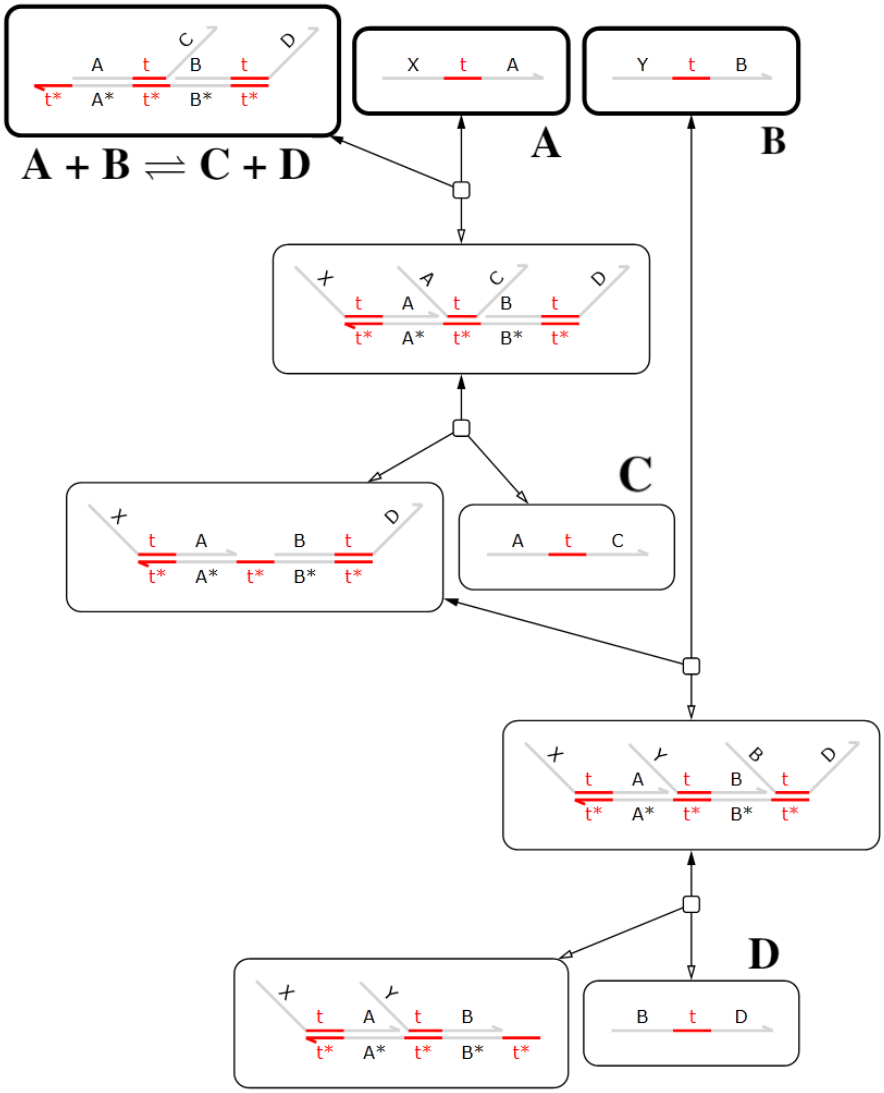

Figure 10: DNA strand displacement: representation of the reaction A + B ⇌ C + D. The initial DNA structures are indicated by a boldface border; reactions between DNA structures (small squares) have hollow heads for direct reactions and filled heads for reverse reactions. Each of the A, B, C, D abstract species is represented by a *3-domain*: a single-stranded DNA sequence logically subdivided into three *domains*, of which the middle one is short (red, ≈ 6 bases) and the others are long (black, ≈ 20 bases). Short domains are such that they bind reversibly to their Watson-Crick complements (indicated by *), while long domains bind irreversibly. A 3-domain is composed of a long *history* domain (left), which participated in past interactions (including X, Y) but does not affect future interactions. Next is a short *toehold* domain, which is used to initiate interactions between 3-domains and *gates* that implement the reactions. Next is a long *identity* domain that is the one that identifies the chemical species (right). A, B, C, D need not be distinct species. The same short sequence t can be used for all toehold occurrences, as successful bindings are determined by matching identity domains. A gate is a double-stranded DNA structure that includes backbone breaks on the top strand; when two breaks or strand-ends are in close proximity, they form an open (i.e., single-stranded) toeholds within the double-strand. A gate accepts 3-domains (the inputs to the reaction) that bind to its open toeholds, and through *strand displacement* releases other 3-domains (the outputs of the reaction). Strand displacement is a reversible random walk that starts at an open toehold and gradually replaces a domain with another identical domain within a double strand. At the end of the random walk, a whole single strand can detach from the double strand. The A + B ⇌ C + D reaction described above is reversible: the outputs can bind back to the gate through the open toehold on the right. However, it is easy to convert this to an irreversible A + B → C + D reaction by attaching a double stranded domain to the right of the gate (not shown), with an auxiliary single strand that irreversibly binds to the right toehold once it is exposed and to the new domain, preventing the outputs from binding back to the gate since no open toeholds are left. In summary, the species in a reaction networks can be uniquely assigned to domains (i.e., to specific sequences of nucleotides) and then a gate can be constructed for each desired reaction. The 3-domain structure is uniformly accepted and produced by the gates, so reactions can be composed.

