DISS. ETH NO. 28394

Adaptive mammalian gene circuits for robust regulation of gene expression and mitigation of cellular burden

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

Timothy Frei
M.Sc. ETH Zurich
born on 16.12.1990
citizen of Switzerland

accepted on the recommendation of

Prof. Dr. Mustafa Khammash (examiner)
Prof. Dr. Francesca Ceroni (co-examiner)
Dr. Christoph Zechner (co-examiner)

2022

Acknowledgments

First and foremost, I would like to express my gratitude to Mustafa Khammash for always believing in my abilities as a researcher and for letting me pursue my research interests as a PhD student in his group. His advisory style gave me the liberty to explore many of my own ideas, while still guiding my efforts when progress was difficult. I am truly grateful for the wonderful and interdisciplinary research group that he created and maintained over the years. It has been an exceptional environment to develop in, both personally and scientifically. I also want to extend many thanks to both Francesca Ceroni and Christoph Zechner for serving as members of my thesis committee and for taking the time to critically read and evaluate my thesis.

I owe a debt of gratitude to my collaborators Federica Cella, Joaquín Gutiérrez, Guy-Bart Stan and Velia Siciliano for their insights and support. I am especially grateful to Federica, whose endurance and ingenuity pushed the work demonstrated in Chapter 2 far beyond what we had initially envisioned. Similarly, I want to express my appreciation to Ching-Hsiang Chang and Maurice Filo for their collaboration and assistance. I want to especially thank Ching-Hsiang for his inventiveness and discipline, which fundamentally shaped the work demonstrated in Chapter 3.

I also want to thank all of the current and past members of the Control Theory and Systems Biology group, it has been an honor to be part of such an extraordinary group of people, thank you all. Much of the skills that lead to this thesis, I owe to the extraordinary training and advice I have had the blessing of receiving from Stephanie Aoki and Gabriele Lillacci. Without Stephanie's continued support and guidance, this thesis would not have been possible. Many thanks also go to my long-term lab and office mate Sara Dionisi, for the great chats and discussions. I also want to thank Maaike Welling for all her support regarding mammalian cell culture and for her feedback on an early draft of this thesis. I further owe thanks to Eline Bijman, Asterios Arampatzis and Tomas Kündig, which I had the privilege of supervising during their research projects / Master's theses. Moreover, I want to thank the D-BSSE single cell facility, especially Verena Jäggi, Telma Lopes, Aleksandra Gumienny and Mariangela Di

Tacchio for their expert support with flow cytometry and single-cell sorting.

A special thank you also goes to my friends from my studies at the BSSE, Noah Ritzman and Dominik Wolf. I still miss our lunch discussions that inspired many fun projects. Lastly, I am beyond grateful to all my friends and especially my family. Without their support throughout this period of my life, none of this would have been possible.

Abstract

Biological systems possess a remarkable ability to adapt to and withstand the dynamic environments they populate. This ability, often referred to as homeostasis, is realized by sophisticated regulatory machinery, frequently in the form of feedforward or negative feedback loops. In engineering, robustness to uncertainty in the system and its environment is often addressed using the principles of control theory. Due to the immense complexity and uncertainty faced when trying to engineer biological systems, synthetic biologists have started adopting these principles for genetic circuit engineering.

One such uncertainty in the engineering of genetic circuits is the strong context dependence of the genetic parts used. Often this leads to unexpected and unintended circuit behavior. An important instance of this context dependence is competition for shared limited resources. On the gene expression level, this manifests as a negative correlation between co-expressed genes, such that the expression level of one gene decreases as the expression of another gene increases. In Chapter 2, we describe the characterization of the competition for shared limited resources in the context of transient transfections in mammalian cells. We show how several common genetic regulatory strategies are affected and in particular focus on the effect of microRNA target number and location on an mRNA on resource allocation. Further, we describe a mathematical modeling framework that enables the extension of existing mathematical models to the resource-limited regime. Based on the collected insights and the mathematical framework, we implement both endogenous microRNA- and synthetic microRNA-based incoherent feedforward loops to mitigate the effects of changes in resource availability on a target gene.

In Chapter 3 we address context sensitivity of genetic circuits more generally by engineering genetic antithetic integral and proportional-integral feedback circuits in mammalian cells. We show that integral feedback reliably adapts to changes in the degradation rate of a regulated transcription factor expression at different expression levels. Furthermore, we demonstrate that the regulated transcription factor is insensitive to the interacting network structure by adding an additional negative feedback loop to its regulation. To further improve the variability in the regulated transcription factor expression, we augmented the antithetic integral feedback motif with proportional negative feedback to realize proportional-integral feedback. We show that in addition to adapting to increased degradation at different expression levels, the proportional-integral feedback circuit also reduces the variability in the expression of the regulated transcription factor. By changing the resource availability through the expression of additional genes, we show that both the integral and proportional-integral feedback circuits successfully mitigate the competition for shared and limited resources. Lastly, we suggest the applicability of our genetic proportional-integral feedback circuit in a simulation study of a cell therapy approach to type 1 and type 2 diabetes.

In this thesis, we have contributed towards a better understanding of context sensitivity of genetic circuits in synthetic biology by identifying resource limitations in mammalian cells and provide potential solutions to the problem via genetically engineered incoherent feedforward, integral and proportional-integral feedback loops. The contributions made in this thesis could help advance genetic circuit engineering in mammalian cells and could contribute to the development of robust biomedical therapies.

Zusammenfassung

Biologische Systeme besitzen eine bemerkenswerte Fähigkeit, sich an die dynamischen Umgebungen, die sie besiedeln, anzupassen und diesen zu widerstehen. Diese Fähigkeit, die oft als Homöostase bezeichnet wird, wird durch ausgeklügelte Regulationsmechanismen realisiert, häufig in Form von Forwärtskopplungs- oder negativen Rückkopplungsschleifen. In der Technik wird die Robustheit gegenüber Ungewissheiten im System und seiner Umgebung häufig anhand der Prinzipien der Regelungstheorie angegangen. Aufgrund der enormen Komplexität und Ungewissheit, mit der man konfrontiert wird, wenn man versucht, biologische Systeme zu konstruieren, haben synthetische Biologen damit begonnen, diese Prinzipien für die genetische Schaltungskreistechnik zu übernehmen.

Eine solche Unsicherheit bei der Konstruktion genetischer Schaltkreise ist die starke Kontextabhängigkeit der verwendeten genetischen Teile. Oft führt dies zu unerwartetem und unbeabsichtigtem Verhalten des Schaltkreises. Ein wichtiges Beispiel für diese Kontextabhängigkeit ist die Konkurrenz um gemeinsam genutzte begrenzte Ressourcen. Auf der Genexpressionsebene manifestiert sich dies als negative Korrelation zwischen koexprimierten Genen, so dass das Expressionsniveau eines Gens abnimmt, wenn die Expression eines anderen Gens zunimmt. In Kapitel 2 beschreiben wir die Charakterisierung der Konkurrenz um gemeinsame begrenzte Ressourcen im Zusammenhang mit transienten Transfektionen in Säugerzellen. Wir zeigen, wie mehrere gängige genetische Regulationsstrategien betroffen sind, und konzentrieren uns insbesondere auf die Auswirkung der Anzahl und Position von microRNA-Bindestellen auf einer mRNA auf die Ressourcenallokation. Darüber hinaus beschreiben wir einen mathematischen Modellierungsrahmen, der die Erweiterung bestehender mathematischer Modelle auf das ressourcenbegrenzte Regime ermöglicht. Basierend auf den gesammelten Erkenntnissen und dem mathematischen Rahmen implementieren wir sowohl endogene microRNA- als auch synthetische microRNA-basierte inkohärente Forwärtskopplungsschleifen, um die Auswirkungen von Änderungen der Ressourcenverfügbarkeit auf ein Zielgen zu mildern.

In Kapitel 3 befassen wir uns allgemeiner mit der Kontextsensitivität, indem wir genetisch antithetische Integral- und Proportional-Integral-Rückkopplungsschaltungen in Säugetierzellen entwickeln. Wir zeigen, dass sich das integrale Rückkopplung zuverlässig an Änderungen der Abbaurate eines regulierten Transkriptionsfaktors auf verschiedenen Expressionsniveaus anpasst. Darüber hinaus zeigen wir, dass der regulierte Transkriptionsfaktor unempfindlich gegenüber der interagierenden Netzwerkstruktur ist, indem wir seiner Regulation eine zusätzliche negative Rückkopplungsschleife hinzufügen. Um die Variabilität in der Expression des regulierten Transkriptionsfaktors weiter zu verbessern, haben wir das antithetische integrale Rückkopplungsmotiv mit proportional negativer Rückkopplung erweitert, um eine proportional-integrale Rückkopplung zu realisieren. Wir zeigen, dass der Proportional-Integral-Rückkopplungsschaltkreis zusätzlich zur Anpassung an den erhöhten Abbau bei verschiedenen Expressionsniveaus auch die Variabilität in der Expression des regulierten Transkriptionsfaktors reduziert. Durch die Änderung der Ressourcenverfügbarkeit durch die Expression eines zusätzlichen Gens zeigen wir, dass sowohl die integralen als auch die proportional-integralen Rückkopplungskreise die Konkurrenz um gemeinsam genutzte und begrenzte Ressourcen erfolgreich mildern. Schließlich schlagen wir die Anwendbarkeit unserer genetischen Proportional-Integral-Rückkopplungsschaltung in einer Simulationsstudie eines Zelltherapieansatzes für Typ-1- und Typ-2-Diabetes vor.

In dieser Dissertation haben wir zu einem besseren Verständnis der Kontextsensitivität in der synthetischen Biologie beigetragen, indem wir Ressourcenbeschränkungen in Säugetierzellen identifiziert haben und mögliche Lösungen für das Problem durch gentechnisch hergestellte inkohärente Forwärtskopplungs-, Integral- und Proportional-Integral-Rückkopplungsschleifen bereitstellen. Wir hoffen, dass die in dieser Dissertation geleisteten Beiträge dazu beitragen können, die genetische Schaltungstechnik in Säugetierzellen voranzutreiben und zur Entwicklung robuster biomedizinischer Therapien beizutragen.

Publications and Contributions

Chapter 2

Timothy Frei: Conceived the project, designed and performed experiments and performed data analysis of generic and transcriptional burden (HDV), developed the mathematical model, performed and analyzed iFFL experiments with synthetic miRNAs, wrote the manuscript and edited the manuscript.

Federica Cella: Conceived the project, designed and performed experiments and performed data analysis of endogenous miRNAs, RBPs, transcriptional and translational burden, endogenous genes' expression, performed and analyzed iFFL experiments with endogenous miRNAs, wrote the manuscript and edited the manuscript.

Fabiana Tedeschi: Performed and analyzed iFFL experiments with endogenous miRNAs and edited the manuscript.

Joaquín Gutiérrez: Developed the mathematical model and edited the manuscript.

Guy-Bart Stan: Conceived the project, supervised the computational work, wrote the manuscript and edited the manuscript.

Mustafa Khammash: Conceived the project, supervised the experimental work and secured funding, supervised the computational work, wrote the manuscript and edited the manuscript.

Velia Siciliano: Conceived the project, supervised the experimental work and secured funding, wrote the manuscript and edited the manuscript.

Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., & Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, *11*(1), 1–14

Chapter 3

Timothy Frei: Designed research, preformed research, performed experiments, analyzed data and wrote the paper.
Ching-Hsiang Chang: Designed research, preformed research, performed experiments, analyzed data and wrote the paper.
Maurice Filo: Designed research, preformed research, analyzed data, performed mathematical modeling and data fitting and wrote the paper.
Asterios Arampatzis: Performed research and wrote the paper.
Mustafa Khammash: Designed research, supervised research, secured funding and wrote the paper.

This chapter represents a manuscript submitted to the Proceedings of the National Academy of Sciences (PNAS). An earlier version was published as a preprint article on the bioRxiv preprint server under a Creative Commons Attributions-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0; https://creativecommons.org/licenses/by-nc-nd/4.0/)

Frei, T., Chang, C.-H., Filo, M., Arampatzis, A., & Khammash, M. (2021). Genetically engineered proportional-integral feedback controllers for robust perfect adaptation in mammalian cells. *bioRxiv*. https://doi.org/10.1101/2020.12.06.412304

Contents

List of Figures

List of Tables

Chapter One

Introduction

This chapter has been adapted from a short review published in Current Opinion in Systems Biology under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license (CC BY-NC-ND 4.0; https://creativecommons.org/licenses/by-nc-nd/4.0/).

Frei, T., & Khammash, M. (2021). Adaptive circuits in synthetic biology.
Current Opinion in Systems Biology, *28*, 100399

1.1 Adaptive circuits in synthetic biology

To survive, biological systems need to adapt to the ever-changing environment that they inhabit. This connection between adaptation and survival is considered a defining property that distinguishes the living from the non-living (Reece et al., 2014). Analogously, in manmade systems, engineers frequently design systems that adapt to unwanted perturbations or disturbances allowing them to robustly carry out their intended function. The properties that enable such adaptive behavior have been a subject of active research in the field of control theory. This shared interest in adaptation in natural and engineered systems is now increasingly being exploited in synthetic biology, where the theoretical insights from control theory are being used to design genetic circuits and genetically engineered cells that robustly operate in variable environments and without complete knowledge of the interactions that affect the system.

Adaptive synthetic biological systems hold great promise in the treatment of disease. In particular, they provide the ability to restore adaptation where it was lost due to disease. This is the case in illnesses that are directly attributed to the loss of homeostasis (Kotas & Medzhitov, 2015), an actively maintained constant internal state. Indeed, recent advances in synthetic biology and ge-

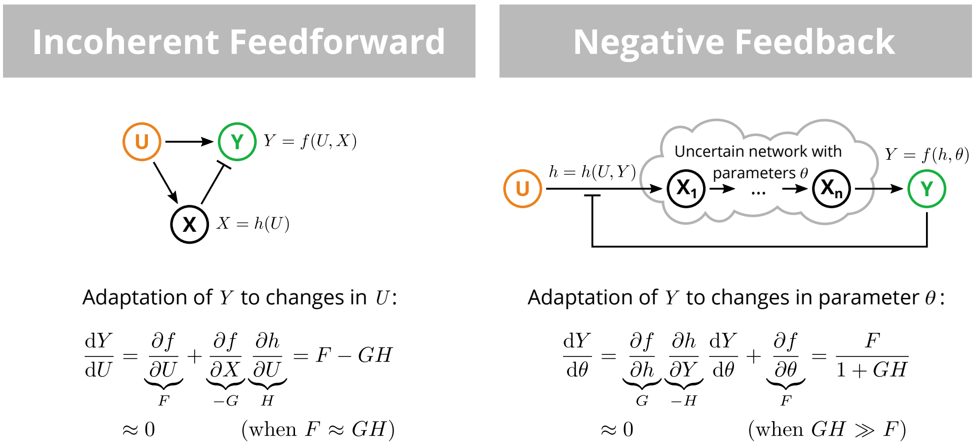


Figure 1.1: A mathematical description of adaptation. Adaptation can be assessed by looking at how the variable of interest Y in steady state changes with respect to what it adapts to. Mathematically, this can be expressed as the derivative of Y with respect to either the input U (incoherent feedforward loops) or the parameters θ (negative feedback) at steady state. The closer this expression is to zero the better the adaptation will be and it is considered perfect if either $\frac{\mathrm{d}Y}{\mathrm{d}U}$ or $\frac{\mathrm{d}Y}{\mathrm{d}\theta}$ are approximately equal to zero. For incoherent feedforward loops adaptation is improved by matching the terms F and GH which can be seen as making the direct and indirect effects of the input U on the variable of interest Y roughly the same. For negative feedback systems, adaptation improves with the term GH becoming larger than F, which is the same as increasing the strength of the feedback by either making the variable of interest Y have a greater effect on h ($H = \frac{\partial h}{\partial Y}$) or analogously by letting h have a greater effect on the variable of interest Y ($G = \frac{\partial f}{\partial h}$).

netic circuit engineering have since enabled the construction of sophisticated biological devices capable of partially restoring homeostasis (Kemmer et al., 2010; Rössger et al., 2013; Xie et al., 2016; Smole et al., 2017).

In the context of this thesis, we consider adaptation not in the evolutionary sense, but as a property of a network that results from suitable dynamic regulation leading to reduced sensitivity to external influences or system perturbations. Theoretical studies have identified several network motifs with characteristic interactions that are capable of producing this type of adaptive behavior (Ma et al., 2009). More recently, classes of network motifs that produce perfect adaptation have been described in the deterministic and stochastic setting (Xiao & Doyle, 2018; Aoki et al., 2019; Gupta & Khammash, 2022). Here, we examine two of the most important and widely prevalent adaptive motifs: incoherent feedforward loops, and negative feedback loops (Figure 1.1), focusing on their use in synthetic circuits.

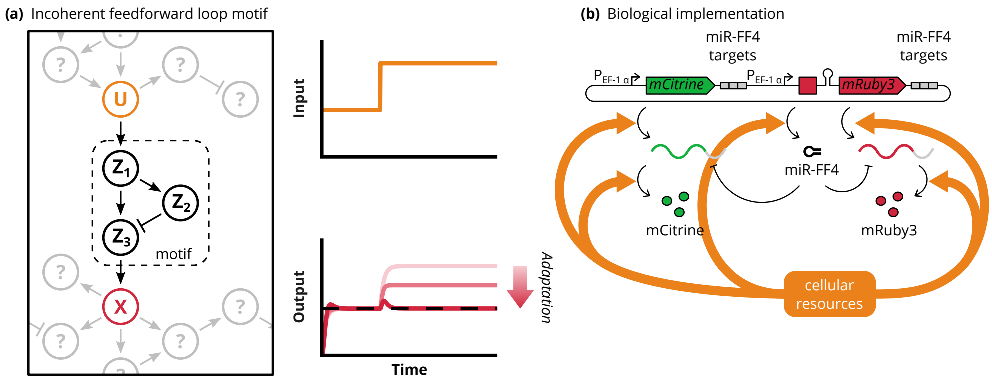


Figure 1.2: Incoherent feedforward loop motifs. (a) The network motif that composes an incoherent feedforward loop is shown embedded between an unknown upstream network and an unknown downstream network. The input species U is sensed through the production of species Z_1. The signal from species Z_1 is then passed on to species Z_2 and Z_3. Species Z_2 negatively affects the levels of species Z_3 to realize the incoherent aspect of the feedforward loop. The Z_3 species finally interacts with a downstream network trough the output species X. Shown on the right is a prototypical response of an incoherent feedforward loop to a step change in its input. Saturation of repression from Z_2 or if Z_3 is only removed when interacting with Z_2 will yield perfect adaptation. (b) A genetic circuit that can adapt to changes in resource availability through a microRNA-based incoherent feedforward loop. By sensing the amount of available resources in the expression levels of the synthetic microRNA FF4 the expression levels of two fluorescent proteins mCitrine and mRuby3 can be adapted to changes in resource allocation.

Incoherent feedforward loops

Discovered as a prototypical motif in biological signaling and natural gene regulatory networks (Mangan & Alon, 2003), the incoherent feedforward loop has become a recurring theme in synthetic biology due to its simple and diverse implementation opportunities. In one of its minimal realizations, the incoherent feedforward loop motif consists of three species (Figure 1.2a). Species Z_1 is produced in response to an input U and activates the production of two additional species Z_2 and Z_3. The species Z_2 then negatively affects the levels of the output species Z_3 to realize the incoherent feedforward component of the network. While negative feedback circuits can adapt to perturbations in the network structure or parameters and to certain exogenous inputs, the incoherent feedforward motif generally only adapts to changes in the input U. Viewed from the perspective of signal processing, the incoherent feedforward loop may also be viewed as a filter that removes unwanted dynamics from an incoming signal (Zechner et al., 2016). The level of adaptation depends on specific reaction rates in the motif (Ferrell Jr, 2016) and can be improved (Figure 1.2a, left panel bottom) through the choice of the constituent components (matching F

and *GH* as in Figure 1.1). Since its discovery, the incoherent feedforward motif has been realized in several synthetic gene circuits (Basu et al., 2004; Entus et al., 2007; Guo & Murray, 2019).

One successful application of the incoherent feedforward loop motif is in the adaptation to variation of gene copy numbers. This was first achieved in mammalian cells where intronically-expressed microRNAs were used to target the mRNA they were spliced from. Here, an incoherent feedforward loop was implemented that has the gene copy number as its input and hence adapts to its fluctuations (Bleris et al., 2011). Other studies have since used the same principle for similar purposes (Strovas et al., 2014; Lillacci et al., 2018; Frei et al., 2020). More recently, bacterial promoters have been engineered that are repressed by transcription-activator-like effectors (TALEs) (Segall-Shapiro et al., 2018). By constitutively expressing the TALE from the same plasmid as the repressible-promoter-driven output, it was shown that it can near-perfectly adapt to many different causes of plasmid copy number variability.

A multi-output incoherent feedforward loop was implemented using a transcriptional repressor in bacteria (Bondí et al., 2017). The incoherent feedforward motif has further also been extended to perform more complex computations (Lillacci et al., 2018; Lormeau et al., 2021; Benzinger et al., 2021).

Negative feedback loops

Another simple way to realize adaptive circuits is through negative feedback using either repressive or degradation-inducing interactions. The ongoing discovery of novel interactions within this category has lead to a wealth of simple synthetic negative feedback circuits (Shimoga et al., 2013; Bloom et al., 2015; Kelly et al., 2018; Guinn & Balázsi, 2019). The earliest reported synthetic instance of such a negative feedback system is in (Becskei & Serrano, 2000), where autoregulatory negative feedback was used to reduce the expression variability of the self-repressing fusion protein TetR-EGFP. This self-repressing realization has later been extended to linearize the non-linear dose-response curves inherent to inducible promoters in yeast (Nevozhay et al., 2009). Moreover, recently this has also been adapted to a pair of orthogonally and linearly inducible promoters in mammalian cells (Szenk et al., 2020). Another variation of this self-repression approach also appeared as a genetic circuit in mammalian cells (Stapleton et al., 2012). Here, the autoregulatory negative feedback was however chosen such that the repression acts on translation rather than transcription. Alternatively to implementing negative feedback on the transcriptional or translational level, a recent study used a de novo-designed protein pair to realize post-translational negative feedback (Ng et al., 2019).

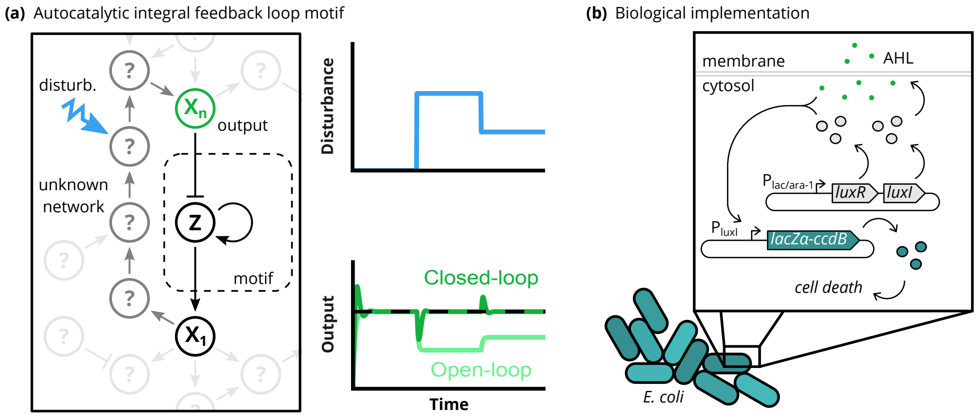


Figure 1.3: Autocatalytic feedback loops. (a) The autocatalytic feedback motif is shown embedded in some unknown network. The species *Z* catalyzes its own production and its levels are negatively affected by the output species of the unknown network X_n. Species *Z* interacts with the unknown network through species X_1. Note that to be a valid feedback loop there has to be a path from species X_1 to species X_n as is shown in dark gray. A disturbance to the dynamics of an arbitrary species in the unknown network is shown as a zigzagging blue arrow. On the right, we show how this motif is expected to react to disturbances in the unknown network. Here, we compare the closed-loop system — one with an embedded motif — to an open-loop system that lacks this motif. Adaptation is only achieved in the closed-loop system. (b) A simple genetic circuit that exploits cell division as the autocatalytic process and induced killing through quorum sensing implements autocatalytic feedback. Constitutive production of LuxI and LuxR produces and senses the small molecule AHL. AHL is secreted and by binding to LuxR induces the expression of the LacZα-ccdB killer gene.

While simple negative feedback circuits are well-understood and relatively easy to implement, they do not in general achieve *perfect* adaptation. As has been recognized in several studies (Shoval et al., 2011; Drengstig et al., 2012; Briat, Zechner, et al., 2016), autocatalysis or positive feedback implements what is sometimes referred to as constrained integral feedback (Xiao & Doyle, 2018). Integral feedback is a particular type of negative feedback that has been shown in control theory to guarantee perfect adaptation. In this motif, as is shown in Figure 1.3a, an unknown network interacts with a molecular species or cell type *Z* through node X_n in an inhibitory way. Furthermore, species *Z* catalyzes its own production and acts on the unknown network through node X_1. For a feedback loop to be closed, there has to be a path along which the levels of species X_1 influence the levels of species X_n. An alternative interpretation of species *Z* is as a cell type. A reason why such an interpretation is useful is the fact that proliferating cells naturally realize autocatalysis or positive feedback. Every cell makes a copy of itself as it divides. In fact, this motif might be at the heart of tissue size control (Lander et al., 2009; Buzi et al., 2015).

A synthetic instance of autocatalytic feedback has been realized in *E. coli* for population size control (You et al., 2004). Here, the authors used the production of the quorum sensing molecule acyl-homoserine-lactone (AHL), which is produced at a rate proportional to the number of cells, to induce the production of a killer gene through which cell density is downregulated (Figure 1.3b). This circuit can be seen as a realization of the autocatalytic feedback motif (Xiao & Doyle, 2018). Although the focus of the study (You et al., 2004) is population density control, an analysis of the genetic circuit as an autocatalytic feedback motif suggests that the controlled variable, X_n is in fact AHL. This observation is clarified by viewing the cells themselves as the autocatalytic entity. Moreover, it was observed that increasing the degradation rate of AHL in the medium by increasing medium pH leads to an increase in cell density, Z, which suggests that the system is compensating for an increased loss of AHL in the medium.

Antithetic integral feedback motif

More recently, a promising class of negative feedback circuits was described which are also capable of realizing integral feedback. The core motif has been termed antithetic integral feedback and can be implemented with two species (Briat, Gupta, et al., 2016). As is shown in Figs. 1.4 and 1.5a, these two species, Z_1 and Z_2 interact in such a way that they sequester each other's function. Furthermore, Z_2 is produced at a rate proportional to the controlled species X_n and acts as a sensor, while species Z_1 is constitutively produced and acts on the unknown network through X_1. A biological implementation of this motif was introduced in (Aoki et al., 2019; Lillacci et al., 2017) and has been shown to be capable of perfectly adapting to disturbances in the controlled variable as depicted in Figure 1.5b. To the best of our knowledge, this is the first genetic circuit to specifically implement integral feedback control in a living organism. A variant of this circuit has been realized in a cell-free system as well (Agrawal et al., 2019). An alternative implementation of the antithetic feedback motif in bacteria utilized small RNAs to realize quasi-integral control (Huang et al., 2018). It is known from theoretical analyses of the antithetic motif that substantial degradation or dilution of the two species in the motif leads to leaky integration. If substantial, this leakiness leads to a departure from the perfect adaptation, unless it is mediated by increased loop gains (Aoki et al., 2019; Qian & Del Vecchio, 2018). Since RNA species are generally short-lived in bacteria this makes RNA-based implementations in fast dividing cells challenging. While RNAs have a fast turnover rate in bacteria, in mammalian cells this is not generally the case (Milo & Phillips, 2015). This has made a recent implementation of the antithetic integral feedback motif in mammalian cells possible, where sense mRNA and anti-sense RNA sequester each other to inhibit the

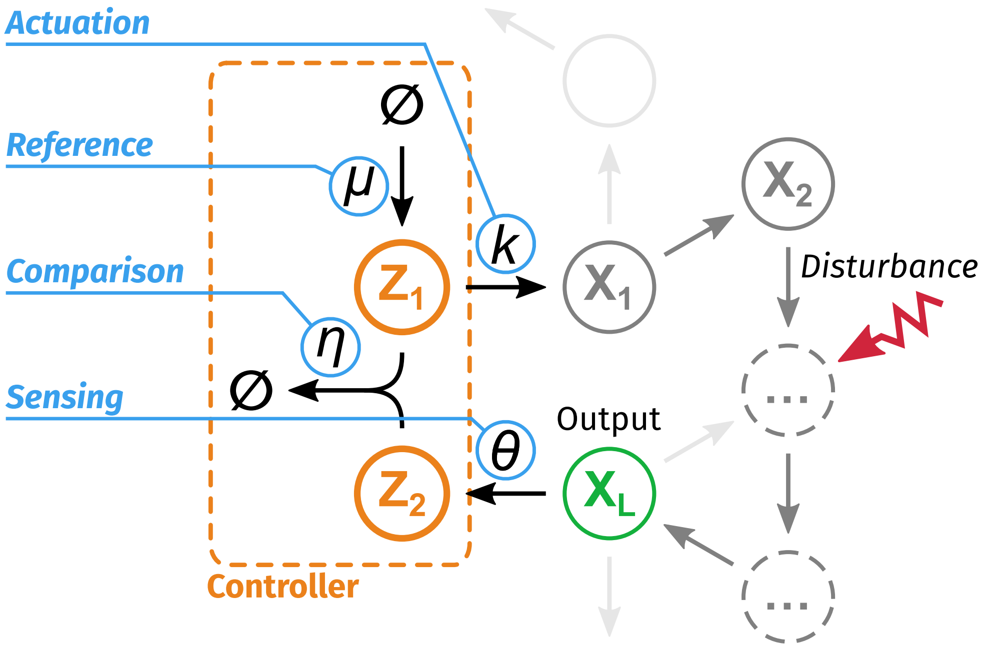
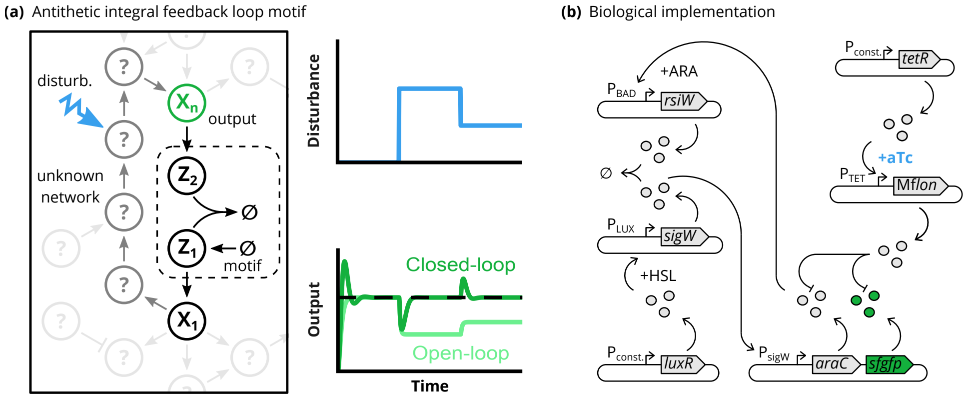


Figure 1.4: Schematic of the canonical antithetic integral feedback motif. The antithetic integral feedback motif is realized by two species, commonly referred to as Z_1 and Z_2. These two species together with four reactions constitutes the antithetic integral feedback motif. Each one of these reactions may be assigned a particular role in the motif. The reaction labeled *actuation* that produces species X_1 at a rate k enables the controller to exert influence on the controlled network and with that on the output species of interest X_n. The reaction that produces species Z_1 at the rate μ, labeled as *reference*, defines the steady-state expression of the output species X_n. In the reaction labeled *comparison* which occurs at a rate η, the expression levels of the two species are compared and the difference between the two levels is computed. Lastly, the levels of the output species X_n are sensed at a rate θ in the reaction labeled *sensing*.

translation of the mRNA and realize the antithetic motif (Frei et al., 2021).

1.2 Translating genetic circuit cartoons into ordinary differential equations

Mathematical modeling is starting to establish itself as an invaluable tool in synthetic biology. It accelerates both the design and troubleshooting of engineered genetic circuits by enabling the construction and testing of hypotheses without needing to run laborious experiments. Many introductory texts that teach mathematical modeling in the context of systems and synthetic biology exist, however these often approach the subject more theoretically than most practicing synthetic biologists are familiar with.



(a) Antithetic integral feedback loop motif

(b) Biological implementation

Figure 1.5: Antithetic integral feedback loops. (a) An antithetic motif is shown embedded in an unknown network. The state of the output species X_n is sensed in the levels of species Z_2. Species Z_1 is produced at a constant rate and reacts with species Z_2 through sequestration. Species Z_1 further interacts with the unknown network through species X_1. For the system to form a valid feedback loop, there has to be a path through which the levels of X_1 affect the levels of species X_n. This is illustrated in dark gray in the graphic. A disturbance to the dynamics of an arbitrary species in the unknown network is depicted as blue zigzagging arrow. The ideal response of the output species X_n to disturbances in the unknown network in closed-loop feedback with an antithetic integral feedback loop is shown on the right. The open-loop response of a system lacking antithetic feedback is shown as comparison. (b) The bacterial transcription factor AraC is sensed by the antithetic motif by activating the expression of the *rsiW* gene. The anti-σ factor RsiW sequesters the σ factor SigW, which is inducibly expressed through the LuxR system. SigW finally activates the expression of AraC to close the feedback loop. As a readout, the fluorescent protein sfGFP is also expressed from the same transcript as AraC. To introduce a targeted disturbance, the *Mf*Lon protease is inducibly expressed through the TetR system. Both AraC and sfGFP are fused to degradation tags which are recognized by the protease. Therefore, upon addition of aTc the expression of the *Mf*Lon protease is induced and the expression levels of AraC and sfGFP are perturbed.

Here, we will attempt to provide a recipe that describes the translation of genetic circuit cartoons into mathematical ordinary differential equation models. The proposed recipe consists of three steps, first, one is given or draws a cartoon of a genetic circuit based on a description or idea. Second, since there are usually many implicit assumptions about the specifics of the interactions depicted in these cartoons, a set of reactions and constraints that represent the aspects of the cartoon one wants to model are written down. Third and finally, the set of reactions is translated into a set of ordinary differential equations with the help of a lookup table.

The first step in the recipe is the most informal as many different depictions of genetic circuits are used in the synthetic biology literature. While there exist endeavors to standardize the representations of genetic circuits (Galdzicki et al.,

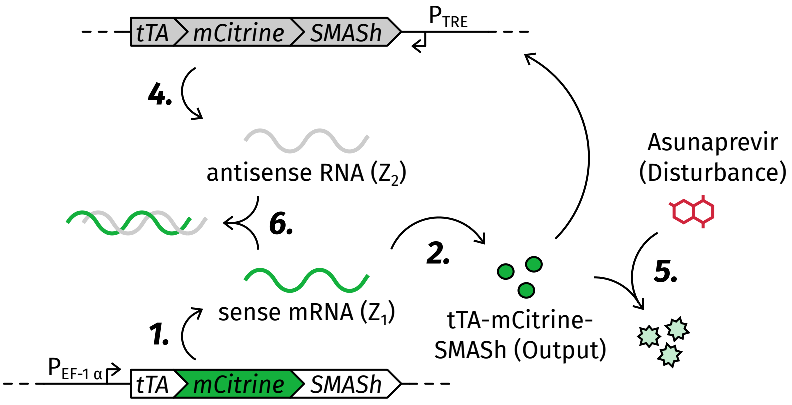


Figure 1.6: Circuit cartoon of sense/antisense implementation of the antithetic integral feedback motif. Details of the motif are elaborated in Chapter 3 of this thesis. Reaction number 3 from Table 1.1 is not depicted in the cartoon.

Table 1.1: Reactions of an implementation of the antithetic integral feedback motif.

Nr.	Reaction	Description
1	$\varnothing \xrightarrow{\mu} Z_1$	Production of Z_1 (reference in Fig. 1.4).
2	$Z_1 \xrightarrow{\mathrm{k}} Z_1 + \text{Out.}$	Z_1-dependent production of the output (actuation in Fig. 1.4).
3	$\text{Out.} \xrightarrow{\gamma} \varnothing$	Degradation/dilution of output.
4	$\text{Out.} \xrightarrow{\theta} \text{Out.} + Z_2$	Output-dependent production of Z_2 (sensing in Fig. 1.4).
5	$\text{Out.} + \text{Dist.} \xrightarrow{\omega} \text{Dist.}$	Output- and disturbance-dependent degradation of the output.
6	$Z_1 + Z_2 \xrightarrow{\eta} \varnothing$	Z_1- and Z_2-dependent sequestration of Z_1 and Z_2 (comparison in Fig. 1.4).

2014; Baig et al., 2021), these are often limited in their expressiveness and do not capture the wealth of interactions possible in biology. Chemical reaction network theory provides a rigorous framework in which most biomolecular circuits may be represented (Gunawardena, 2003). However, most depictions of genetic circuits do not adhere to such a rigorous framework, although there does appear to be some convergence to a limited set of common syntax choices for cartoon representations of genetic circuits.

To illustrate the recipe, we will use an implementation of the antithetic integral feedback motif as an example of which we will derive a simple ordinary differential equation model. The cartoon of the circuit is depicted in Fig. 1.6 and in this case already concludes step 1. For step two, we will elaborate on the assumptions not explicitly made in the cartoon that lead to the set of reactions collected in Table 1.1. To derive a useful model in terms of ordinary differential equations, we generally assume that all biomolecular species are well mixed and in sufficiently high copy numbers such that all molecules belonging to the same species may be treated equally. Since the production of the sense mRNA (labeled as reaction 1 in Fig. 1.6), which we will refer to as species Z_1, does not depend on any species that has time-dependent amounts we model it as zeroth order reaction. The Z_1-dependent production of tTA-mCitrine-SMASh (reaction 2 in Fig. 1.6), referred to as *Output*, however depends on the amounts of the species Z_1 present in the system. Since the species Z_1 is not consumed in the reaction, we model it as a first order catalytic production reaction. Although not depicted explicitly in Fig. 1.6, we commonly assume that all molecular species within a cell are either actively degraded or dilute over time due to cell growth and division. We therefore include a reaction that models the removal of the Output species. This reaction falls into the category of first order reactions, since the rate of removal depends on the amount of the output species present in the system. One may notice that we have not assumed any degradation or dilution of the two controller species Z_1 and Z_2. This is due to the fact that the antithetic integral feedback motif specifically requires that the two controller species are only removed by sequestration. More details on this are provided in Section 1.4 and Chapter 3. The production of the antisense RNA, referred to as Z_2, depends on the amount of Output species. However, the Output species does not get consumed in the reaction and we therefore model it as a first order catalytic reaction. In the depicted circuit the disturbance used to test for perfect adaptation is shown as the interaction of the Output species with the small molecule Asunaprevir, which we will call the *Disturbance*. Since this reaction depends both on the amount of Output species and Disturbance species we model it as a second order reaction that only has the Disturbance species as a product, due to the induced degradation of the Output species. Finally and arguably most importantly for the antithetic integral feedback motif, we model the interaction of the two controller species Z_1 and Z_2 as a sequestration where the reaction rate depends on both species. Therefore, this is a again a second order reaction. However, this time the reaction does not have a product since both species remove each other's function.

Once all the reactions one wants to model have been identified and listed, similar to what has been done in Table 1.1, one can use the lookup table Table 1.2 to construct a system of ordinary differential equations that model the time evolution of the concentrations of the individual species. To do so, the reactions

Table 1.2: Reaction lookup table.

	Nr.	**Reaction**	**Term**	
			Substrate	Product
0th	1	$\varnothing \xrightarrow{\mathrm{k}} \mathrm{A}$		$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = k$
1st	3	$\mathrm{A} \xrightarrow{\mathrm{k}} \varnothing$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -kA(t)$	
		$\mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{B}$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -kA(t)$	$\frac{\mathrm{d}B(t)}{\mathrm{d}t} = kA(t)$
		$\mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{A} + \mathrm{A}$		$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = kA(t)$
	2, 4	$\mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{A} + \mathrm{B}$		$\frac{\mathrm{d}B(t)}{\mathrm{d}t} = kA(t)$
		$\mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{B} + \mathrm{C}$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -kA(t)$	$\frac{\mathrm{d}B(t)}{\mathrm{d}t} = \frac{\mathrm{d}C(t)}{\mathrm{d}t} = kA(t)$
2nd	6	$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{k}} \varnothing$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = \frac{\mathrm{d}B(t)}{\mathrm{d}t} = -kA(t)B(t)$	
		$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{k}} \mathrm{A}$	$\frac{\mathrm{d}B(t)}{\mathrm{d}t} = -kA(t)B(t)$	
	5	$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{k}} \mathrm{B}$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -kA(t)B(t)$	
		$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{k}} \mathrm{C}$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = \frac{\mathrm{d}B(t)}{\mathrm{d}t} = -kA(t)B(t)$	$\frac{\mathrm{d}C(t)}{\mathrm{d}t} = kA(t)B(t)$
		$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{k}} \mathrm{A} + \mathrm{B} + \mathrm{C}$		$\frac{\mathrm{d}C(t)}{\mathrm{d}t} = kA(t)B(t)$
		$\mathrm{A} + \mathrm{A} \xrightarrow{\mathrm{k}} \varnothing$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -2kA(t)^2$	
		$\mathrm{A} + \mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{A}$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -kA(t)^2$	
		$\mathrm{A} + \mathrm{A} \xrightarrow{\mathrm{k}} \mathrm{B}$	$\frac{\mathrm{d}A(t)}{\mathrm{d}t} = -2kA(t)^2$	$\frac{\mathrm{d}B(t)}{\mathrm{d}t} = kA(t)^2$

in Table 1.1 are identified with the equivalent reactions in the lookup table and the corresponding terms are added to get a differential equation for every species. This is illustrated in Eq. (1.1).

$$\begin{aligned}
\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \left(\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t}\right)_{P1} + \left(\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t}\right)_{S6} \\
\frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \left(\frac{\mathrm{d}Z_2(t)}{\mathrm{d}t}\right)_{P4} + \left(\frac{\mathrm{d}Z_2(t)}{\mathrm{d}t}\right)_{S6} \\
\frac{\mathrm{d}X(t)}{\mathrm{d}t} &= \left(\frac{\mathrm{d}X(t)}{\mathrm{d}t}\right)_{P2} + \left(\frac{\mathrm{d}X(t)}{\mathrm{d}t}\right)_{S3} + \left(\frac{\mathrm{d}X(t)}{\mathrm{d}t}\right)_{S5}
\end{aligned} \tag{1.1}$$

Here, indices are used to indicate if the species is either a substrate (S) or product (P) and is combined with the reaction number that relates it to the reactions in Table 1.1. Once all the terms have been identified, we arrive at the model shown in Eq. (1.2).

$$\begin{aligned}
\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu - \eta Z_1(t)Z_2(t) \\
\frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \theta X(t) - \eta Z_1(t)Z_2(t) \\
\frac{\mathrm{d}X(t)}{\mathrm{d}t} &= kZ_1(t) - \gamma X(t) - \omega A(t)X(t)
\end{aligned} \tag{1.2}$$

In the above equations, we use $X(t)$ for the output and $A(t)$ for the disturbance.

General antithetic integral feedback motif model

To arrive at a more general model of the antithetic integral feedback controller, we allow the controlled network to be of an arbitrary size. For this we replace $X(t)$ in Eq. (1.2) with the state vector of the controlled network $\mathbf{X}(t)$ of length n corresponding to the number of species in the network. The networks dynamics are given by $g(\mathbf{X}(t))$ and we actuate on the system through species X_1. This is represented by the term $e_1kZ_1(t)$, where $e_1 = [1, 0, \dots, 0]^\mathrm{T}$ of length n and k is the reaction rate of the actuation reaction. We assume that the last species X_n in the network is sensed and controlled robustly. The full system of equations reads:

$$\begin{aligned}
\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu - \eta Z_1(t)Z_2(t) \\
\frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \theta X_n(t) - \eta Z_1(t)Z_2(t) \\
\frac{\mathrm{d}\mathbf{X}(t)}{\mathrm{d}t} &= g(\mathbf{X}(t)) + e_1kZ_1(t).
\end{aligned} \tag{1.3}$$

1.3 Resource limitations

Work in bacterial synthetic biology (Ceroni et al., 2015; Qian et al., 2017) has identified competition for shared limited resources as a source of unexpected behavior when engineering genetic circuits. In these studies, it was postulated that resources required for gene expression, such as RNA polymerases and ribosomes, exist in limited and conserved quantities within a cell. Due to this, genes will compete for these resources when they are expressed and indirectly affect each others expression levels. More specifically, this leads to the expression levels of genes being anticorrelated, an increase in expression of one gene leads to a decrease in expression of all other genes that share a common pool of limited resources. The details of this phenomena are elaborated in Chapter 2 of this thesis.