References

[1] Drew Endy. "Foundations for engineering biology". In: *Nature* 438.7067 (2005), pp. 449–453.

[2] Ahmad S Khalil and James J Collins. "Synthetic biology: applications come of age". In: *Nature Reviews Genetics* 11.5 (2010), pp. 367–379.

[3] Warren C Ruder, Ting Lu, and James J Collins. "Synthetic biology moving into the clinic". In: *Science* 333.6047 (2011), pp. 1248–1252.

[4] George M Church et al. "Realizing the potential of synthetic biology". In: *Nature Reviews Molecular Cell Biology* 15.4 (2014), pp. 289–294.

[5] Fankang Meng and Tom Ellis. "The second decade of synthetic biology: 2010–2020". In: *Nature Communications* 11.1 (2020), pp. 1–4.

[6] Christopher A Voigt. "Synthetic biology 2020–2030: six commercially-available products that are changing our world". In: *Nature Communications* 11.1 (2020), pp. 1–6.

[7] Hana El-Samad. "Biological feedback control—Respect the loops". In: *Cell Systems* 12.6 (2021), pp. 477–487.

[8] Domitilla Del Vecchio and Richard M Murray. *Biomolecular feedback systems*. Princeton University Press Princeton, NJ, 2015.

[9] Domitilla Del Vecchio, Aaron J Dy, and Yili Qian. "Control theory meets synthetic biology". In: *Journal of The Royal Society Interface* 13.120 (2016), p. 20160380.

[10] Harrison Steel et al. "Challenges at the interface of control engineering and synthetic biology". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2017, pp. 1014–1023.

[11] Victoria Hsiao, Anandh Swaminathan, and Richard M Murray. "Control theory for synthetic biology: recent advances in system characterization, control design, and controller implementation for synthetic biology". In: *IEEE Control Systems Magazine* 38.3 (2018), pp. 32–62.

[12] Domitilla Del Vecchio et al. "Future systems and control research in synthetic biology". In: *Annual Reviews in Control* 45 (2018), pp. 5–17.

[13] Iacopo Ruolo et al. "Control engineering meets synthetic biology: Foundations and applications". In: *Current Opinion in Systems Biology* 28 (2021), p. 100397.

[14] Mustafa H Khammash. “Perfect adaptation in biology”. In: *Cell Systems* 12.6 (2021), pp. 509–521.

[15] Nika Shakiba et al. “Context-aware synthetic biology by controller design: Engineering the mammalian cell”. In: *Cell Systems* 12.6 (2021), pp. 561–592.

[16] Jinsu Kim and German Enciso. “Absolutely robust controllers for chemical reaction networks”. In: *Journal of the Royal Society Interface* 17.166 (2020), p. 20200031.

[17] Aivar Sootla et al. “Dichotomous Feedback: A Signal Sequestration-based Feedback Mechanism for Biocontroller Design”. In: *bioRxiv* (2021).

[18] Max Whitby et al. “PID control of biochemical reaction networks”. In: *IEEE Transactions on Automatic Control* (2021).

[19] Nuno MG Paulino et al. “PID and state feedback controllers using DNA strand displacement reactions”. In: *IEEE Control Systems Letters* 3.4 (2019), pp. 805–810.

[20] Sigurd Skogestad and Ian Postlethwaite. *Multivariable feedback control: analysis and design*. Vol. 2. Citeseer, 2007.

[21] Agostino Guarino, Davide Fiore, and Mario Di Bernardo. “In-silico feedback control of a MIMO synthetic toggle switch via pulse-width modulation”. In: *2019 18th European Control Conference (ECC)*. IEEE. 2019, pp. 680–685.

[22] Jean-Baptiste Lugagne et al. “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”. In: *Nature communications* 8.1 (2017), pp. 1–8.

[23] Benjamin H Weinberg et al. “Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells”. In: *Nature biotechnology* 35.5 (2017), pp. 453–462.

[24] Deepro Bonnerjee, Sayak Mukhopadhyay, and Sangram Bagh. “Design, fabrication, and device chemistry of a 3-input-3-output synthetic genetic combinatorial logic circuit with a 3-input AND gate in a single bacterial cell”. In: *Bioconjugate Chemistry* 30.12 (2019), pp. 3013–3020.

[25] Chelsea Y Hu and Richard M Murray. “Layered Feedback Control Overcomes Performance Trade-off in Synthetic Biomolecular Networks”. In: *bioRxiv* (2021).