To mitigate this effect, incoherent feedforward loops have been applied. Two recent studies (one of which is reproduced as Chapter 2 in this thesis) apply different incoherent feedforward loop implementations in mammalian cells to adapt gene expression to changes in cellular resource allocation. In one, an endoribonuclease that cleaves RNA upon recognition of a particular secondary structure — a so-called hairpin — was used to knock down the translation of an mRNA that encoded said hairpin. By placing both the gene for the endoribonuclease and the hairpin-encoding mRNA on the same plasmid, an incoherent feedforward motif was constructed (Jones et al., 2020). In another study, the microRNA-based incoherent feedback loop from (Lillacci et al., 2018) was modified to adapt multiple genes to variations in resource availability (Figure 1.2b) (Frei et al., 2020,Chapter 2). Furthermore, this work implemented a minimal incoherent feedforward motif capable of mitigating resource burden by exploiting endogenous microRNAs.

Negative feedback circuits have also been used to address issues that arise from competition for limited and shared cellular resources. To confront this issue in bacteria, a promoter that natively activates under burdened conditions was used to expresses a guide RNA that together with the dCas9 protein adaptively represses the expression of other synthetic constructs in the cell (Ceroni et al., 2018). In another recent study, self-repression on the promoter level by a dCas9/guide RNA complex was utilized to eliminate competition for the shared dCas9 protein in CRISPRi circuits in bacteria (Huang et al., 2021).

1.4 Limitations of antithetic integral feedback control

While its simple structure makes the antithetic feedback motif an attractive topology to implement genetically, some technicalities related to specific choices about its implementation may demand a deviation from the conditions required for achieving robust perfect adaptation. In this section, we will briefly describe some of these technicalities. In particular, we will discuss the effects of resource dependence, the effects of saturation of either the sensing or actuation reactions and the effect of degradation or dilution of the controller species on the robustness of the set point.

Resource dependence

As briefly introduced in Section 1.3 and elaborated on more in Chapter 2, competition for limited resources induces an indirect coupling of expressed genes. Most and especially high copy number genetic implementations of the antithetic integral feedback motif will also be affected by this phenomena and its consequences may directly influence the robustness properties of the output species. We will illustrate this by extending the model of the canonical antithetic integral feedback motif introduced in Section 1.2 with resource limited production reactions.

$$\begin{aligned}\frac{\mathrm{d}C_{Z_1}(t)}{\mathrm{d}t} &= \alpha_{Z_1} R_i Y_m(t) - (\beta_{Z_1} + \mu) C_{Z_1}(t)\\ \frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu C_{Z_1}(t) - \eta Z_1(t) Z_2(t)\\ \frac{\mathrm{d}C_{Z_2}(t)}{\mathrm{d}t} &= \alpha_{Z_2} R_j X_n(t) - (\beta_{Z_2} + \theta) C_{Z_2}(t)\\ \frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \theta C_{Z_2}(t) - \eta Z_1(t) Z_2(t)\end{aligned} \tag{1.4}$$

In Eq. (1.4) we have introduced the two additional species $\mathrm{C}_{\mathrm{Z}_1}$ and $\mathrm{C}_{\mathrm{Z}_2}$ to denote the complex formed by binding of the substrate species with a resource. We have further also introduced a substrate species Y_m for the controller species Z_1. The set point for the output species X_n is now no longer $\overline{X}_m = \frac{\mu}{\theta}$ but also includes the steady-state of the free resource species R_i, R_j and the substrate species Y_m. Explicitly, the new set point is

$$\overline{X}_n = \frac{\mu}{\theta}\frac{\overline{C}_{Z_1}}{\overline{C}_{Z_2}} = \frac{\mu}{\theta}\frac{\zeta_{Z_1}^{-1}}{\zeta_{Z_2}^{-1}}\frac{\overline{R}_i}{\overline{R}_j}\overline{Y}_m. \tag{1.5}$$

Here, we have lumped the parameters regarding the complexes as $\zeta_{Z_1} := (\beta_{Z_1} + \mu)/\alpha_{Z_1}$ and $\zeta_{Z_2} := (\beta_{Z_2} + \theta)/\alpha_{Z_2}$. In the expression shown in Eq. (1.5), the steady states of the species R_i, R_j and Y_m may depend on parameters of the unknown network and would therefore not be robust to their variation. In particular, the steady state of the resource species R_i and R_j will depend on the steady states of all the species whose production they catalyze. That way, shared resources couple the expression of the controller species Z_1 and Z_2 to the species in the unknown network they are controlling. Furthermore, the substrate species Y_m will often also depend on the steady-state level of a shared resources, further complicating the situation.

To set up a framework for assessing how robustness may be recovered for resource burdened systems, we assume an order on the resource pools considered. For example, following the central dogma of molecular biology, we may assume a general pool of transcriptional $\mathrm{R_{txs}}$ and a pool of translational resources $\mathrm{R_{tsl}}$, we order them as $\emptyset \to \mathrm{R_{txs}} \to \mathrm{R_{tsl}} \to \mathrm{R_{txs}} \to \dots$. The consequence of this order is that species created by resources from the transcriptional pool will catalyze the creation of other species together with resources from the translational pool. We also assume the order to be circular, which may be interpreted in the example as transcription factors, produced by translational resources, catalyzing the production of mRNA with the help of transcriptional resources. To account for an arbitrary number of resource pools, we define the order $\emptyset \to \mathrm{R_1} \to \dots \to \mathrm{R_n} \to \mathrm{R_k} \to \dots$ with $\mathrm{k} \in \{1, \dots, n-1\}$ and where every $\mathrm{R_i}$ with $\mathrm{i} \in \{1, \dots, \mathrm{n}\}$ is a distinct set of resources. In this order, we allow for resource pools that precede the loop by allowing $\mathrm{R_k}$ to be any resource pool except for the last resource pool $\mathrm{R_n}$.

Assuming sufficient freedom in the choice of implementation robustness may be recovered. This may be done by choosing the controller species such that they are produced by the initial pool of resources. In that case, we have $R_i = R_j = R_1$ and the fraction in Eq. (1.5) will cancel. Since we chose the resource to be the first in the order there is also no resource affecting the species $\mathrm{Y_m}$ and the set point will now be decoupled from changes in resource availability. This gives the steady state for $\mathrm{X_n}$ as:

$$\overline{X}_n = \frac{\mu}{\theta} \frac{\zeta_{Z_1}^{-1}}{\zeta_{Z_2}^{-1}} \sigma_m. \tag{1.6}$$

Here, we introduce σ_m to indicate the steady state of $\mathrm{Y_m}$ that does not depend on any resource pool.

Sensing saturation

The gene expression response of an inducible promoter often has a sigmoidal shape. This can be partially explained by both the cooperative nature of transcription factor binding and limited binding capacity of the promoters. Specifically the latter property leads to the saturation of the gene expression response. Under saturation the change in expression is effectively zero even when more transcription factor is being produced. Early work in enzyme kinetics has derived mathematical expressions that model these sigmoidal response curves. Most famously, work by Hill, and Michaelis and Menten (Hill, 1910; Michaelis, Menten, et al., 1913) has produced compact expression commonly used to model saturation in biomolecular reactions.

$$f(x) = \alpha + (1-\alpha)\frac{x^h}{\kappa^h + x^h} \tag{1.7}$$

In this equation we have also introduced a parameter α to model basal expression that is independent of the amount of x. Specifically, we use the constraint $0 \leq \alpha \leq 1$ to vary the amount of basal expression. In the extremes, if we let $\alpha = 0$, then there would be no basal expression at all, whereas if we let $\alpha = 1$ then all expression would be basal. The exponent h can be interpreted as a measure of cooperativity and the parameter κ may be derived akin to a Michaelis-Menten constant as a fraction of the sum of the forward and the catalytic reactions over the backward reaction.

To potentially capture more realistic implementations of the antithetic integral feedback motif, we replace the sensing reaction term in Eq. (1.2) with a more general production reaction.

$$\begin{aligned} \frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu - \eta Z_1(t)Z_2(t) \\ \frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \hat{\theta} f(X_n(t)) - \eta Z_1(t)Z_2(t) \end{aligned} \tag{1.8}$$

Specifically, we have replaced the $\theta X_n(t)$ term with $\hat{\theta} f(X_n(t))$ in Eq. (1.8) to then arrive at an expression for the steady state of species $\mathrm{X_n}$ with a more general sensing reaction. We introduce $\hat{\theta}$ since in the formulation in Eq. (1.8) θ no longer has units $1/[Time]$ but now has units $[Molecules]/[Time]$. The steady-state expression in Eq. (1.9) will hold for any invertible f.

$$\overline{X}_n = f^{-1}\left(\frac{\mu}{\hat{\theta}}\right) \tag{1.9}$$

If we then consider the case of a hill-type activation function as was introduced in Eq. (1.7), then the inverse function is

$$f^{-1}(x) = \kappa\left(\frac{x-\alpha}{1-x}\right)^{1/h}. \tag{1.10}$$

Plugging the inverse function from Eq. (1.10) into the steady-state expression for $\mathrm{X_n}$ in Eq. (1.9), we get

$$\overline{X}_n = \kappa\left(\frac{\frac{\mu}{\hat{\theta}} - \alpha}{1 - \frac{\mu}{\hat{\theta}}}\right)^{1/h}. \tag{1.11}$$

Given that we require that $\overline{X}_n \in \mathbb{R}_{\geq 0}$, we find that the expression in Eq. (1.11) provides both an upper and a lower bound on the set point. The bounds can be deduced from the inequality in Eq. (1.12).

$$\kappa\left(\frac{\frac{\mu}{\hat{\theta}}-\alpha}{1-\frac{\mu}{\hat{\theta}}}\right)^{1/h} \geq 0 \tag{1.12}$$

Reducing the above expression, we find that $\alpha \leq \frac{\mu}{\hat{\theta}} < 1$. Interpreting these bounds, we find that to obtain an admissible set point we want to chose it such that it exceeds the basal fraction of the sensing reaction and remains below 1.

Actuation saturation

Similar to sensing saturation, actuation saturation puts bounds on parameters for which robust perfect adaptation remains possible. To illustrate this, we adapt Eq. (1.3) by changing the term $e_1 k Z_1(t)$ to $e_1 \hat{k} f(Z_1(t))$. Again, in this formulation the units of k change from $1/[Time]$ to $[Molecules]/[Time]$ and we replace it by $\hat{k}$ to reflect this. This gives the following system of equations:

$$\begin{aligned}\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu - \eta Z_1(t) Z_2(t)\\ \frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \theta X_n(t) - \eta Z_1(t) Z_2(t)\\ \frac{\mathrm{d}\mathbf{X}(t)}{\mathrm{d}t} &= g(\mathbf{X}(t)) + e_1 \hat{k} f(Z_1(t)).\end{aligned} \tag{1.13}$$

To see how the saturation of the actuation reaction imposes bounds on the parameters, we solve for the steady state of species Z_1. This gives

$$\overline{Z}_1 = f^{-1}(-\frac{1}{\hat{k}} g_1(\overline{\mathbf{X}})). \tag{1.14}$$

To go from the more general solution obtained in Eq. (1.14) to a more particular form, we assume the same f as previously introduced in Section 1.4 and Eq. (1.7). Since we require that $\overline{Z}_1 \in \mathbb{R}_{\geq 0}$, we again obtain an inequality

$$\kappa\left(\frac{-\frac{1}{\hat{k}} g_1(\overline{\mathbf{X}}) - \alpha}{1 + \frac{1}{\hat{k}} g_1(\overline{\mathbf{X}})}\right)^{1/h} \geq 0. \tag{1.15}$$

Reducing this inequality gives the constraints $\alpha \leq -\frac{1}{\hat{k}} g_1(\overline{\mathbf{X}}) < 1$. Unfortunately, this formulation requires knowledge of the function g to rationally select parameters such that the constraints hold. Furthermore, generally when engineering a biological system one cannot manipulate the parameters in the controlled network. However, since we may have the ability to change the actuation reaction rate $\hat{k}$, one may attempt to satisfy these bounds by experimental trial and error. Additionally, the steady state of the controlled network will depend on the set point $\frac{\mu}{\theta}$. Therefore, the bounds may also be satisfied by changing the set point. However, the steady state of Z_1 may depend non-trivially on the set point and direct satisfaction of the bounds may be hard to achieve.

Degradation / dilution

As has been shown in Briat, Gupta, et al., 2016, the integral action of the controller is revealed by taking the difference $\frac{dZ_1(t)}{dt} - \frac{dZ_2(t)}{dt}$ and integrating, which gives

$$(Z_1 - Z_2)(t) = \theta \int_0^t \left(\frac{\mu}{\theta} - X_n(s) \right) ds. \tag{1.16}$$

From Eq. (1.16) it can be seen that the integral over the error between the set point and the controlled species X_n is realized in the difference between the two controller species Z_1 and Z_2.

In biology, most molecules are constantly turned over and therefore a more accurate model of the antithetic integral feedback circuit should incorporate dilution and/or degradation of the controller species. Adding these reactions to the model in Eq. (1.3) gives

$$\begin{aligned} \frac{dZ_1(t)}{dt} &= \mu - \eta Z_1(t) Z_2(t) - \phi Z_1(t) \\ \frac{dZ_2(t)}{dt} &= \theta X_n(t) - \eta Z_1(t) Z_2(t) - \psi Z_2(t) \\ \frac{d\mathbf{X}(t)}{dt} &= g(\mathbf{X}(t)) + e_1 k Z_1(t). \end{aligned} \tag{1.17}$$

Here, we have introduced dilution/degradation by subtracting the terms $\phi Z_1(t)$ and $\psi Z_2(t)$ to their respective equations. Reevaluating the integral action with the updated model gives

$$(Z_1 - Z_2)(t) = \theta \int_0^t \left(\frac{\mu}{\theta} - X_n(s) - \frac{\phi}{\theta} Z_1(s) + \frac{\psi}{\theta} Z_2(s) \right) ds. \tag{1.18}$$

Now, the error integrated over does no longer only depend on the set point but also the levels of the controller species. Because the levels of the controller species will also depend on the reaction rates in the controlled network, the controlled species X_n will no longer be robust to changes in these parameters as well as changes in the network structure.

The most straightforward way to mitigate this effect is by ensuring that the dilution/degradation rates ϕ and ψ are low such that their contribution is minimized. Alternatively, increasing the sensing reaction rate θ will decrease the contribution of the dilution/degradation of the controller species Z_1 and Z_2. However, when increasing the rate θ one should also increase the reference reaction rate μ to maintain the set point at the desired level.

1.5 Bibliography

Frei, T., & Khammash, M. (2021). Adaptive circuits in synthetic biology. *Current Opinion in Systems Biology*, *28*, 100399.

Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., Jackson, R. B., et al. (2014). *Campbell biology*. Pearson Boston.

Kotas, M. E., & Medzhitov, R. (2015). Homeostasis, inflammation, and disease susceptibility. *Cell*, *160*(5), 816–827.

Kemmer, C., Gitzinger, M., Daoud-El Baba, M., Djonov, V., Stelling, J., & Fussenegger, M. (2010). Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nature biotechnology*, *28*(4), 355–360.

Rössger, K., Charpin-El-Hamri, G., & Fussenegger, M. (2013). A closed-loop synthetic gene circuit for the treatment of diet-induced obesity in mice. *Nature communications*, *4*(1), 1–9.

Xie, M., Ye, H., Wang, H., Charpin-El Hamri, G., Lormeau, C., Saxena, P., Stelling, J., & Fussenegger, M. (2016). β-cell–mimetic designer cells provide closed-loop glycemic control. *Science*, *354*(6317), 1296–1301.

Smole, A., Lainšček, D., Bezeljak, U., Horvat, S., & Jerala, R. (2017). A synthetic mammalian therapeutic gene circuit for sensing and suppressing inflammation. *Molecular therapy*, *25*(1), 102–119.

Ma, W., Trusina, A., El-Samad, H., Lim, W. A., & Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell*, *138*(4), 760–773.

Xiao, F., & Doyle, J. C. (2018). Robust perfect adaptation in biomolecular reaction networks. *2018 IEEE Conference on Decision and Control (CDC)*, 4345–4352.

Aoki, S. K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., & Khammash, M. (2019). A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, *570*(7762), 533–537.

Gupta, A., & Khammash, M. (2022). Universal structural requirements for maximal robust perfect adaptation in biomolecular networks. *bioRxiv*. https://doi.org/10.1101/2022.02.01.478605

Mangan, S., & Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences*, *100*(21), 11980–11985.

Zechner, C., Seelig, G., Rullan, M., & Khammash, M. (2016). Molecular circuits for dynamic noise filtering. *Proceedings of the National Academy of Sciences*, *113*(17), 4729–4734.

Ferrell Jr, J. E. (2016). Perfect and near-perfect adaptation in cell signaling. *Cell systems*, *2*(2), 62–67.

Basu, S., Mehreja, R., Thiberge, S., Chen, M.-T., & Weiss, R. (2004). Spatiotemporal control of gene expression with pulse-generating networks. *Proceedings of the National Academy of Sciences*, *101*(17), 6355–6360.

Entus, R., Aufderheide, B., & Sauro, H. M. (2007). Design and implementation of three incoherent feed-forward motif based biological concentration sensors. *Systems and synthetic biology*, *1*(3), 119–128.

Guo, S., & Murray, R. M. (2019). Construction of incoherent feedforward loop circuits in a cell-free system and in cells. *ACS synthetic biology*, *8*(3), 606–610.

Bleris, L., Xie, Z., Glass, D., Adadey, A., Sontag, E., & Benenson, Y. (2011). Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Molecular systems biology*, *7*(1), 519.

Strovas, T. J., Rosenberg, A. B., Kuypers, B. E., Muscat, R. A., & Seelig, G. (2014). Microrna-based single-gene circuits buffer protein synthesis rates against perturbations. *ACS synthetic biology*, *3*(5), 324–331.

Lillacci, G., Benenson, Y., & Khammash, M. (2018). Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic acids research*, *46*(18), 9855–9863.

Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., & Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, *11*(1), 1–14.

Segall-Shapiro, T. H., Sontag, E. D., & Voigt, C. A. (2018). Engineered promoters enable constant gene expression at any copy number in bacteria. *Nature biotechnology*, *36*(4), 352.

Bondí, R., Longo, F., Messina, M., D'Angelo, F., Visca, P., Leoni, L., & Rampioni, G. (2017). The multi-output incoherent feedforward loop constituted by the transcriptional regulators lasr and rsal confers robustness to a subset of quorum sensing genes in pseudomonas aeruginosa. *Molecular BioSystems*, *13*(6), 1080–1089.

Lormeau, C., Rudolf, F., & Stelling, J. (2021). A rationally engineered decoder of transient intracellular signals. *Nature communications*, *12*(1), 1–8.

Benzinger, D., Ovinnikov, S., & Khammash, M. (2021). Synthetic gene networks recapitulate dynamic signal decoding and differential gene expression. *bioRxiv*, 2021–01. https://doi.org/10.1101/2021.01.07.425755

Shimoga, V., White, J. T., Li, Y., Sontag, E., & Bleris, L. (2013). Synthetic mammalian transgene negative autoregulation. *Molecular systems biology*, *9*(1), 670.

Bloom, R. J., Winkler, S. M., & Smolke, C. D. (2015). Synthetic feedback control using an rnai-based gene-regulatory device. *Journal of biological engineering*, *9*(1), 1–13.

Kelly, C. L., Harris, A. W. K., Steel, H., Hancock, E. J., Heap, J. T., & Papachristodoulou, A. (2018). Synthetic negative feedback circuits using engineered small rnas. *Nucleic acids research*, *46*(18), 9875–9889.

Guinn, M. T., & Balázsi, G. (2019). Noise-reducing optogenetic negative-feedback gene circuits in human cells. *Nucleic acids research*, *47*(14), 7703–7714.

Becskei, A., & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, *405*(6786), 590–593.

Nevozhay, D., Adams, R. M., Murphy, K. F., Josić, K., & Balázsi, G. (2009). Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. *Proceedings of the National Academy of Sciences*, *106*(13), 5123–5128.

Szenk, M., Yim, T., & Balázsi, G. (2020). Multiplexed gene expression tuning with orthogonal synthetic gene circuits. *ACS synthetic biology*, *9*(4), 930–939.

Stapleton, J. A., Endo, K., Fujita, Y., Hayashi, K., Takinoue, M., Saito, H., & Inoue, T. (2012). Feedback control of protein expression in mammalian cells by tunable synthetic translational inhibition. *ACS synthetic biology*, *1*(3), 83–88.

Ng, A. H., Nguyen, T. H., Gomez-Schiavon, M., Dods, G., Langan, R. A., Boyken, S. E., Samson, J. A., Waldburger, L. M., Dueber, J. E., Baker, D., et al. (2019). Modular and tunable biological feedback control using a de novo protein switch. *Nature*, *572*(7768), 265–269.

Shoval, O., Alon, U., & Sontag, E. (2011). Symmetry invariance for adapting biological systems. *SIAM Journal on Applied Dynamical Systems*, *10*(3), 857–886.

Drengstig, T., Ni, X., Thorsen, K., Jolma, I., & Ruoff, P. (2012). Robust adaptation and homeostasis by autocatalysis. *The Journal of Physical Chemistry B*, *116*(18), 5355–5363.

Briat, C., Zechner, C., & Khammash, M. (2016). Design of a synthetic integral feedback circuit: Dynamic analysis and dna implementation. *ACS synthetic biology*, *5*(10), 1108–1116.

Lander, A. D., Gokoffski, K. K., Wan, F. Y., Nie, Q., & Calof, A. L. (2009). Cell lineages and the logic of proliferative control. *PLoS Biol*, *7*(1), e1000015.

Buzi, G., Lander, A. D., & Khammash, M. (2015). Cell lineage branching as a strategy for proliferative control. *BMC biology*, *13*(1), 1–15.

You, L., Cox, R. S., Weiss, R., & Arnold, F. H. (2004). Programmed population control by cell–cell communication and regulated killing. *Nature*, *428*(6985), 868–871.

Briat, C., Gupta, A., & Khammash, M. (2016). Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell systems*, *2*(1), 15–26.

Lillacci, G., Aoki, S., Schweingruber, D., & Khammash, M. (2017). A synthetic integral feedback controller for robust tunable regulation in bacteria. *Biorxiv*, 170951.