[26] F Veronica Greco et al. “Harnessing the central dogma for stringent multi-level control of gene expression”. In: *Nature communications* 12.1 (2021), pp. 1–11.

[27] Alexander PS Darlington and Declan G Bates. “Architectures for Combined Transcriptional and Translational Resource Allocation Controllers”. In: *Cell Systems* 11.4 (2020), pp. 382–392.

[28] Corentin Briat, Ankit Gupta, and Mustafa Khammash. “Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks”. In: *Cell Systems* 2.1 (2016), pp. 15–26.

[29] Eduardo D Sontag. “Monotone and near-monotone biochemical networks”. In: *Systems and synthetic biology* 1.2 (2007), pp. 59–87.

[30] Alexander Y Mitrophanov and Eduardo A Groisman. “Positive feedback in cellular control systems”. In: *Bioessays* 30.6 (2008), pp. 542–555.

[31] Uri Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press, 2019.

[32] Indra Bervoets et al. “A sigma factor toolbox for orthogonal gene expression in Escherichia coli”. In: *Nucleic Acids Research* 46.4 (2018), pp. 2133–2144.

[33] Jason R. Kelly et al. “Measuring the activity of BioBrick promoters using an in vivo reference standard”. In: *Journal of Biological Engineering* 3.1 (Mar. 2009), p. 4.

[34] Limor Nahary, Alla Trahtenherts, and Itai Benhar. “Isolation of scFvs that Inhibit the NS3 Protease of Hepatitis C Virus by a Combination of Phage Display and a Bacterial Genetic Screen”. In: *Antibody Phage Display: Methods and Protocols*. Ed. by Robert Aitken. Totowa, NJ: Humana Press, 2009, pp. 115–132.

[35] Emmanouil Alexis et al. “Biomolecular mechanisms for signal differentiation”. In: *iScience* 24.12 (2021).

[36] David Chen and Adam P Arkin. “Sequestration-based bistability enables tuning of the switching boundaries and design of a latch”. In: *Molecular Systems Biology* 8 (2012), p. 620.

[37] Stephanie K. Aoki et al. “A universal biomolecular integral feedback controller for robust perfect adaptation”. In: *Nature* 570.7762 (2019), pp. 533–537.

[38] Bernard Yurke et al. “A DNA-fuelled molecular machine made of DNA”. In: *Nature* 406.6796 (2000), pp. 605–608.

[39] David Soloveichik, Georg Seelig, and Erik Winfree. “DNA as a universal substrate for chemical kinetics”. In: *Proceedings of the National Academy of Sciences* 107.12 (2010), pp. 5393–5398.

[40] Robert M Dirks et al. “Thermodynamic analysis of interacting nucleic acid strands”. In: *SIAM review* 49.1 (2007), pp. 65–88.

[41] Yuan-Jyue Chen et al. “Programmable chemical controllers made from DNA”. In: *Nature nanotechnology* 8.10 (2013), p. 755.

[42] Niranjan Srinivas et al. “Enzyme-free nucleic acid dynamical systems”. In: *Science* 358.6369 (2017).

[43] Kevin M Cherry and Lulu Qian. “Scaling up molecular pattern recognition with DNA-based winner-take-all neural networks”. In: *Nature* 559.7714 (2018), pp. 370–376.

[44] Avigdor Eldar and Michael B Elowitz. “Functional roles for noise in genetic circuits”. In: *Nature* 467.7312 (2010), pp. 167–173.

[45] Arjun Raj and Alexander Van Oudenaarden. “Nature, nurture, or chance: stochastic gene expression and its consequences”. In: *Cell* 135.2 (2008), pp. 216–226.

[46] Mads Kaern et al. “Stochasticity in gene expression: from theories to phenotypes”. In: *Nature Reviews Genetics* 6.6 (2005), pp. 451–464.

[47] Luca Laurenti et al. “Molecular filters for noise reduction”. In: *Biophysical Journal* 114.12 (2018), pp. 3000–3011.

[48] Corentin Briat, Ankit Gupta, and Mustafa Khammash. “Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks”. In: *Journal of The Royal Society Interface* 15.143 (2018), p. 20180079.

[49] Luca Cardelli, Marta Kwiatkowska, and Luca Laurenti. “Stochastic analysis of chemical reaction networks using linear noise approximation”. In: *Biosystems* 149 (2016), pp. 26–33.

[50] David J Warne, Ruth E Baker, and Matthew J Simpson. “Simulation and inference algorithms for stochastic biochemical reaction networks: from basic concepts to state-of-the-art”. In: *Journal of the Royal Society Interface* 16.151 (2019), p. 20180943.

[51] Yili Qian and Domitilla Del Vecchio. “Realizing ‘integral control’in living cells: how to overcome leaky integration due to dilution?” In: *Journal of The Royal Society Interface* 15.139 (2018), p. 20170902.