Agrawal, D. K., Marshall, R., Noireaux, V., & Sontag, E. D. (2019). In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nature communications*, *10*(1), 1–12.

Huang, H.-H., Qian, Y., & Del Vecchio, D. (2018). A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nature communications*, *9*(1), 1–12.

Qian, Y., & Del Vecchio, D. (2018). Realizing ‘integral control’in living cells: How to overcome leaky integration due to dilution? *Journal of The Royal Society Interface*, *15*(139), 20170902.

Milo, R., & Phillips, R. (2015). *Cell biology by the numbers*. Garland Science.

Frei, T., Chang, C.-H., Filo, M., Arampatzis, A., & Khammash, M. (2021). Genetically engineered proportional-integral feedback controllers for robust perfect adaptation in mammalian cells. *bioRxiv*. https://doi.org/10.1101/2020.12.06.412304

Galdzicki, M., Clancy, K. P., Oberortner, E., Pocock, M., Quinn, J. Y., Rodriguez, C. A., Roehner, N., Wilson, M. L., Adam, L., Anderson, J. C., et al. (2014). The synthetic biology open language (sbol) provides a community standard for communicating designs in synthetic biology. *Nature biotechnology*, *32*(6), 545–550.

Baig, H., Fontanarossa, P., McLaughlin, J., Scott-Brown, J., Vaidyanathan, P., Gorochowski, T., Misirli, G., Beal, J., Myers, C., et al. (2021). Synthetic biology open language visual (sbol visual) version 3.0. *Journal of integrative bioinformatics*, *18*(3).

Gunawardena, J. (2003). Chemical reaction network theory for in-silico biologists. *Notes available for download at http://vcp. med. harvard. edu/papers/crnt. pdf*, *5*.

Ceroni, F., Algar, R., Stan, G.-B., & Ellis, T. (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature methods*, *12*(5), 415–418.

Qian, Y., Huang, H.-H., Jiménez, J. I., & Del Vecchio, D. (2017). Resource competition shapes the response of genetic circuits. *ACS synthetic biology*, *6*(7), 1263–1272.

Jones, R. D., Qian, Y., Siciliano, V., DiAndreth, B., Huh, J., Weiss, R., & Del Vecchio, D. (2020). An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nature communications*, *11*(1), 1–16.

Ceroni, F., Boo, A., Furini, S., Gorochowski, T. E., Borkowski, O., Ladak, Y. N., Awan, A. R., Gilbert, C., Stan, G.-B., & Ellis, T. (2018). Burden-driven feedback control of gene expression. *Nature methods*, *15*(5), 387–393.

Huang, H.-H., Bellato, M., Qian, Y., Cárdenas, P., Pasotti, L., Magni, P., & Del Vecchio, D. (2021). Dcas9 regulator to neutralize competition in crispri circuits. *Nature Communications*, *12*(1), 1–7.

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *j. physiol.*, *40*, 4–7.

Michaelis, L., Menten, M. L. et al. (1913). Die kinetik der invertinwirkung. *Biochem. z*, *49*(333-369), 352.

Chapter Two

Characterization and mitigation of gene expression burden in mammalian cells

---

Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., & Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, *11*(1), 1–14

2.1 Abstract

Despite recent advances in circuit engineering, the design of genetic networks in mammalian cells is still painstakingly slow and fraught with inexplicable failures. Here, we demonstrate that transiently expressed genes in mammalian cells compete for limited transcriptional and translational resources. This competition results in the coupling of otherwise independent exogenous and endogenous genes, creating a divergence between intended and actual function. Guided by a resource-aware mathematical model, we identify and engineer natural and synthetic miRNA-based incoherent feedforward loop (iFFL) circuits that mitigate gene expression burden. The implementation of these circuits features the use of endogenous miRNAs as elementary components of the engineered iFFL device, a versatile hybrid design that allows burden mitigation to be achieved across different cell-lines with minimal resource requirements. This study establishes the foundations for context-aware prediction and im-

provement of in vivo synthetic circuit performance, paving the way towards more rational synthetic construct design in mammalian cells.

2.2 Introduction

Mammalian synthetic biology facilitates the study of diverse biological processes including gene regulation (Brinkman et al., 1995), developmental patterns (Bamforth et al., 2001), evolution (Farquhar et al., 2019), and cancer progression (Liu et al., 2000). More recently, it has gained clinical relevance, offering powerful new tools for the engineering of recombinant proteinproducing cells (Stuible et al., 2018) and for the creation of novel cell-based therapies for clinical use (M. Xie et al., 2016; Siciliano et al., 2018; Caliendo et al., 2019). Prior to cell engineering, the synthetic parts and the behavior of their resulting devices are tested and characterized via transient transfection in the desired mammalian cell lines. However, often the discrepancy between expected and actual behavior leads to numerous design–build–test–learn iterations (MacDonald & Siciliano, 2017; di Bernardo et al., 2012), which are particularly expensive and time consuming (Xiang et al., 2018) in mammalian cells.

At the core of the problem is the poor predictability of gene expression (di Bernardo et al., 2012) in transfected cells arising from the dependence of gene expression on the cellular context. In particular, the often overlooked dependence of exogenous genetic circuits on limited host resources that are shared with endogenous pathways frequently leads to unanticipated and counterintuitive circuit behaviors (Borkowski et al., 2016). In bacterial cells, substantial progress towards increasing the predictability of gene expression has been made by showing that exogenous genetic material imposes a significant burden, resulting in decreased growth rates and degraded cellular performance (Ceroni et al., 2015). This has been attributed to the diversion of the pool of resources available for gene expression (Li et al., 2014; Scott et al., 2010) towards transcription and translation of the newly introduced synthetic payloads. These observations prompted the development of models that consider gene expression in a resource-limited context (Carbonell-Ballestero et al., 2016; Qian et al., 2017; Weiße et al., 2015; Gyorgy et al., 2015) and led to approaches for mitigating the impact of resource burden in bacteria (Huang et al., 2018; Ceroni et al., 2018). Analogous studies in Saccharomyces cerevisiae showed that transcription and translation are limiting processes (Kafri et al., 2016). For example, the use of potent transactivators — such as the DOX-inducible rtTA — causes a squelching shortage of general transcription factors for native gene expression in yeast (Gouda et al., 2019). In mammalian cells, while performance shortcomings of synthetic circuits due to transactivator dosage and plasmid uptake variation (Lillacci et al., 2018) have been observed, a deeper understanding of the problem of resource burden and methods for its mitigation are still missing. Competition for endogenous resources can have detrimental effects on basic and translational biology. For instance, in studies based on transient DNA expression, genes that are used to normalize the results might be subject to resource-dependent expression coupling (e.g. protein levels measured by flow cytometry are usually normalized to the expression levels of the transfection marker, which is also used as a measure of transfection efficiency).

Here, we investigate the burden imposed by transiently expressed synthetic circuits on host cells (Fig. 2.1). Through the design of genetic constructs that allow us to uncouple transcription and translation processes, we separately study transcriptional and translational burden caused by cellular resource sharing. In particular, we engineer several regulatory circuits composed of a tunable load, called X-tra (eXtra Transgene), which we genetically express in the host cell in varying amounts. We then measure the impact of this tunable load on a “sensor” gene, which we refer to as the capacity monitor (Fig. 2.1a). We demonstrate in different mammalian cell lines that the sharing of transcriptional and translational resources in the host cell can tightly couple otherwise independently coexpressed synthetic genes and lead to trade-offs in their expression (Fig. 2.1a). To enhance the predictability of synthetic devices in mammalian cells, we explicitly incorporate these load-sharing effects in a general mathematical model in which we replace the rates of resource-dependent reactions with adjusted effective rates (Fig. 2.1b). This framework follows ideas originally used to capture the competitive interaction of multiple inhibitors with an enzyme (Chou & Talaly, 1977) and has been applied to describe shared cellular resources in previous studies (Carbonell-Ballestero et al., 2016; Qian et al., 2017; Weiße et al., 2015; Gyorgy et al., 2015; Rondelez, 2012). We demonstrate the usefulness of this modeling framework by showing that it successfully recapitulates the non-monotonic dose–response behavior of a simple inducible gene expression system observed in Lillacci et al. (Lillacci et al., 2018). Additionally, we investigate the role of posttranscriptional regulators, like RNA-binding proteins (RBPs) and microRNAs (miRNAs), in mitigating the impact of burden-induced coupling and find that both are able to reallocate resources, making them candidates for use in burden-mitigation circuits. Using these observations, and guided by our modeling framework, we identify the incoherent feedforward loop (iFFL) as a network topology that is particularly effective at resource burden mitigation, and then we use endogenous and synthetic miRNA regulation to engineer iFFL-based, burden-mitigating synthetic circuits (Fig. 2.1c). While miRNA-based iFFL circuits have been previously constructed and proposed to buffer gene expression against noise (Siciliano et al., 2013; Carignano et al., 2018) and fluctuations in external inducer concentration (Strovas et al., 2014),

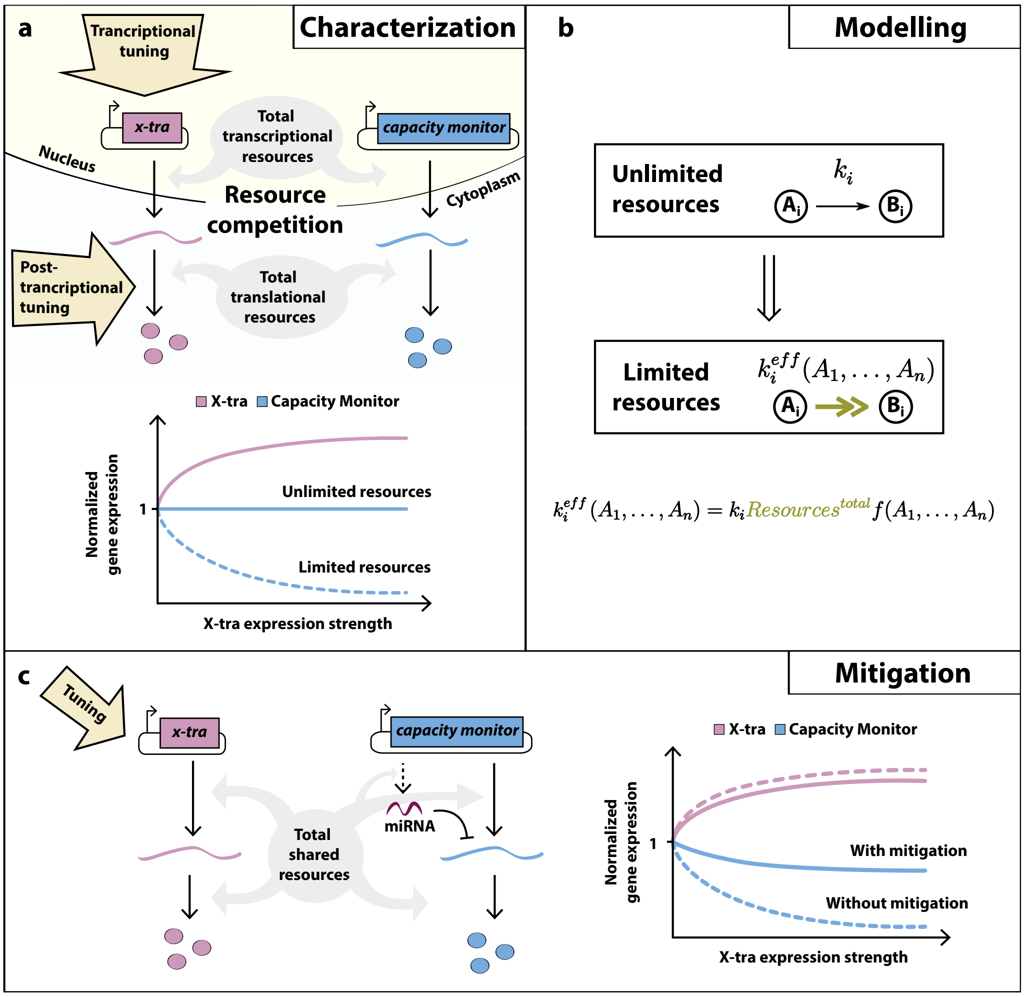


Figure 2.1: Resource sharing and the origin of gene expression burden. a Characterization of gene expression burden. Expression of independent exogenous genes impacts on host cellular resources. Thus, perturbations in one gene's expression (hereby named X-tra) affect the expression of a second gene (hereby named capacity monitor). b Modeling of gene expression in a resource-limited environment. Modeling of gene expression is generally performed under the assumption of unlimited resources. A simple framework enables the straightforward transformation of such a model to a system that incorporates resources explicitly. The transformation involves a simple function that scales the original reaction rate. c Mitigation of gene expression burden. A simple microRNA-based circuit motif is capable of mitigating the burden-induced coupling of X-tra and the capacity monitor. It should be noted that the dynamic range of X-tra also slightly decreases as a consequence of mitigation. However, as it will be discussed in the results section and shown in Supplementary Fig. A.19, the absolute expression of X-tra is higher with mitigation.

in this study we demonstrate that they also act to rescue the expression level of genes of interest despite changes in available cellular resources due to the loading effects of transgene constructs (Fig. 2.1c). Our findings pave the way to more realistic output predictions and optimal synthetic construct design in mammalian cells.

2.3 Results

Genetic circuits compete for limited shared resources

We reasoned that competition for finite cellular resources would introduce an indirect coupling in the expression levels of two otherwise independently expressed genes. To test this, we cotransfected HEK293T cells with two constitutively expressed fluorescent proteins mCitrine and mRuby3 driven by EF1α promoters, in molar ratios ranging from 1:4 to 4:1, for a total of 50 ng (low) or 500 ng (high) of encoding plasmid (Fig. 2.2a). The competition for limited resources is expected to shape gene expression as presented in Fig. 2.2a, according to the modeling framework that will be introduced in Fig. 2.4a (model described in Supplementary Note A.2). As expected, the total amount of 500 ng of encoding plasmids results in a dramatic drop of encoded-gene expression as compared to 50 ng (Fig. 2.2a, right). Furthermore, in both experimental conditions mCitrine and mRuby3 fluorescence levels are negatively correlated; the higher the amount of expressed mCitrine, the lower that of mRuby3 and vice versa (Fig. 2.2a, right); this correlation was also more severe for 500 ng of transfected plasmid than for 50 ng.

We demonstrated that the negative correlation is promoter independent: using a CMV and a PGK promoter (Qin et al., 2010) that have different expression strength in HEK293T and H1299 (Supplementary Fig. 2.1a), we observed analogous outcomes (Supplementary Fig. 2.1b–e). Further, by combining different molar ratios of mCitrine and mRuby3 encoding plasmids driven by two promoters of different strengths (EF1α or EFS) a similar behavior to Fig. 2.2a was observed (Supplementary Fig. 2.2). Finally, as many synthetic circuits rely on tunable gene expression, we next tested resource competition on transcriptional inducible systems, by modulating X-tra repression with a Doxycycline (Dox)-repressed promoter (Fig. 2.2b) at different concentrations of Dox (from 0 to 1 μg/mL) while keeping capacity monitor amounts constant (Fig. 2.2b, left). Consistent with previous results, we observed that increased repression of X-tra corresponds to increased capacity monitor levels (Fig. 2.2b, right).

To exclude any bias of fluorescent protein expression on resource competition, we transfected a plasmid encoding a human codon optimized variant of

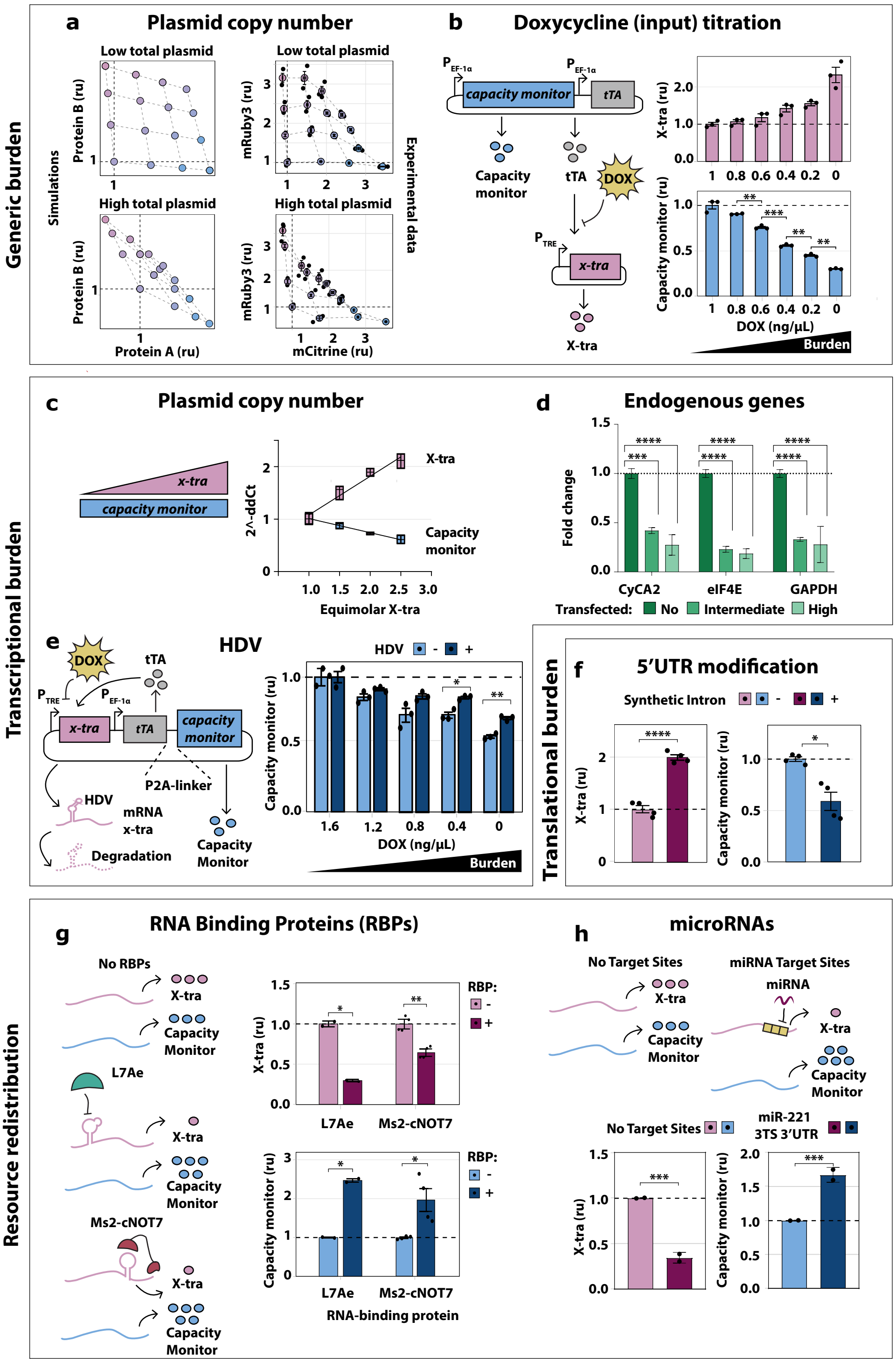


Figure 2.2: See caption on the next page.

Figure 2.2: Burden imposed by genetic circuits in mammalian cells. a Left: As the total plasmid amount increases, the total expression plateaus. Right: Titration of two plasmids expressing the fluorescent proteins mCitrine and mRuby3 from EF1α promoters in ratios from 1:4 to 4:1 (total of 50 ng, top right; or 500 ng of DNA, bottom right). N = 3 biological replicates. Source data are provided as a Source Data file. b Two plasmids were co-transfected, one constitutively expressing capacity monitor and tTA from a strong constitutive promoter and the other expressing X-tra from a tTA responsive promoter. Capacity monitor levels counterbalance the increase in X-tra expression. Flow cytometry data are normalized to the expression at maximal Dox. N = 3 biological replicates. Source data are provided as a Source Data file. c mRNA quantification of X-tra and a capacity monitor expressed at different molar ratios. As the X-tra increases, the mRNA levels of the capacity monitor decreases. N = 4 biological replicates. qPCR analysis was performed 48 h post-transfection and data show fold change ± SE. Source data are provided as a Source Data file. d Cells transfected with a plasmid expressing two fluorescent proteins from a bidirectional promoter were sorted according to high, intermediate, or no fluorescence (Supplementary Fig. A.4) for mRNA extraction. mRNA levels expressed from endogenous genes decrease in cells with intermediate and high fluorescence. N = 3 biological replicates. Data show fold change ± SE. Individual values are plotted in Supplementary Fig. A.28. Source data are provided as a Source Data file. e Capacity monitor levels are higher with an HDV ribozyme rapidly degrading the capacity monitor mRNA than with an inactive mutant, suggesting a sequestration of transcriptional resources. N = 3 biological replicates (N = 2 for HDV, 1.6 ng/µL DOX). Source data are provided as a Source Data file. f The synthetic intron shows higher X-tra levels compared to a control and leads to reduced capacity monitor levels. N = 4 biological replicates. Source data are provided as a Source Data file. g Repressed X-tra expression leads to increased capacity monitor levels. N = 2 biological replicates for L7Ae and N = 4 for Ms2-cNOT7. Source data are provided as a Source Data file. h When X-tra is downregulated by miR-221 endogenously expressed in HEK293T cells, the capacity monitor levels increase. All flow cytometry data were acquired 48 h post-transfection and are plotted as mean ± SE. SE standard error, r.u. relative units. N = 2 biological replicates. Source data are provided as a Source Data file. Unpaired two-sided T-test. P value: ****<0.0001, ***<0.0005, **<0.005, *<0.05.

the bacterial σ-factor sigW in increasing amounts with a fixed concentration of the mCitrine capacity monitor plasmid, and demonstrated similar behavior to fluorescent protein expression (Supplementary Fig. A.3).

Finally, to avoid any experimental confounds as the source of our observations, we showed that neither cell seeding nor nutrient supply had any apparent effect on the expression levels of the two genes, one of which was titrated whereas the second was held at a constant copy number (Supplementary Fig. A.4).

These proof-of-concept experiments demonstrate that (i) gene expression in mammalian synthetic circuits is connected even in the absence of direct regulation and (ii) expression of exogenous genes is limited by cellular resource availability.

Transcriptional and translational resources are limiting

Since several different resource pools could be responsible for the observed effects described above, we set out to characterize the individual contributions of transcriptional and translational resource limitation to cellular burden in HEK293T and H1299 cells (Fig. 2.2). To evaluate potential limitations in transcriptional resources and the consequent gene competition for mRNA expression, we quantified mRNA levels in cells expressing X-tra/capacity monitor molar ratios from 1:1 to 2.5:1 in H1299 cells for a total of 500 ng of plasmid DNA (corresponding protein data in Supplementary Fig. A.8a). We observed that as the X-tra mRNA increased, the capacity monitor mRNA levels decreased (Fig. 2.2c), supporting the hypothesis that shared transcriptional resources are indeed a limiting factor in mammalian synthetic gene co-expression.

To investigate whether the expression of endogenous genes is also affected by heterologous genetic payloads, we transfected H1299 cells with a plasmid encoding for EGFP and mKate under the control of a bidirectional promoter. We then sorted transfected cells according to high and intermediate levels of fluorescent markers as well as non-transfected cells (absence of fluorescence) (Supplementary Fig. A.5). We then quantified the mRNA levels of three endogenous genes (CyCA2, eIF4E, GAPDH, Fig. 2.2d, Supplementary Fig. A.28). Notably, in transfected cells that express high and intermediate levels of EGFP and mKate, the expression of CyCA2, eIF4E, and GAPDH decreases when compared to the non-transfected population. We also measured the mRNA levels of CyCA2, eIF4E, and GAPDH in cells transfected with X-tra/capacity monitor molar ratios from 1:1 to 2:1 and observed a progressive, albeit not dramatic decrease with higher amounts of X-tra when compared to the 1:1 ratio (Supplementary Fig. A.6). Of note, in the latter experiment cells were not sorted before mRNA extraction.

To provide further support to the observations on transcriptional burden on exogenous genes (Fig. 2.2c), we implemented a genetic circuit that can selectively overload the transcriptional resource pool without sequestering translational resources. The system is based on the self-cleaving hepatitis delta virus (HDV) ribozyme, which ensures that most of the transcribed mRNA is cleaved and thus destabilized (Fig. 2.2e, left). The circuit is composed of a single plasmid with two transcriptional units (TUs). One TU contains a tTA transcription factor co-expressed with the mRuby3 (capacity monitor) via the P2A peptide, driven by a constitutive promoter. The second TU includes the HDV-Xtra expression regulated by the TRE promoter. In this setup, Dox can be used to modulate the amount of burden imposed, similar to what was already shown in Fig. 2.2b.

We compared this circuit to a catalytically inactive mutant of the HDV ribozyme in HEK293T cells. As expected, we observed that when the HDV ribozyme is inactive, X-tra protein levels increase with decreasing amounts of Dox (Supplementary Fig. A.7, top pale pink bar), whereas those of the capacity monitor decrease (Fig. 2.2e, bottom pale blue bar). In contrast, when the HDV ribozyme is active, X-tra expression is strongly reduced and only minorly increasing with lower Dox concentrations (Supplementary Fig. A.7, top dark purple bar). Here, the capacity monitor levels decrease to a smaller extent than in the previous condition, supporting the observations in Fig. 2.2c that transcriptional resources are limited to a certain extent (Fig. 2.2e, dark blue bar). Interestingly, the expression levels of the capacity monitor with active HDV ribozyme are higher compared to the inactive mutant (Supplementary Fig. A.7, bottom dark blue bar). We suggest that, assuming that the X-tra mRNA with an active HDV ribozyme is decapped and rapidly degraded, it is likely to sequester fewer translational resources, which should result in higher expression of the capacity monitor.

Transcriptional resource pool sharing is therefore at least partially responsible for the described gene expression trade-offs, and translational resources may represent an additional bottleneck to the overall expression of synthetic genes. We confirmed this hypothesis by adding a synthetic intron (Lu et al., 2013) in the 5′ untranslated region (UTR) of the X-tra fluorescent protein (Fig. 2.2f, top). The synthetic intron enhances translation by augmenting mRNA export from the nucleus to the cytoplasm (Lu et al., 2013) and therefore imposes specific translational load. Indeed, we observed higher expression of X-tra in HEK293T (Fig. 2.2f) and H1299 (Supplementary Fig. A.8b) cell lines in the presence of a synthetic intron, accompanied by lower capacity monitor levels, confirming that resources employed for translational regulation are also limiting. Thus our data collectively indicate that exogenous genes compete for resources both at the transcriptional and translational levels, overall imposing a gene expression burden on mammalian cells.

Since one of the goals in synthetic biology is output predictability, reproducibility, and robustness, gene expression burden is a key issue to address. We reasoned that posttranscriptional and translational regulators, such as RBPs and miRNAs, may free up cellular resources (Kallehauge et al., 2017) by repressing target mRNA translation or inducing its degradation. If true, they could be exploited in more robust circuit topologies to reduce gene expression load, resulting in improved performance and predictability of engineered circuits. Therefore, we tested two RBPs, L7Ae and Ms2-cNOT7 (Wroblewska et al., 2015; Cella et al., 2018), as well as endogenous miRNAs, miR-221 and miR-31, in HEK293T (Fig. 2.2g, h) and H1299 (Supplementary Fig. A.8c, d) respectively. For each system, a fluorescent protein encoding mRNA targeted by either RBPs or miRNAs (X-tra) was co-expressed with a second, constitutively expressed fluorescent readout (capacity monitor). L7Ae binds the 5′UTR of the X-tra mRNA inhibiting its translation, whereas Ms2 binds target sites (TS) in the 3′UTR of the X-tra transcript, allowing cNOT7 to cut the polyA tail to destabilize the target mRNA (Wroblewska et al., 2015). We consistently observed in both cell lines that X-tra downregulation by RBPs results in increased levels of the capacity monitor (Fig. 2.2g, Supplementary Fig. A.8c).

miRNAs operate by either translation inhibition or mRNA degradation, according to complete (Gam et al., 2018) or partial (Ameres & Zamore, 2013) complementarity to the mRNA target. To evaluate the effect of miRNA regulation on cellular resource reallocation, we placed three perfect complementary TS in the 3′UTR of X-tra, which respond to the endogenous miR-221 and miR-31 highly expressed in HEK293T and H1299 cells. The capacity monitor expression levels increased when the X-tra mRNA was downregulated by miRNAs, as compared to controls lacking miRNA TS (Fig. 2.2h, Supplementary Fig. A.8d).

To further demonstrate that the burden imposed by synthetic circuits is cell-type independent, we performed the same set of experiments of Supplementary Fig. 2.1d and Fig. 2.2f–h in U2OS, HeLa, and CHO-K1 cells, obtaining similar results (Supplementary Figs. A.9–A.11). Interestingly, even CHO-K1 cells, which are the workhorses of the biopharmaceutical industry due to their high productive capability (Meleady et al., 2011) show cellular burden. Redistribution of resources was also observed by the RBPs L7Ae and MS2-cNot7 and the highly expressed endogenous miR-221 and miR-21 in U2OS and HeLa/CHO-K1 cells, respectively.

These results confirm that post-transcriptional regulators can redistribute intracellular resources and, importantly, that this phenomenon is cell-context independent. The extent of negative correlation between X-tra and capacity monitor expression, as well as the amount of repression by post-transcriptional regulators, differs across cell lines; this could be the consequence of several factors, such as the relative abundance of transcriptional, post-transcriptional, and translational resources.

A major advantage of miRNAs over RBPs is that they are endogenously expressed and cell line specific. Thus, their expression does not impose an additional burden, and since several thousand endogenous miRNAs with different TS are naturally present in mammalian cells (Alles et al., 2019), the design space is rather large, giving rise to a tremendous number of circuits that can be easily tailored to the cell/tissue of interest. Based on the results presented here, we envision that genetic circuits that mitigate resource competition via miRNAs may be designed for any mammalian cell line with a very broad set of potential applications.

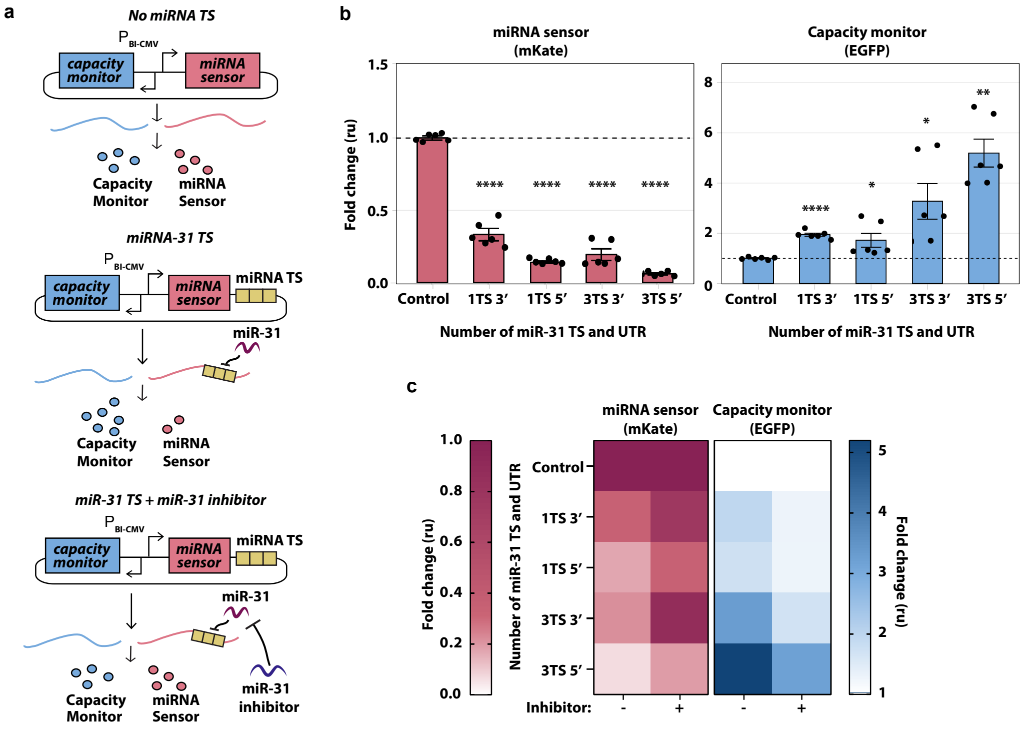


Figure 2.3: Impact of miRNA target sites number and location on burden. a Schematics of experimental design to infer miRNA-mediated cellular resources redistribution. EGFP (capacity monitor) and mKate (miRNA sensor) are encoded on the same bidirectional CMV promoter plasmid. One or 3 TS for miR-31 (TS) are added either in the 3′ or 5′UTR of mKate. Control: no miR-31 TS. Hypothesis: in the absence of miR-31 regulation, capacity monitor and miRNA sensor are expressed to a certain level (top). In the presence of miR-31, lower miRNA sensor levels correlate with higher capacity monitor expression (middle). This condition is reversed by an miR-31 inhibitor (bottom). b Fold change of miRNA sensor and capacity monitor protein levels compared to control (set to 1). EGFP increases up to fivefold with the strongest downregulation of mKate (3 TS 5′UTR). Flow cytometry data were acquired 48 h posttransfection and are plotted as mean ± SE. SE standard error, r.u. relative units. N = 6 biological replicates. Source data are provided as a Source Data file. Unpaired two-sided T-test. P value: ****<0.0001, **<0.005, *<0.05. c When miR-31 activity was impaired by a miR-31 inhibitor, the rescue of mKate expression corresponds to reduced EGFP levels, whereas both fluorescent proteins do not vary in the control. The heatmaps represent the fold change derived by flow cytometry data, calculated as the ratio between the geometric mean of six biological replicates and the corresponding geometric mean in the control condition. Source data are provided as a Source Data file. Bar plots and statistical analysis are reported in Supplementary Fig. A.12.

Characterizing the effect of miRNAs on resource distribution

We sought to characterize the correlation between miRNA-mediated downregulation and resource redistribution by building a library of miRNA sensors for miR-31, which is endogenously expressed in H1299 lung cancer cells (Thomson et al., 2011). The miRNA sensor is composed of the fluorescent reporter mKate with or without miR-31 TS, encoded along with the capacity monitor (EGFP) on a single plasmid with a bidirectional promoter (Fig. 2.3a). The library includes 0, 1, or 3 fully complementary miR-TS in the 3′ or 5′UTR of mKate.

Similar to what was previously observed (Supplementary Fig. A.8d), when the miRNA sensor's levels decrease as a consequence of miR-31 regulation, the expression of the capacity monitor increases. The strongest repression was achieved with 3 TS in the 5′UTR and was accompanied by corresponding higher capacity monitor levels (Fig. 2.3b). Conversely, when we rescued mKate expression by a miR-31 inhibitor (Fig. 2.3c, left and Supplementary Fig. A.12, red bars), the capacity monitor levels decreased (Fig. 2.3c, right and Supplementary Fig. A.12, dark blue bars) demonstrating that miRNA sensor and capacity monitor levels are linked. Interestingly, the effect of the miRNA inhibitor was more pronounced with TS placed in the 3′UTR. Synthetic miRNA inhibitors bind to endogenous miRNAs in an irreversible manner (Robertson et al., 2010), but differences in their action (e.g. when TS are placed in the 3′ versus 5′UTR), as well as mechanistic insights into these differences, are still missing.

To confirm that miRNA-mediated resource redistribution is independent of experimental setting and plasmid design, we encoded the miRNA sensor and capacity monitor on two separate plasmids. Similar to previous results, miRNA sensor and capacity monitor were negatively correlated (Supplementary Fig. A.13a), suggesting that cellular burden and miRNA-dependent resource reallocation are a common challenge and solution respectively. Downregulation of the miRNA sensor was also confirmed by qPCR (Supplementary Fig. A.13b). Finally, when the miR-31 sensor was transfected in low miR-31 cell lines such as U2OS and HEK293T, neither the miRNA sensor nor the capacity monitor levels varied (Supplementary Fig. A.14), further confirming the miRNA-dependent resource reallocation.

We showed in Fig. 2.2h and Supplementary Figs. A.8d, A.9d, A.10d and A.11d that miRNA-dependent resource reallocation is observed across different cell lines, by expressing cell-specific miRNA sensors which include 3 TS in the 3′UTR. We then built a library of sensors with different numbers and locations of TS for miRNA-221 and -21 which are highly expressed in U2OS and HeLa cells, respectively. We also confirmed here that miRNA sensor and capacity monitor are inversely correlated, consistent with our observations in H1299 cells (Supplementary Figs. A.15 and A.16).

Overall these data show that miRNAs can be used to develop resource-aware plasmid-designs harboring burden-mitigating circuit topologies, and that the number and location of TS can be tuned to achieve desired protein expression levels.

A resource-aware model framework

In order to provide a better understanding of our results, we developed a general resourceaware model, which offers a simple and convenient framework for extending existing models of biochemical reactions allowing them to incorporate the effects of shared limited resources.

Fig. 2.4a illustrates an overview of the framework. The main idea is to replace the rates of reactions that involve a shared resource with an effective reaction rate that captures the reduced availability of that resource due to the presence of competing genes. To create a distinction between regular reactions and resource-limited ones, we use double-headed reaction arrows to denote resource-limited reactions as illustrated at the bottom of Fig. 2.4a. This double-headed arrow summarizes the set of intermediate interactions shown in more detail at the top left of Fig. 2.4a. Here, the substrate A_i binds resource R with rate k_i^+ to form the complex C_i. This reaction is also assumed to be reversible with rate k_i^-. With a rate k_i^{cat} the complex gives rise to the product B_i, while also freeing up both the substrate A_i and the resource R. We assume that the total amount of resource, R^{total}, is conserved and remains constant at the time scale of the considered reactions. Considering all possible substrates that require resource R and assuming that C_i is in quasi-steady state, the rate for resource-limited production can be expressed as k_i^{eff}, shown in the top right corner of Fig. 2.4a. A more detailed derivation can be found in Supplementary Note A.1. k_i^{eff} is a function of the total amount of resources and the current concentration of all substrates competing for this resource. This expression can be readily used to substitute all reaction rates that involve shared and limited resources.

To demonstrate the effectiveness of our modeling framework, we extend the models of different circuit topologies introduced in Lillacci et al. (Lillacci et al., 2018) to include limited resources and show that the resulting extended models recapitulate the previously unexplained non-intuitive experimental observations.

The four topologies considered in Lillacci et al. (Lillacci et al., 2018) were split into two groups based on the presence of negative feedback from the fluorescent protein DsRed to the transcriptional activator (tTA). The first group consisted of the open-loop (OLP) and incoherent feedforward (IFF) topologies. In both these circuits, the constitutively expressed transcriptional transactivator, fused to the fluorescent protein Cerulean (tTA-Cer), activates the expression of the fluorescent protein DsRed. Furthermore, the gene of DsRed intronically encodes the synthetic miRNA FF4 (miR-FF4). In the IFF topology, the matched target of this miRNA is present in the 3′UTR of the DsRed gene. This target is replaced by a mismatched target for the miRNA FF5 in the OLP. These detailed interactions are depicted here in Fig. 2.4b, left side. To observe poten-

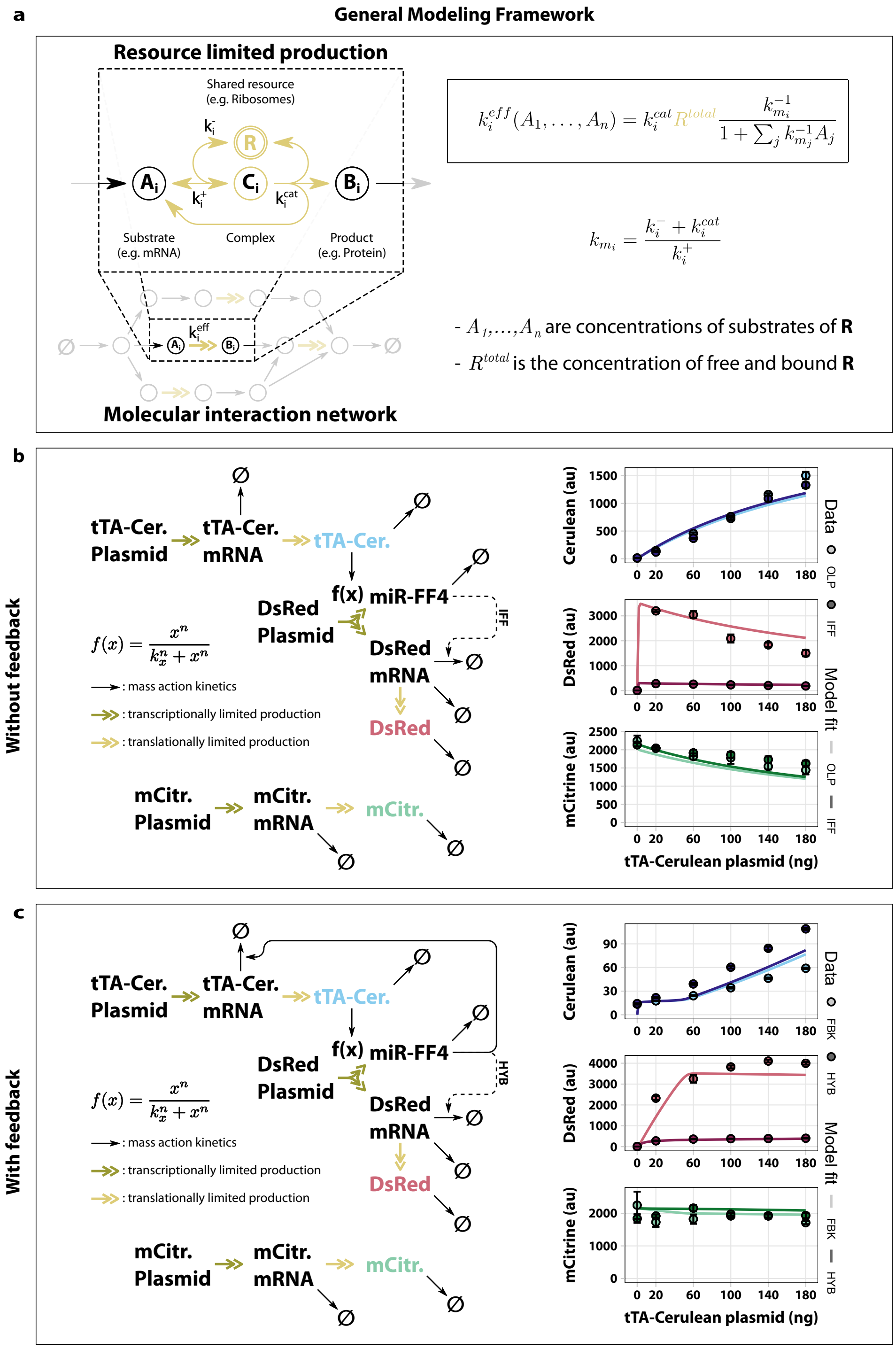


Figure 2.4: See caption on the next page.

Figure 2.4: A resource-aware mathematical modeling framework. a General framework for transforming molecular interaction network models. Existing models of molecular interaction networks can be transformed to include shared limiting resources by substituting ki, the reaction rate of a resource-limited production, with keffi. Shown above an exemplary resource-limited production are the detailed interactions between the substrate and the shared resource. b Limited shared resources reproduce non-monotonous dose response in open-loop and incoherent feedforward circuit topologies. On the left, a graphical representation of a model for both the open-loop (OLP) and incoherent feedforward (IFF) topologies from Lillacci et al. (Lillacci et al., 2018). Transcriptional activation is modeled by a Hill-type function. The solid arrows denote reactions assumed to follow the law of mass action. The model incorporates resources as introduced in panel a. These reactions are depicted as double-headed arrows. The model was fit to data obtained by transiently transfecting HEK293T cells with increasing amounts of plasmid encoding tTA-Cerulean. The data and the fit are shown on the right. c Limited shared resources reproduce nonmonotonous dose-response in feedback and hybrid circuit topologies. The model shown on the left is the same as in panel b with an additional negative feedback from miR-FF4 to tTA-mRNA. These topologies correspond to the feedback (FBK) and hybrid (HYB) topologies from Lillacci et al. (Lillacci et al., 2018). The activation of gene expression by tTA-Cerulean is modeled by a Hill-type function as shown in the center. Reactions with double-headed arrows denote resourcelimited production reactions as introduced in panel a. Solid arrows are assumed to follow the law of mass action. The model was fit to experimental data obtained from transient transfections with increasing amounts of plasmid encoding tTA-Cerulean. A description of the models can be found in Supplementary Note A.4 and the parameter values obtained by fitting are summarized in Supplementary Table A.28. Data were obtained 48 h after transfection and are plotted as mean ± SE. SE standard error. N = 3 biological replicates. Source data are provided as a Source Data file.

tial shifts in the allocation of resources, we generated dose–response curves by increasing the amount of transfected tTA-Cer plasmid, while the other two plasmids, containing DsRed and the constitutively expressed fluorescent transfection reporter mCitrine, were held constant. As can be seen from the model fit, plotted as a solid line in the data graph, the extended model reproduces the non-monotonic behavior of the dose responses (Fig. 2.4b, right).

The second group of topologies considered by Lillacci et al. (Lillacci et al., 2018) consisted of the feedback (FBK) and the FBK + IFF hybrid (HYB) topologies. In addition to all the interactions described for the OLP and IFF circuits, the FBK and the HYB circuits possess miR-FF4 targets in the 3′UTR of the tTA-Cer gene, which introduces negative feedback. Furthermore, the FBK and HYB differ from each other by the presence of a matched target for miR-FF4 in the HYB topology, which introduces incoherent feedforward and is replaced by a mismatched FF5 target in the FBK circuit. All the interactions are illustrated in detail in Fig. 2.4c, left. The dose–response curves for the two circuits were obtained as described above. Again, the fit of the extended model to the data captures its rather unexpected behavior (Fig. 2.4c, right).

Lastly, we also apply our framework to model the gene expression systems presented in Figs. 2.2b, e, g and 2.3b. The resulting model fits are shown in Supplementary Fig. A.17. The models are described in Supplementary Note A.6 and the parameter values obtained by fitting are summarized in Supplementary Table A.34–A.38.

Our simple framework adapts existing models of gene expression to include pools of shared and limited resources. We show that it can be used to provide an explanation for unintuitive dose responses in tTA-based circuits. With this framework as a tool, we believe that performance issues attributed to gene expression burden can be addressed head-on in the design phase of circuit-building, thereby reducing the need for costly subsequent build-test-learn iterations.

Mitigating burden with iFFL circuits

We implemented a strategy that exploits miRNA to reduce the indirect coupling between co-expressed genes. In particular, we took advantage of the fact that miRNA production also requires (pre-translational) cellular resources, therefore acting as a sensor for resource availability. Because of this, it is possible to reduce the coupling between genes co-expressed via a common resource pool by introducing miRNA-mediated repression of those genes (as long as the miRNA itself is also affected by the same resource pool). Since both the miRNA and the miRNA-repressed gene are affected by the availability of resources, miRNA-mediated repression implements an iFFL similar to previously published circuits (Lillacci et al., 2018; Strovas et al., 2014; Bleris et al., 2011) (Fig. 2.5a). Interestingly, this iFFL-based circuit constitutes a biological implementation of the miRNA circuit proposed by Zechner et al. (Zechner et al., 2016). In this setting, the miRNA can be interpreted as an estimator of its cellular context (e.g. amount of free resources) and acts to filter out this context, thereby minimizing its impact on the output of interest.

We explored this strategy for an endogenously expressed miRNA (Fig. 2.5b, c) and a synthetic miRNA encoded on a plasmid (Fig. 2.5d, e). More specifically, Fig. 2.5b describes a strategy that exploits endogenous miRNAs to reduce the coupling of a gene of interest (GOI) to the expression level of other genes, introduced by the limitation in resources. Implementation of this strategy only requires adding the TS of an endogenous miRNA to the 5′ UTR of the gene of interest (mKate). In our experimental setup, when the copy number of a second gene (X-tra) is increased, resources are drawn away from the expression of mKate and allocated to the expression of X-tra. The shift in resource allocation is expected to also affect miR-31, which acts as a capacity monitor. This leads to a reduction in the repression of mKate, effectively compensating for the burden imposed by the co-expression of the X-tra gene.

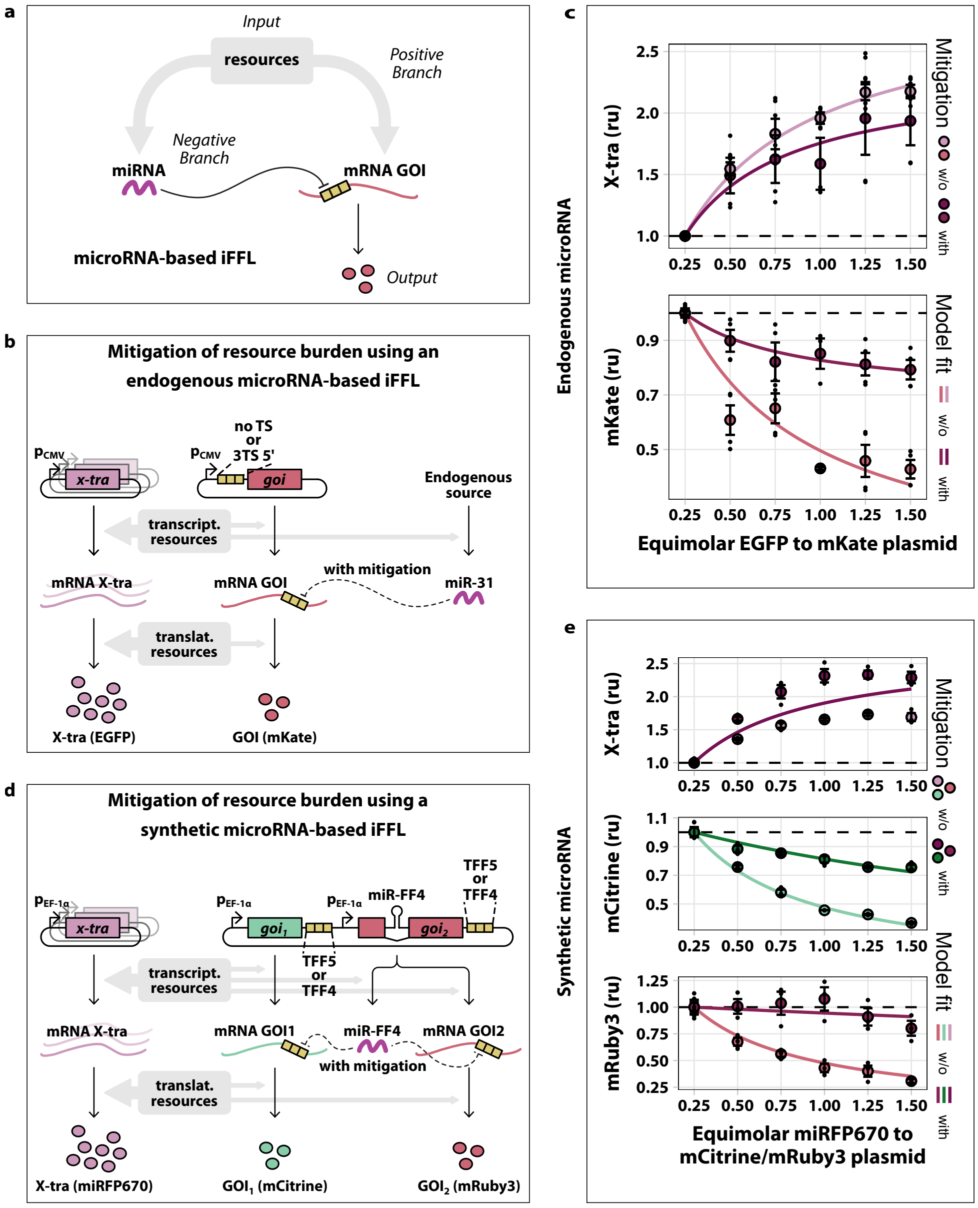


Figure 2.5: See caption on the next page.

Figure 2.5: Mitigating the effects of resource limitation with microRNA-based iFFL. a The microRNA-based incoherent feedforward loop (iFFL) motif. b Mitigation system based on endogenous microRNA. At high copy number of the X-tra, resources are drawn away from the production of the GOI and miR-31. By sensing the resource availability and repressing the GOI less when there are fewer resources, the miRNA reduces the effect of limited resources. c Two plasmids were co-transfected into H1299 cells which respectively express the X-tra and GOI genes (EGFP and mKate respectively (b)), and the molar ratio of the X-tra:GOI plasmid was progressively increased. The presence of miR-31 TS in mKate 5′UTR mitigates effects due to resource sharing. The parameter values obtained by fitting are summarized in Supplementary Table A.29. N = 3 biological replicates. d Mitigation system based on synthetic miRNA. In the presence of many copies of the X-tra gene, resources are drawn away from the production of both the GOIs and the miR-FF4. Due to lower production of miR-FF4 the GOIs are less repressed. This compensates for the reduced availability of resources. e A plasmid encoding both the fluorescent protein mCitrine and an intronic microRNA expressed from the mRuby3 gene (GOI1, GOI2 and miR-FF4 (d)) was co-transfected into HEK293T cells with increasing amounts of a plasmid expressing the X-tra gene (miRFP670 (d)). The impact of resource limitation on both GOIs was reduced when they contained three miR-FF4 targets in their 3′UTRs compared to when they contained three mismatched miR-FF5 targets. The parameter values obtained by fitting are summarized in Supplementary Table A.30. N = 3 biological replicates. Source data are provided as a Source Data file. A description of the models can be found in Supplementary Note A.5. Flow cytometry data were acquired 48 h post-transfection and are plotted as mean ± SE. SE standard error, r.u. relative units.

To demonstrate this mitigation approach experimentally, we co-transfected H1299 cells with increasing amounts of EGFP (X-tra), along with a constant amount of mKate (GOI) that either includes (for mitigation) or omits (no mitigation) three miR-31 TS in the 5′UTR. As expected, the expression level of X-tra approached saturation as the plasmid copy number increased, both for the targeted and non-targeted GOI variants (Fig. 2.5c). In agreement with previous results, the expression of the nontargeted GOI strongly decreased with increased expression of X-tra. Conversely, the decrease in expression of the targeted GOI was only about a third of that of the non-targeted variant, indicating improved adaptation to changes in resource availability (Fig. 2.5c and Supplementary Fig. A.18). This observation was also captured well by a model of the system that explicitly considered resources, as described in the previous section. It should be noted that while the relative dynamic output range of X-tra is slightly reduced (fold change of 1.94× with mitigation versus 2.18× without mitigation (Fig. 2.5c), our data show that the absolute levels of X-tra increases about 2× in the presence of miR-31-based iFFL, de facto benefiting from this network topology (Supplementary Fig. A.19). Analogously, miR-221 iFFL circuits specific for U2OS and HEK293T cells (Tian et al., 2012) (Supplementary Fig. A.21) show improved robustness to burden imposed by increasing exogenous gene load (Supplementary Figs. A.20 and A.22). Models used for fitting and the resulting parameter values are summarized in Supplementary Note A.5 and Supplementary Table A.31–A.32.

Importantly, the delivery of genetic payloads also affects the expression of endogenous genes (CyCA2, elF4E, and GAPDH), as shown in Fig. 2.2d. We then sought to compare the expression of the same endogenous genes in the presence or absence of miR-31 sensor in H1299 cells. After 48 h from transfection of EGFP and mKate on a bidirectional plasmid, with mKate either including (miRNA sensor) or not (noTS) TS for miR-31, we sorted cells according to high, intermediate, or absence of fluorescence expression (Supplementary Fig. A.23a) and performed qPCR. Curiously, we observed that in cells transfected with miR-31 sensor, the decrease in the expression of the endogenous genes was much lower than in its absence (Supplementary Fig. A.23c). Furthermore, the expression of endogenous genes was inversely proportional to the levels of fluorescent proteins (Supplementary Fig. A.22b). Thus, the lower expression of endogenous genes due to the burden imposed by exogenous payloads is counteracted by the miR-31 sensor. To investigate whether the use of endogenous miRNAs may impair the regulation of native targets, we measured the expression of SATB2 mRNA, a natural target of miR-31 (Aprelikova et al., 2010) in cells transfected with miR-31 sensor versus the noTS control, and observed no difference between the two conditions (Supplementary Fig. A.24).

Motivated by our desire to achieve portability across cell lines and multiple-output regulation, we implemented and tested a synthetic miRNA-iFFL circuit that tunes two GOIs (Fig. 2.5d). Similar to the endogenous case, the genes of interest, mCitrine (GOI1) and mRuby3 (GOI2), encode TS for the miRNA-FF4 in their 3′UTRs. In contrast to endogenous miRNA expression, however, here the miRNA is expressed intronically from GOI2. In this way, the circuit forms a self-contained unit that can be easily transferred between cell types.

We co-transfected HEK293T cells with a plasmid encoding constitutively expressed miRFP670 (X-tra) and a plasmid composed of two TUs, each expressed under the constitutive promoter EF1α (Fig. 2.5d). The first TU encodes mCitrine, whereas the second drives mRuby3. Furthermore, the 3′UTR of mCitrine and mRuby3 contained either three TS for the synthetic miRNAFF4 or three mismatched miR-FF5 TS (negative control). The miRNA-FF4 was intronically encoded in the mRuby3 gene. Identically to the endogenous case, the amount of X-tra plasmid was increased while keeping the GOIs plasmid constant. Again, expression of X-tra increased and approached saturation with increasing molar amounts and consequently, the non-targeted variants of the GOIs decreased (TFF5 in Fig. 2.5e). Conversely, the expression of the targeted variants (TFF4 in Fig. 2.5e) decreased to a lesser extent than the non-targeted ones, analogously to what was observed for endogenous miRNAs, albeit with lower efficiency. Finally, to demonstrate the portability of the device we tested the approach in mouse embryonic stem cells (Supplementary Fig. A.25). Here, adaptation to shifts in resource availability was similar to the endogenous miRNA-based regulation (Fig. 2.5c). The model used for fitting and the resulting parameter values are summarized in Supplementary Note A.5 and Supplementary Table A.33. Thus, we showed that also in entirely synthetic systems, adaptation to shifts in resource availability was achieved. To ensure that the observed mitigation was not caused by a higher tolerance to changes in availability at lower expression levels, we showed analytically using the described modeling framework that the normalized expression at lower levels was more sensitive to burden (Supplementary Note A.3).

Indeed, mitigation comes at the cost of the maximal achievable expression levels for the capacity monitor. Moreover, tuning the iFFL circuit to become even less sensitive to changes in available resources will necessarily further limit the maximal expression. This trade-off is intrinsic to the iFFL mitigation strategy. Nevertheless, these results suggest that our approach can be used to mitigate resource-mediated coupling of gene expression despite cell-to-cell variability, demonstrating the portability and broad applicability of our findings. Our results demonstrate that iFFL circuits can mitigate burden from transgene expression in mammalian cells. Importantly, by using miRNAs one can either opt for endogenous miRNAs to specifically tailor a circuit to a desired cell line or create a portable circuit by using a synthetic miRNA such as miR-FF4.

2.4 Discussion

Our study demonstrated that the sharing of limited cellular resources represents a general bottleneck for the predictability and performance of transiently transfected synthetic circuits in mammalian cells, with important consequences for mammalian synthetic biology and biotechnology applications. Due to resource limitations, transient heterologous gene expression results in the coupling of independent exogenous genes and affects the expression of endogenous ones. We presented a detailed characterization of the distinct contributions of transcriptional and translational processes to resource competition and showed that RBPs and miRNAs can redistribute cellular resources thereby alleviating burden. To get a deeper understanding of the mechanisms behind gene expression coupling, we described a modeling framework that captures the indirect interdependence of gene expression in a resource-limited context. Our resource-aware model successfully recapitulated the non-intuitive behavior of the dose responses for the family of controllers described in Fig. 2.4, demonstrating its potential to aid the design of circuits that are less prone to burden effects.

The modeling framework also suggested that an iFFL is a particularly well-suited circuit motif for mitigating burden effects. The iFFL itself is one of the core gene regulatory motifs in biology, and unsurprisingly it has served as inspiration for many synthetic genetic circuits that exploit its adaptation properties (Lillacci et al., 2018; Strovas et al., 2014; Bleris et al., 2011; Segall-Shapiro et al., 2018). In this study we adopt a miRNA implementation of iFFL circuits for the purpose of burden mitigation. Previously, synthetic miRNA-based iFFLs have been demonstrated to increase robustness to gene dosage variability (Lillacci et al., 2018; Bleris et al., 2011) and external perturbations (Strovas et al., 2014). In contrast to synthetic miRNAs, endogenous miRNAs have seen far more limited use in synthetic circuits (e.g. as inputs to synthetic cell-type classifiers (Z. Xie et al., 2011; Matsuura et al., 2018)). Regardless of their origin, miRNA-based iFFL circuits were shown here to decouple the expression of both exogenous and endogenous genes. We speculate that this positive effect is attributed to the freeing up of translational resources, leading to an increase in the expression of proteins involved in the transcription of endogenous genes. At the same time, as already proposed in Gambardella et al. (Gambardella et al., 2017), the downregulation of mKate by miRNAs may lead to a "queueing effect" for the degradation of the other mRNAs, similar to what was shown with two independent proteins tagged for degradation by the proteasome (Cookson et al., 2011).

An implementation of iFFL could alternatively be achieved using RBPs (e.g. L7Ae and Ms2-cNOT7), or using endoribonucleases as is done in a concurrent study by Jones et al. (Jones et al., 2020). Here, we opted for a miRNA-based approach (both endogenous and synthetic) due to several considerations. RBPs impose additional burden, limiting their suitability to mitigate burden itself, while miRNAs are endogenously or intronically expressed with the GOI, thus channeling a negligible amount of resources. To achieve minimal load as in our endogenous miRNA-based iFFL, the RBP alone, or the iFFL should be integrated into the genome. However, a single-copy integration may not guarantee burden mitigation, whereas multiple copy integration may constitute itself a new source of burden. Such systems would need to be tested to assess their usefulness for burden mitigation. Moreover, RBPs rely on specific binding sites that are not as easy to tune as miRNA TS. Lastly, miRNA circuits do not use genetic components that derive from different organisms, circumventing potential toxicity and immunogenicity concerns that could limit their application in medical therapy (Charlesworth et al., 2019; You et al., 2019).

At the same time, miRNAs offer several inherent benefits. Specifically, iFFL circuits that exploit endogenous miRNAs enable cell-type specificity (Z. Xie et al., 2011; Miki et al., 2015), whereas synthetic miRNAs enable portability of circuits across different cell lines. Furthermore, flexibility at the sequence level allows scaling up to many orthogonally operating circuits. The specificity of a miRNA can be easily engineered to target any synthetic or endogenous gene without the need to engineer the target itself (Geisler & Fechner, 2016; Seyhan, 2016; Laganà et al., 2014) (programmability). Finally, tunability of repression strength can be easily achieved both through the number and the placement of the targets, and can be used to enhance adaptation to variations in resource availability. It should be noted that stronger repression will yield lower expression levels of the gene of interest (GOI, Fig. 2.5c). This trade-off is unavoidable, and is an inherent limitation to all implementations of iFFL-based burden-mitigation circuits, including endoRNase implementations.

A potential limitation of miRNA-based iFFL circuits is the diversion of endogenous miRNAs from native targets to synthetic ones. Although this is not what we observe in our miR-31 iFFL (Supplementary Fig. A.24) this may however give rise to an inevitable trade-off similar to what has been observed for competing endogenous RNA (ceRNA). ceRNAs are known to naturally regulate other RNAs by competing for miRNA-binding. To attempt to remedy this, one could use partially complementary TS, which would decrease the affinity of the miRNA to the target and diminish the competition. However this would make the system less efficient and potentially decrease the mitigation effect. Alternatively, the incorporation of multiple TS that respond to different highly expressed miRNAs would distribute the competition between multiple miRNAs and reduce the detrimental effects on their native targets.

Besides iFFLs, negative feedback motifs (Lillacci et al., 2018; Nevozhay et al., 2013; Guinn & Balázsi, 2019; Aoki et al., 2019; Becskei & Serrano, 2000) can also be used to mitigate resource burden, as was shown in a series of studies in Escherichia coli (Ceroni et al., 2015; Qian et al., 2017; Huang et al., 2018; Ceroni et al., 2018; Boo et al., 2019). While negative feedback circuits possess well-established robustness properties, iFFL circuits have several advantages for burden mitigation. In particular, iFFL circuits are considerably simpler to implement and easier to tune than negative feedback circuits, which usually require more components and can become dynamically unstable if not properly designed and tuned. In terms of dynamic response, iFFL circuits are also generally faster in rejecting disturbances like a sudden change in resource availability. Indeed iFFL regulation responds to the disturbance itself, while negative feedback begins to act only after the impact of the disturbance on the regulated output has been detected.

In this study we characterized the contribution of transcriptional and translational processes to resource competition in mammalian cells. In yeast, it was previously reported that squelching, a shortage of general transcription factors, is responsible for the evolutionary breakdown of synthetic gene circuits following exogenous gene expression under a rtTAresponsive promoter (Gouda et al., 2019). However, a deeper understanding of similar effects in mammalian cells is currently lacking. For example, the activation domain of tTA, VP16, interacts with essential components of the transcription machinery such as TFIIB, TFIID, TFIIH, and dTAFII40 (Baron et al., 1997), whose abundances or sub-compositions are unknown and may vary widely across cells (Gelev et al., 2014; Dikstein et al., 1996). Uncovering the key players responsible for gene coupling and endogenous genes' dysregulation will enable the implementation of even more robust and resource-aware solutions to mitigate gene expression burden.

Ultimately, the goal of gene circuit engineering is the creation of cell lines that stably express circuits of interest. Although the presented work focused on the effects of limited resources as induced by transient transfection, it would be natural to investigate if similar effects also occur in the context of genomic integration of highly expressed genes. Moreover, while using a transiently transfected capacity monitor enables the quantification of cellular expression capacity by providing a comparative measure of the geometric mean of free resources in a burdened population relative to a minimally burdened baseline population, stable integration of the capacity monitor would permit a more direct measure in terms of arithmetic mean of free cellular resources (Supplementary Note A.7).

Understanding the impact of resource availability during the engineering of biological systems will have important consequences for biological studies and for improved mammalian cell engineering. For example, studies of biological functions that employ perturbations by exogenous gene expression often lack accuracy and exhibit highly variable results due to less-thanoptimal genetic circuit designs. Using burden-aware designs, cell therapies that rely on finely tuned expression and secretion of therapeutic molecules can now be engineered with resourceaware circuits. Our findings suggest that, when choosing a host cell line, one of the key factors to consider should be its transcriptional and translational capacity (Kafri et al., 2016) not only in terms of productivity but also in terms of the ability of the cells to maintain their fitness while performing their engineered function. Our study presents a portable design capable of enhancing the insulation of transgene expression and will thus contribute to the development of robust-by-design mammalian synthetic circuits, with important implications for basic science and applications in industrial biotechnology and medical therapy.

2.5 Methods

Cell culture

HEK293T, U2OS, and HeLa cells (all from the ATCC) used in this study were maintained in Dulbecco's modified Eagle medium (DMEM, Gibco); H1299 (ATCC) were maintained in Roswell Park Memorial Institute medium (RPMI, Gibco); CHO-K1 were maintained in minimum essential medium α (α-MEM, Gibco). All media were supplemented with 10 % FBS (Atlanta BIO), 1 % penicillin/streptomycin/L-glutamine (Sigma-Aldrich), and 1 % non-essential amino acids (HyClone). HEK239T cells (ATCC, strain number CRL-3216) used for part of this study were maintained in DMEM (Sigma-Aldrich or Gibco) supplemented with 10% FBS (Sigma-Aldrich), 1× GlutaMAX (Gibco) and 1 mM Sodium Pyruvate (Gibco). E14 mouse embryonic stem (mES) (a kind gift from Dr. Maaike Welling) cells were grown in DMEM (Gibco) supplemented with 15 % FBS (PAN Biotech; specifically for ES cell culture), 1 % penicillin/streptomycin (Sigma-Aldrich), 1 % non-essential amino acids (Gibco), 2 mM L-glutamine (GlutaMAX; Gibco), 0.1 mM beta-mercaptoethanol (Sigma-Aldrich), and 100 U mL^{-1} Leukemia inhibitory factor (LIF; Preprotech). At every passage the media was additionally supplemented with fresh CHIR99021 to 3 µM and PD0390125 to 1 µM to support naïve pluripotency (2i conditions (Ying et al., 2008)). All labware used was coated with 0.1 % gelatin (prepared ourselves) prior to plating the ES cells. The cells were maintained at 37 °C and 5 % CO_2.

Transfection

Transfections were carried out in a 24-well plate for flow cytometry analysis or in a 12-well plate format for flow cytometry and qPCR analysis run on the same biological replicates (Table A.1). Transfections for Fig. 2.2d and Supplementary Figs. A.5 and A.22 were carried out in 6 cm dishes. H1299, HeLa, U2OS, HEK293T, and CHO-K1 cells were transfected with Lipofectamine® 3000 (ThermoFisher Scientific) according to the manufacturer's instructions and 300 ng total DNA (500 ng in Fig. 2.2c, d and Supplementary Figs. A.1, A.8a, A.9a, A.10a, and A.11a) in 24-well plates. DNA and transfection reagents were scaled up according to the Lipofectamine® 3000 manufacturer's instructions. miR-31 inhibitor (Invitrogen™ mirVana™ miRNA Inhibitors) was co-transfected using the same method as for DNA (Fig. 2.3c). HEK293T cells used for experiments shown in Figs. 2.2a, b, e, 2.4 and 2.5e were plated approximately 24 h before transfection at 62,500–75,000 cells per well in 24-well plates. The transfection solution was prepared using polyethylenimine (PEI) "MAX" (Mw 40,000, Polysciences, Inc.) in a 1:3 (µg DNA to µg PEI) ratio with a total of 500 ng of plasmid

DNA per well. Both DNA and PEI were diluted in OptiMEM I reduced serum media (Gibco) before being mixed and incubated for 25 min prior to addition to the cells. E14 mouse embryonic stem cells were transfected using Lipofectamine® 2000 (ThermoFisher Scientific) in a 1:3 (µg DNA to µg Lipofectamine® 2000) with 300 ng of plasmid DNA per well. The transfection was performed on cells in suspension immediately after plating at approximately 30,000 cells per well. All wells were coated with 0.1 % gelatin before the addition of the cells.

Flow cytometry and data analysis

H1299, HEK293T, U2OS, HeLa, and CHO-K1 cells were analyzed with a BD Facsaria™ cell analyzer (BD Biosciences) or BD Celesta™ cell analyzer (BD Biosciences) using 488 and 561 lasers. For each sample >20,000 singlet events were collected and fluorescence data were acquired with the following cytometer settings: 488 nm laser and 530/30 nm bandpass filter for EGFP, 561 nm laser and 610/20 nm filter for mKate. Cells transfected in 12-well plates were washed with DPBS, detached with 100 µL of Trypsin-EDTA (0.25 %), and resuspended in 600 µL of DPBS (Thermo Fisher). Two hundred microliters of cell suspension were used for flow cytometry and 400 µL for RNA extraction. HEK293T used for experiments shown in Figs. 2.2a, b, e, 2.4 and 2.5e cells were measured 48 h after transfection on a BD LSRFortessa™ Special Order and Research Product (SORP) cell analyzer. mCitrine fluorescence was excited via a 488 nm laser and was detected through a 530/11 nm bandpass filter. mRuby3 was excited via a 561 nm laser and measured through a 610/20 nm bandpass filter. miRFP670 was excited at 640 nm and measured through a 670/14 nm bandpass filter. E14 mES cells were measured 48 h after transfection on a Beckman Coulter CytoFLEX S flow cytometer. mCitrine fluorescence was excited using a 488 nm laser and was detected through a 525/40+OD1 bandpass filter. mRuby3 was excited with 561 nm laser light and measured through a 610/20+OD1 bandpass filter. miRFP670 was excited at 638 nm and measured through a 660/10 bandpass filter. The cells were collected for measurement by washing with DBPS (Sigma-Aldrich or Gibco) and detaching in 70 µL to 180 µL of Accutase solution (Sigma-Aldrich). For each sample between 10,000 and 200,000 singlet events were collected. Fluorescence intensity in arbitrary units (a.u.) was used as a measure of protein expression. For each experiment a compensation matrix was created using unstained (wild type cells), and single-color controls (mKate/mCherry only, EGFP only or mCitrine only, mRuby3 only, miRFP670 only). Live cell population and single cells were selected according to FCS/SSC parameters (Supplementary Figs. A.26 and A.27). Data analysis was performed with Cytoflow or a custom R script. Data fitting was performed using Mathematica’s NonlinearModelFit function and the InteriorPoint method.

Cell sorting

H1299 cells used for the experiment shown in Fig. 2.2d were trypsinized from 6 cm dishes and counted. They were then centrifuged at 500 g for 5 min and resuspended at a concentration of 5 mln mL^{-1} in sorting buffer (PBS 1× + 3 mM EDTA + 0.8 % Trypsin + 1 % FBS). Cells were sorted with a BD FACSMelody™ cell sorter according to their fluorescence levels (Supplementary Fig. A.5). In total, 150,000 cells per gate were collected.

DNA cloning and plasmid construction

Plasmid vectors carrying gene cassettes were created using In-Fusion HD cloning kit (Clonetch), Gibson Assembly, via digestion and ligation or using the yeast toolkit (YTK) (Lee et al., 2015) with custom parts for mammalian cells. Gibson Assembly master mixes were created from Taq DNA Ligase (NEB), Phusion High-Fidelity DNA Polymerase (NEB), and T5 Exonuclease (Epicentre) in 5× isothermal buffer (Supplementary Table A.27). Ligation reactions were performed in 1:2–5 molar ratios of plasmid backbone:gene insert starting with 50 ng to 100 ng of vector backbone digested with selected restriction enzymes. Assemblies using the YTK were performed according to the original publication (Lee et al., 2015). Newly created constructs were transformed into XL10-Gold or TOP10 *E. coli* strains. For plasmids with miRNA TS, the target sequences were selected using miRBase database (http://www.mirbase.org/) and are listed in Supplementary Table A.25. List of oligos used to clone endogenous miRNAs TS are listed in Supplementary Table A.24. All plasmids were confirmed by sequencing analysis and deposited to addgene. To perform western blot analysis, an His-tag composed of six Histidine residues was inserted after the start codon of mKate encoding plasmids.

mRNA extraction and reverse transcription

RNA extraction was performed with E.Z.N.A.® Total RNA Kit I (Omega Biotek). The protocol was followed according to manufacturer's instructions and RNA was eluted in 30 µL of RNAse free water. RNA samples were conserved at -80 °C. PrimeScript RT Reagent Kit with gDNA Eraser—Perfect Real Time (Takara) was used according to the manufacturer's instructions. The protocol was performed on ice in a RNAse free environment to avoid RNA degradation. A negative control without PrimeScript RT Enzyme Mix I was always prepared to investigate genomic DNA contamination.

qPCR

Fast SYBR Green Master Mix (ThermoFisher Scientific) was used to perform qPCR of cDNAs obtained from 500 ng of RNA and diluted 1:5. Samples were loaded in MicroAmp™ Fast Optical 96-Well Reaction Plate (0.1 mL) and the experiment was carried out with a CFX96 Touch Real-Time PCR Detection System (BioRad) machine. Each well contained 20 µL of final volume (7 µL SYBR Green Master Mix, 10 µL ddHO, 1 µL of each primer, 1 µL of template). Also, a control without template (blank) was set. Primers were designed to amplify a region of 60 bp to 200 bp (Supplementary Table A.26) and with a temperature of annealing between 50 °C and 65 °C. Data were analyzed using the comparative Ct method according to the Applied Biosystems Protocols.

Statistics and reproducibility

Each experiment was repeated independently at least twice with similar results, with the exception of Supplementary Fig. A.2 and condition w/o Mitigation, 1.5 equimolar EGFP to mKate plasmid in Supplementary Fig. A.20. All models used for parameter fitting are contained in Supplementary Notes A.4–A.6. The obtained parameter values are summarized in Supplementary Table A.28–A.38.

2.6 Bibliography

Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., & Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, *11*(1), 1–14.

Brinkman, B., Zuijdeest, D., Kaijzel, E. L., Breedveld, F. C., & Verweij, C. L. (1995). Relevance of the tumor necrosis factor alpha (tnf alpha)-308 promoter polymorphism in tnf alpha gene regulation. *Journal of inflammation*, *46*(1), 32–41.

Bamforth, S. D., Bragança, J., Eloranta, J. J., Murdoch, J. N., Marques, F. I., Kranc, K. R., Farza, H., Henderson, D. J., Hurst, H. C., & Bhattacharya, S. (2001). Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking cited2, a new tfap2 co-activator. *Nature genetics*, *29*(4), 469–474.

Farquhar, K. S., Charlebois, D. A., Szenk, M., Cohen, J., Nevozhay, D., & Balázsi, G. (2019). Role of network-mediated stochasticity in mammalian drug resistance. *Nature communications*, *10*(1), 1–14.

Liu, W., Dong, X., Mai, M., Seelan, R. S., Taniguchi, K., Krishnadath, K. K., Halling, K. C., Cunningham, J. M., Qian, C., Christensen, E., et al. (2000). Mutations in axin2 cause colorectal cancer with defective mismatch repair by activating β-catenin/tcf signalling. *Nature genetics*, *26*(2), 146–147.

Stuible, M., Burlacu, A., Perret, S., Brochu, D., Paul-Roc, B., Baardsnes, J., Loignon, M., Grazzini, E., & Durocher, Y. (2018). Optimization of a high-cell-density polyethylenimine transfection method for rapid protein production in cho-ebna1 cells. *Journal of biotechnology*, *281*, 39–47.

Xie, M., Ye, H., Wang, H., Charpin-El Hamri, G., Lormeau, C., Saxena, P., Stelling, J., & Fussenegger, M. (2016). β-cell–mimetic designer cells provide closed-loop glycemic control. *Science*, *354*(6317), 1296–1301.

Siciliano, V., DiAndreth, B., Monel, B., Beal, J., Huh, J., Clayton, K. L., Wroblewska, L., McKeon, A., Walker, B. D., & Weiss, R. (2018). Engineering modular intracellular protein sensor-actuator devices. *Nature communications*, *9*(1), 1–7.

Caliendo, F., Dukhinova, M., & Siciliano, V. (2019). Engineered cell-based therapeutics: Synthetic biology meets immunology. *Frontiers in bioengineering and biotechnology*, *7*, 43.

MacDonald, J. T., & Siciliano, V. (2017). Computational sequence design with r2odna designer. *Mammalian synthetic promoters* (pp. 249–262). Springer.

di Bernardo, D., Marucci, L., Menolascina, F., & Siciliano, V. (2012). Predicting synthetic gene networks. *Synthetic gene networks* (pp. 57–81). Springer.

Xiang, Y., Dalchau, N., & Wang, B. (2018). Scaling up genetic circuit design for cellular computing: Advances and prospects. *Natural computing*, *17*(4), 833–853.

Borkowski, O., Ceroni, F., Stan, G.-B., & Ellis, T. (2016). Overloaded and stressed: Whole-cell considerations for bacterial synthetic biology. *Current opinion in microbiology*, *33*, 123–130.

Ceroni, F., Algar, R., Stan, G.-B., & Ellis, T. (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature methods*, *12*(5), 415–418.

Li, G.-W., Burkhardt, D., Gross, C., & Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, *157*(3), 624–635.

Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z., & Hwa, T. (2010). Interdependence of cell growth and gene expression: Origins and consequences. *Science*, *330*(6007), 1099–1102.

Carbonell-Ballestero, M., García-Ramallo, E., Montañez, R., Rodríguez-Caso, C., & Macía, J. (2016). Dealing with the genetic load in bacterial synthetic biology circuits: Convergences with the ohm's law. *Nucleic acids research*, *44*(1), 496–507.

Qian, Y., Huang, H.-H., Jiménez, J. I., & Del Vecchio, D. (2017). Resource competition shapes the response of genetic circuits. *ACS synthetic biology*, *6*(7), 1263–1272.

Weiße, A. Y., Oyarzún, D. A., Danos, V., & Swain, P. S. (2015). Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, *112*(9), E1038–E1047.

Gyorgy, A., Jiménez, J. I., Yazbek, J., Huang, H.-H., Chung, H., Weiss, R., & Del Vecchio, D. (2015). Isocost lines describe the cellular economy of genetic circuits. *Biophysical journal*, *109*(3), 639–646.

Huang, H.-H., Qian, Y., & Del Vecchio, D. (2018). A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nature communications*, *9*(1), 1–12.

Ceroni, F., Boo, A., Furini, S., Gorochowski, T. E., Borkowski, O., Ladak, Y. N., Awan, A. R., Gilbert, C., Stan, G.-B., & Ellis, T. (2018). Burden-driven feedback control of gene expression. *Nature methods*, *15*(5), 387–393.

Kafri, M., Metzl-Raz, E., Jona, G., & Barkai, N. (2016). The cost of protein production. *Cell reports*, *14*(1), 22–31.

Gouda, M. K., Manhart, M., & Balázsi, G. (2019). Evolutionary regain of lost gene circuit function. *Proceedings of the National Academy of Sciences*, *116*(50), 25162–25171.

Lillacci, G., Benenson, Y., & Khammash, M. (2018). Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic acids research*, *46*(18), 9855–9863.

Chou, T.-C., & Talaly, P. (1977). A simple generalized equation for the analysis of multiple inhibitions of michaelis-menten kinetic systems. *Journal of Biological Chemistry*, *252*(18), 6438–6442.

Rondelez, Y. (2012). Competition for catalytic resources alters biological network dynamics. *Physical review letters*, *108*(1), 018102.

Siciliano, V., Garzilli, I., Fracassi, C., Criscuolo, S., Ventre, S., & Di Bernardo, D. (2013). Mirnas confer phenotypic robustness to gene networks by suppressing biological noise. *Nature communications*, *4*(1), 1–7.

Carignano, A., Mukherjee, S., Singh, A., & Seelig, G. (2018). Extrinsic noise suppression in micro rna mediated incoherent feedforward loops. *2018 IEEE Conference on Decision and Control (CDC)*, 4353–4359.

Strovas, T. J., Rosenberg, A. B., Kuypers, B. E., Muscat, R. A., & Seelig, G. (2014). Microrna-based single-gene circuits buffer protein synthesis rates against perturbations. *ACS synthetic biology*, *3*(5), 324–331.

Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., Xiang, A. P., Ren, B.-Z., & Lahn, B. T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PloS one*, *5*(5), e10611.

Lu, J., Zhang, F., & Kay, M. A. (2013). A mini-intronic plasmid (mip): A novel robust transgene expression vector in vivo and in vitro. *Molecular Therapy*, *21*(5), 954–963.

Kallehauge, T. B., Li, S., Pedersen, L. E., Ha, T. K., Ley, D., Andersen, M. R., Kildegaard, H. F., Lee, G. M., & Lewis, N. E. (2017). Ribosome profiling-guided depletion of an mrna increases cell growth rate and protein secretion. *Scientific reports*, *7*(1), 1–12.

Wroblewska, L., Kitada, T., Endo, K., Siciliano, V., Stillo, B., Saito, H., & Weiss, R. (2015). Mammalian synthetic circuits with rna binding proteins for rna-only delivery. *Nature biotechnology*, *33*(8), 839–841.

Cella, F., Wroblewska, L., Weiss, R., & Siciliano, V. (2018). Engineering protein-protein devices for multilayered regulation of mrna translation using orthogonal proteases in mammalian cells. *Nature communications*, *9*(1), 1–9.

Gam, J. J., Babb, J., & Weiss, R. (2018). A mixed antagonistic/synergistic mirna repression model enables accurate predictions of multi-input mirna sensor activity. *Nature communications*, *9*(1), 1–12.

Ameres, S. L., & Zamore, P. D. (2013). Diversifying microrna sequence and function. *Nature reviews Molecular cell biology*, *14*(8), 475–488.

Meleady, P., Doolan, P., Henry, M., Barron, N., Keenan, J., O'Sullivan, F., Clarke, C., Gammell, P., Melville, M. W., Leonard, M., et al. (2011). Sustained productivity in recombinant chinese hamster ovary (cho) cell lines: Proteome analysis of the molecular basis for a process-related phenotype. *BMC biotechnology*, *11*(1), 1–11.

Alles, J., Fehlmann, T., Fischer, U., Backes, C., Galata, V., Minet, M., Hart, M., Abu-Halima, M., Grässer, F. A., Lenhof, H.-P., et al. (2019). An estimate of the total number of true human mirnas. *Nucleic acids research*, *47*(7), 3353–3364.

Thomson, D. W., Bracken, C. P., & Goodall, G. J. (2011). Experimental strategies for microrna target identification. *Nucleic acids research*, *39*(16), 6845–6853.

Robertson, B., Dalby, A. B., Karpilow, J., Khvorova, A., Leake, D., & Vermeulen, A. (2010). Specificity and functionality of microrna inhibitors. *Silence*, *1*(1), 1–9.

Bleris, L., Xie, Z., Glass, D., Adadey, A., Sontag, E., & Benenson, Y. (2011). Synthetic incoherent feedforward circuits show adaptation to the amount of their genetic template. *Molecular systems biology*, *7*(1), 519.

Zechner, C., Seelig, G., Rullan, M., & Khammash, M. (2016). Molecular circuits for dynamic noise filtering. *Proceedings of the National Academy of Sciences*, *113*(17), 4729–4734.

Tian, W., Dong, X., Liu, X., Wang, G., Dong, Z., Shen, W., Zheng, G., Lu, J., Chen, J., Wang, Y., et al. (2012). High-throughput functional micrornas profiling by recombinant aav-based microrna sensor arrays. *PloS one*, *7*(1), e29551.

Aprelikova, O., Yu, X., Palla, J., Wei, B.-R., John, S., Yi, M., Stephens, R., Simpson, R. M., Risinger, J. I., Jazaeri, A., et al. (2010). The role of mir-31 and its target gene satb2 in cancer-associated fibroblasts. *Cell cycle*, *9*(21), 4387–4398.

Segall-Shapiro, T. H., Sontag, E. D., & Voigt, C. A. (2018). Engineered promoters enable constant gene expression at any copy number in bacteria. *Nature biotechnology*, *36*(4), 352.

Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., & Benenson, Y. (2011). Multi-input rnai-based logic circuit for identification of specific cancer cells. *Science*, *333*(6047), 1307–1311.

Matsuura, S., Ono, H., Kawasaki, S., Kuang, Y., Fujita, Y., & Saito, H. (2018). Synthetic rna-based logic computation in mammalian cells. *Nature communications*, *9*(1), 1–8.

Gambardella, G., Carissimo, A., Chen, A., Cutillo, L., Nowakowski, T. J., Di Bernardo, D., & Blelloch, R. (2017). The impact of micrornas on transcriptional heterogeneity and gene co-expression across single embryonic stem cells. *Nature communications*, *8*(1), 1–11.

Cookson, N. A., Mather, W. H., Danino, T., Mondragón-Palomino, O., Williams, R. J., Tsimring, L. S., & Hasty, J. (2011). Queueing up for enzymatic processing: Correlated signaling through coupled degradation. *Molecular systems biology*, *7*(1), 561.

Jones, R. D., Qian, Y., Siciliano, V., DiAndreth, B., Huh, J., Weiss, R., & Del Vecchio, D. (2020). An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nature communications*, *11*(1), 1–16.

Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Camarena, J., Lemgart, V. T., Cromer, M. K., Vakulskas, C. A., Collingwood, M. A., Zhang, L., Bode, N. M., et al. (2019). Identification of preexisting adaptive immunity to cas9 proteins in humans. *Nature medicine*, *25*(2), 249–254.

You, L., Tong, R., Li, M., Liu, Y., Xue, J., & Lu, Y. (2019). Advancements and obstacles of crispr-cas9 technology in translational research. *Molecular Therapy-Methods & Clinical Development*, *13*, 359–370.

Miki, K., Endo, K., Takahashi, S., Funakoshi, S., Takei, I., Katayama, S., Toyoda, T., Kotaka, M., Takaki, T., Umeda, M., et al. (2015). Efficient detection and purification of cell populations using synthetic microrna switches. *Cell stem cell*, *16*(6), 699–711.

Geisler, A., & Fechner, H. (2016). Microrna-regulated viral vectors for gene therapy. *World journal of experimental medicine*, *6*(2), 37.

Seyhan, A. A. (2016). A multiplexed mirna and transgene expression platform for simultaneous repression and expression of protein coding sequences. *Molecular BioSystems*, *12*(1), 295–312.

Laganà, A., Acunzo, M., Romano, G., Pulvirenti, A., Veneziano, D., Cascione, L., Giugno, R., Gasparini, P., Shasha, D., Ferro, A., et al. (2014). Mir-synth: A computational resource for the design of multi-site multi-target synthetic mirnas. *Nucleic acids research*, *42*(9), 5416–5425.

Nevozhay, D., Zal, T., & Balázsi, G. (2013). Transferring a synthetic gene circuit from yeast to mammalian cells. *Nature communications*, *4*(1), 1–11.

Guinn, M. T., & Balázsi, G. (2019). Noise-reducing optogenetic negative-feedback gene circuits in human cells. *Nucleic acids research*, *47*(14), 7703–7714.

Aoki, S. K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., & Khammash, M. (2019). A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, *570*(7762), 533–537.

Becskei, A., & Serrano, L. (2000). Engineering stability in gene networks by autoregulation. *Nature*, *405*(6786), 590–593.

Boo, A., Ellis, T., & Stan, G.-B. (2019). Host-aware synthetic biology. *Current Opinion in Systems Biology*, *14*, 66–72.

Baron, U., Gossen, M., & Bujard, H. (1997). Tetracycline-controlled transcription in eukaryotes: Novel transactivators with graded transactivation potential. *Nucleic acids research*, *25*(14), 2723–2729.

Gelev, V., Zabolotny, J. M., Lange, M., Hiromura, M., Yoo, S. W., Orlando, J. S., Kushnir, A., Horikoshi, N., Paquet, E., Bachvarov, D., et al. (2014). A new paradigm for transcription factor tfiib functionality. *Scientific reports*, *4*(1), 1–8.

Dikstein, R., Zhou, S., & Tjian, R. (1996). Human tafii105 is a cell type–specific tfiid subunit related to htafii130. *Cell*, *87*(1), 137–146.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., & Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *nature*, *453*(7194), 519–523.

Lee, M. E., DeLoache, W. C., Cervantes, B., & Dueber, J. E. (2015). A highly characterized yeast toolkit for modular, multipart assembly. *ACS synthetic biology*, *4*(9), 975–986.

Chapter Three

A genetic mammalian proportional-integral feedback control circuit for robust and precise gene regulation

This chapter represents a manuscript submitted to the Proceedings of the National Academy of Sciences (PNAS). An earlier version was published as a preprint article on the bioRxiv preprint server under a Creative Commons Attributions-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0; https://creativecommons.org/licenses/by-nc-nd/4.0/)

Frei, T., Chang, C.-H., Filo, M., Arampatzis, A., & Khammash, M. (2021). Genetically engineered proportional-integral feedback controllers for robust perfect adaptation in mammalian cells. *bioRxiv*. https://doi.org/10.1101/2020.12.06.412304

3.1 Abstract

The processes that keep a cell alive are constantly challenged by unpredictable changes in its environment. Cells manage to counteract these changes by employing sophisticated regulatory strategies that maintain a steady internal milieu. Recently, the antithetic integral feedback motif has been demonstrated to be a minimal and universal biological regulatory strategy that can guarantee robust perfect adaptation for noisy gene regulatory networks in *E. coli*. Here, we present the first realization of the antithetic integral feedback motif in a synthetic gene circuit in mammalian cells. We show that the motif robustly main-

tains the expression of a synthetic transcription factor at tunable levels even when it is perturbed by increased degradation or its interaction network structure is perturbed by a negative feedback loop with an RNA-binding protein. We further demonstrate an improved regulatory strategy by augmenting the antithetic integral motif with additional negative feedback to realize antithetic proportional-integral control. We show that this motif produces robust perfect adaptation while also reducing the variance of the regulated synthetic transcription factor. We demonstrate that the integral and proportional-integral feedback motifs can mitigate the impact of gene expression burden and we computationally explore their use in cell therapy. We believe that the engineering of precise and robust perfect adaptation will enable substantial advances in industrial biotechnology and cell-based therapeutics.

3.2 Significance Statement

To survive in the harsh environments they inhabit, cells have evolved sophisticated regulatory mechanisms that can maintain a steady internal milieu or homeostasis. This robustness, however, does not generally translate to engineered genetic circuits, such as the ones studied by synthetic biology. Here, we introduce an implementation of a minimal and universal gene regulatory motif that produces robust perfect adaptation for mammalian cells and we improve on it by enhancing the precision of its regulation.

3.3 Introduction

The ability to maintain a steady internal environment in the presence of a changing and uncertain external environment — called homeostasis — is a defining characteristic of living systems (Urry et al., 2017). Homeostasis is maintained by various regulatory mechanisms, often in the form of negative feedback loops. The importance of homeostasis is clearly exemplified in physiology and medicine, where a loss of homeostasis is often attributed to the development of disease (Kotas & Medzhitov, 2015; Mullur et al., 2014; Yu et al., 2015).

Feedback control systems in engineering use the error or more specifically the difference between the desired output — commonly referred to as the setpoint — and the current output of the system which is to be regulated to determine the effort that the control system applies to steer the system under control. In a simple proportional feedback system the effort is determined by the instantaneous difference. Therefore, if the difference becomes zero when the desired output is reached no effort is applied and the output is free to deviate from the desired state. Compared to this simpler proportional feedback, integral feedback does not just feed back the instantaneous difference between the desired state and the current state, but uses the entire history of the difference to determine the control effort applied (see Figure 3.1(d)). This difference may build up and will provide control effort even as the error has decayed to zero. Therefore it can guarantee that a zero difference between desired output and current output is achieved and adaptation is perfect (Åström & Murray, 2010). A system with integral feedback is known to reject constant disturbances in parameters and structure of the system under control and is also able to perfectly track a desired output commonly referred to as the setpoint. More recently, it has become increasingly evident that integral feedback is a regulatory strategy that drives biological adaptation in different systems (Yi et al., 2000; El-Samad et al., 2002; Miller & Wang, 2006; Muzzey et al., 2009; Ben-Zvi & Barkai, 2010). Although integral feedback guarantees robust perfect adaptation, it does not in general prevent large transient deviations. To mediate this, control engineers often augment proportional feedback to their integral feedback control systems. By counteracting such large deviations, proportional-integral feedback also suppresses large stochastic fluctuations around the setpoint and therefore provides preciser regulation than integral feedback can achieve (Briat et al., 2018).

Here, we demonstrate perfect adaptation in a sense/antisense mRNA implementation of the antithetic integral feedback circuit in mammalian cells and show that the resulting closed-loop control system is highly robust to network changes and parameter disturbances. By further incorporating proportional feedback on the sensed output to achieve proportional-integral feedback control, we also increase the precision of the resulting adaptation. Furthermore, we derive a mathematical (mechanistic) model that describes the various interactions in the system. We show that the obtained model fits the experimentally obtained data well, and is also capable of predicting the robustness features of our implementation of the antithetic integral controller. Lastly, we demonstrate the applicability of our integral and proportional-integral controllers by demonstrating perfect mitigation of gene expression burden and show that the proportional-integral controller provides superior precision over integral feedback.

Over the last decade, several experimental studies have constructed genetic systems and cell-based therapies that implement negative feedback to mitigate disease (Kemmer et al., 2010; Wei et al., 2012; Rössger et al., 2013; Xie et al., 2016). These, however, rely solely on proportional feedback rather than integral or proportional-integral feedback and are therefore not guaranteed to achieve precise and robust regulation. In 2016, Briat et al. introduced a biomolecular circuit topology that implements integral feedback control for general biomolecular systems (Briat et al., 2016). Figure 3.1(a) depicts an abstract representation of

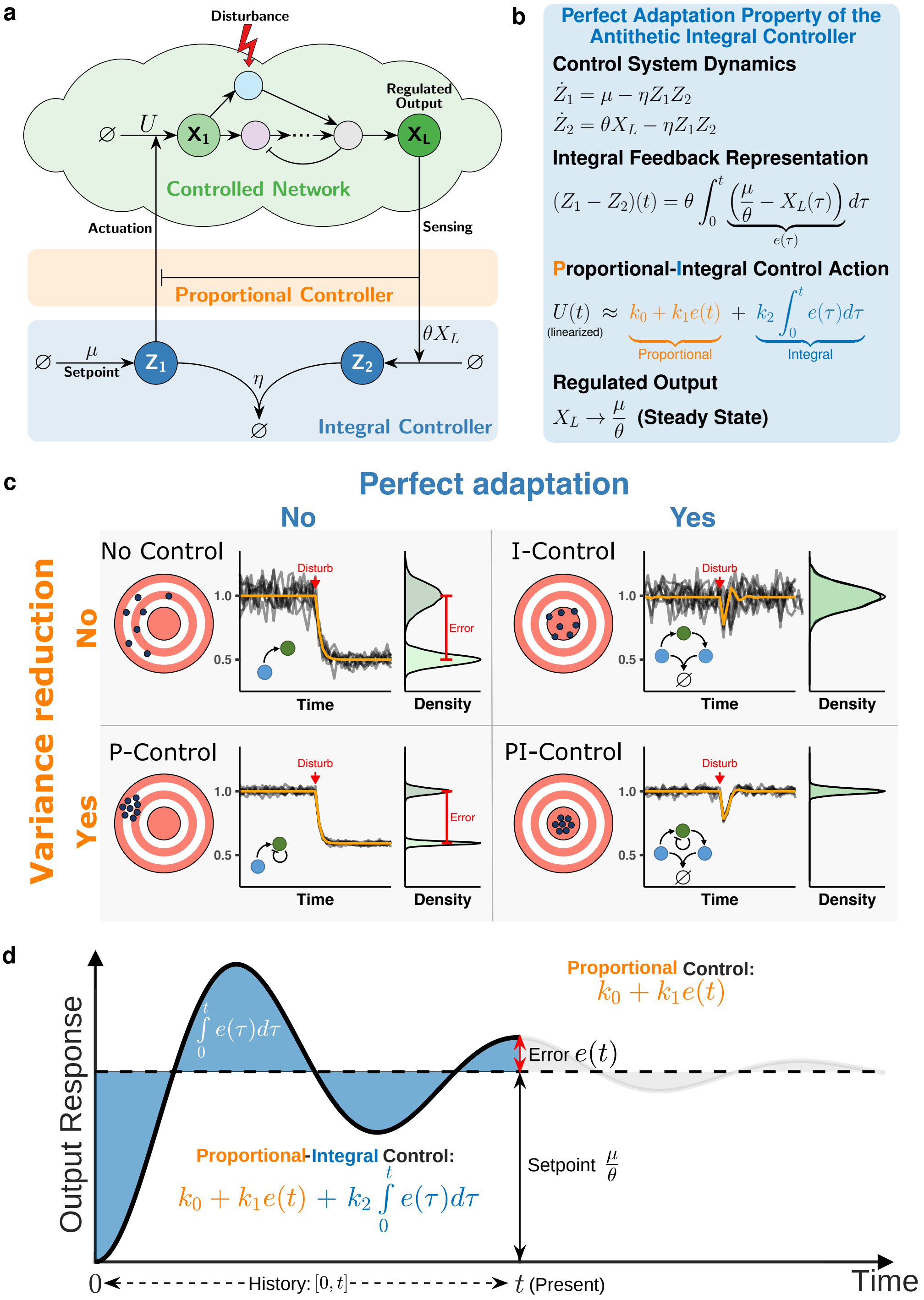


Figure 3.1: See caption on the next page.

Figure 3.1: The antithetic proportional-integral feedback motif. (a) Network topology of an arbitrary molecular network interacting with an antithetic proportional-integral feedback motif. The nodes labelled with Z_1 and Z_2 together compose the antithetic motif responsible for realizing integral feedback. Species Z_1 is produced at a rate μ and is functionally annihilated when it interacts with species Z_2 at a rate η. Furthermore, Z_1 interacts with the controlled network by promoting the production of species X_1. To close the feedback loop, species Z_2 is produced at a reaction rate that is proportional to θ and the regulated output species X_L. An additional negative feedback from the output to the production reaction extends the motif to proportional-integral feedback. (b) Dynamics of the antithetic integral controller. Subtracting the differential equations of Z_1 and Z_2 reveals the integral action of the controller that ensures that the steady state of the output converges to a value that is independent of the controlled network parameters. Additionally, through linearization (Filo & Khammash, 2021), the individual integral and the proportional control actions of the antithetic proportional-integral motif can be expressed separately. (c) The elements of proportional-integral feedback. Without any feedback control, the output of the controlled network may be highly variable and will likely respond drastically to a disturbance in the network. By adding integral feedback, it can be assured that the output will adapt perfectly to disturbances. Conversely, by adding proportional feedback the variability in the output can be reduced. Combining the two types of feedback reduces the variability of the output while also ensuring perfect adaptation. (d) Graphical demonstration of integral and proportional control. Integral control accounts for error history by mathematically integrating it in time. Consequently, integral controllers have memory and ”remember” the past. However, proportional controllers act instantaneously by only accounting for the present error. Consequently, proportional controllers are memory-less and ”forget” the past.

this control motif. A subsequent publication by the same authors showed that additional proportional negative feedback further reduces variance in the controlled output (Briat et al., 2018). Central to this strategy — termed antithetic proportional-integral feedback — is the so-called annihilation (or sequestration) reaction between the two species that implement the controller (reaction with rate η in Figure 3.1(a)). The annihilation refers to the requirement that both controller species abolish each other’s function when they interact. Another stringent requirement to achieve integral feedback is that the two controller species on their own remain fairly stable over time. Given these conditions, any network interconnected in a stable way with this antithetic integral controller will achieve robust adaptation (Figure 3.1(c)). The incorporation of additional proportional negative feedback from the output of the controlled network to the actuation reaction then yields proportional-integral feedback (Figure 3.1(a)). Independent of integral feedback, this proportional feedback introduces a reduction in the variance of the controlled output (Figure 3.1(c)).

The initial theoretical work has motivated the implementation of antithetic integral control in bacteria (Lillacci et al., 2017; Aoki et al., 2019) and in vitro (Agrawal et al., 2019). A quasi-integral controller in *E. coli* (Huang et al., 2018) also relies on a similar topology. In realizing antithetic integral feedback, one of the main challenges is identifying a suitable implementation of the annihilation (or sequestration) reaction (Aoki et al., 2019). In the bacterial implementation of the antithetic integral feedback motif (Aoki et al., 2019) stable proteins (a σ and anti-σ factor pair) were used to realize the sequestration reaction. However, this approach is not directly applicable to mammalian cells. Instead, in this work we exploit hybridization of complementary mRNAs to realize this critical reaction. (Figure 3.2(a)). For the antithetic integral controller to function properly, the sense and antisense RNAs have to be stable such that their degradation is predominantly due to their mutual interaction (via the hybridization reaction). Unlike bacterial RNAs where the majority of mRNAs have half-lives between 3 and 8 minutes Milo and Phillips, 2015, mammalian RNAs are much more stable with typical mRNA half lives of several hours Pérez-Ortín et al., 2013. Indeed in human cells, the majority of mRNAs have half-lives between 6 and 18 hours, with an overall mean value of 10 hours (Yang et al., 2003; Schwanhäusser et al., 2011). The hybridization of the mammalian sense/antisense RNAs and their stability allow us to realize the antithetic integral controller in mammalian cells. Sense and antisense mRNA have previously been employed to control gene expression in yeast (Bayer & Smolke, 2005) and to build a genetic oscillator in mammalian cells (Tigges et al., 2009). Furthermore, antisense RNA has shown promise in the treatment of cancer and other genetic diseases as well as infections (Haberman et al., 2002; Fuchs et al., 2004; Tebas et al., 2013).

3.4 Results

Integral Feedback

A schematic depiction of the sense/antisense RNA implementation of the antithetic integral feedback circuit is shown in Figure 3.2(a). The basic circuit consists of two genes, which are encoded on separate plasmids. The gene in the *activator plasmid* is the synthetic transcription factor tTA (tetracycline transactivator) (Gossen & Bujard, 1992) fused to the fluorescent protein mCitrine. The expression of this gene is driven by the strong mammalian EF-1 α promoter. This transcription factor drives the expression of the other gene in the *antisense plasmid* via the tTA-responsive TRE promoter. This gene expresses an antisense RNA that is complementary to the *activator* mRNA. The hybridization of these two species realizes the annihilation reaction and closes the negative feedback loop. As an experimental control incapable of producing integral feedback, we built an open-loop analog of the closed-loop circuit, in which the TRE promoter was replaced by a non-cognate promoter. The closed-loop configuration is set up to regulate the expression levels of the *activator* tTA-mCitrine. To introduce specific perturbations to the *activator* we additionally fused an Asunaprevir

(ASV) inducible degradation tag (SMASh) to tTA-mCitrine (Chung et al., 2015).

To show that our genetic implementation of the circuit performs integral feedback we apply constant disturbances with ASV at a concentration of 0.033 µM to HEK293T cells which were transiently transfected with either the open- or the closed-loop circuit. Additionally, we vary the setpoint by transfecting the two plasmids at ratios ranging from 1/16 to 2 (Activator Plasmid/Antisense Plasmid). The fluorescence of the cells was measured 48 hours after transfection using flow cytometry. As the setpoint ratio increases, so does the fluorescence of tTA-mCitrine, indicating that our circuit permits setpoint control (Figure 3.2(b) and Supplementary Figure B.5). Note that this fluorescence is a monotonically-increasing function of the plasmid ratios (see also the function θ in Figure B.1(b)). We consider a circuit to be adapting if its normalized fluorescence intensity stays within 10% of the undisturbed control. Under this criterion, adaptation is achieved for all the setpoints tested below 2 in the closed-loop configuration. In contrast, none of the open-loop configurations manage to meet this adaptation requirement (Figure 3.2(c)).

Next, we sought to demonstrate that our implementation of the antithetic integral controller will provide disturbance rejection at different setpoints regardless of the network topology it regulates. Therefore, we added a negative feedback loop from tTA-mCitrine to its own production. This negative feedback was realized by the RNA-binding protein L7Ae (Saito et al., 2010), which is expressed under the control of a tTA-responsive TRE promoter and binds the kink-turn hairpin on the sense mRNA to inhibit translation (Figure 3.2(a)). The closed- and open-loop circuits were transiently transfected either with or without this negative feedback plasmid to introduce a perturbation to the regulated network. The setpoints 1/4 and 1/2 were tested by transfecting an appropriate ratio of the *activator* to *antisense* plasmids. These different conditions were further perturbed at the molecular level by adding 0.033 µM ASV to induce degradation of tTA-mCitrine. As shown in Figure 3.2(d) (see also Supplementary Figure B.6), the closed-loop circuit rejects both perturbations nearly perfectly in all cases, whereas again the open-loop circuit fails to adapt.

Proportional-integral feedback

The capability of the antithetic integral controller to reject topological network perturbations, as demonstrated previously in Figure 3.2(d), allowed us to further improve the controller performance by increasing its complexity. In particular, we implement a common control strategy that is extensively applied in various engineering disciplines, referred to as Proportional-Integral (PI) control. This control strategy adds to the Integral (I) controller Proportional (P) feedback action to enhance dynamic performance, such as transient dynamics and

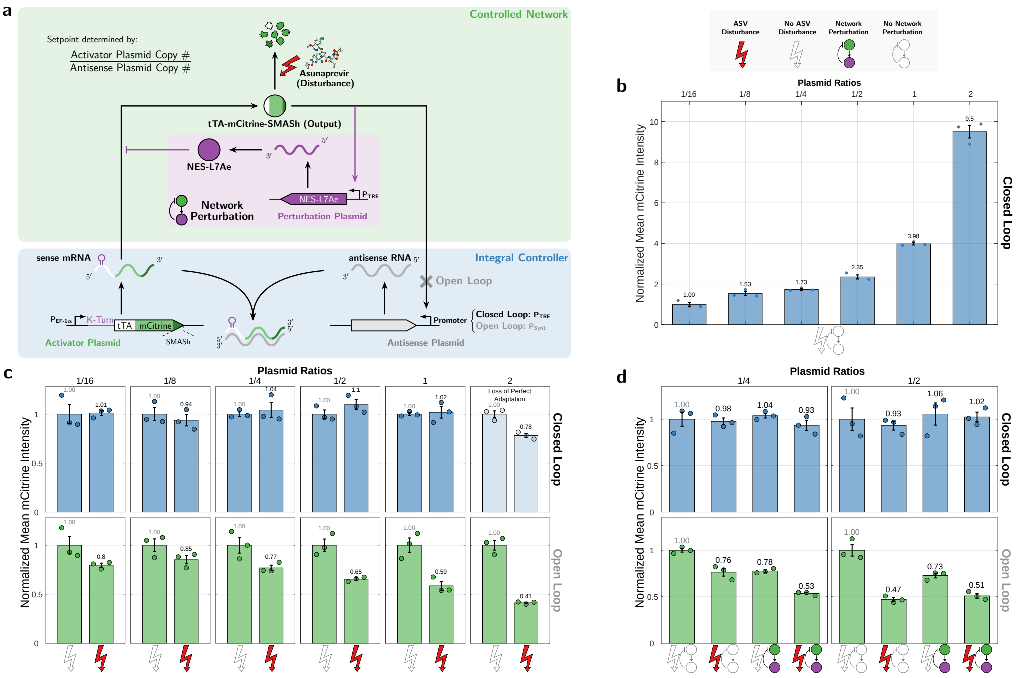


Figure 3.2: See caption on the next page.

variance reduction (Briat et al., 2018; Filo & Khammash, 2021; Chevalier et al., 2019), while maintaining the adaptation property. To implement proportional feedback control that acts faster than the integral feedback, we use a proxy protein, namely the RNA-binding protein L7Ae, which is produced in parallel with mCitrine-tTA from a single mRNA via the use of P2A self-cleavage peptide (Figure 3.3a). Therefore, the expression level of L7Ae is expected to proportionally reflect the level of tTA-mCitrine. The negative feedback is hence realized via the proxy protein that inhibits translation by binding the 5' untranslated region of the sense mRNA. Note that, as opposed to the circuit in Figure 3.2(a), the production of L7Ae in the PI controller is not regulated by the tTA responsive TRE promoter. Instead, it is directly controlled by the sense mRNA. Furthermore, the proportional feedback realized in the PI controller is expected to act faster than the feedback implemented by the tTA-dependent production of L7Ae (Figure 3.2(a)) because it does not require additional transcription and translation steps.

As illustrated in Figure 3.3(b), with a standalone Proportional (P) controller, increasing the proportional feedback strength via introducing additional L7Ae binding hairpins has the effect of reducing the steady-state error induced by the drug disturbance. Nonetheless, despite the error reduction, our criteria of adaptation is not met. On the other hand, with a Proportional Integral (PI) con-

Figure 3.2: Perfect Adaptation of a Synthetic Antithetic Integral Feedback Circuit in Mammalian Cells. (a) Genetic implementation of open- and closed-loop circuits. Both circuits consist of two genes, realized on separate plasmids. The gene in the *activator plasmid* encodes the synthetic transcription factor tTA (tetracycline transactivator) tagged with the fluorescent protein mCitrine and a chemically-inducible degradation tag (SMASh). Its expression is driven by a strong constitutive promoter ($P_{\text{EF-1}\,\alpha}$). The gene in the *antisense plasmid* expresses the antisense RNA under the control of a tTA responsive promoter (P_{TRE}). In the open-loop configuration, the TRE promoter was exchanged for a non-cognate promoter. In this setting the controlled species is the tTA protein, which can be perturbed externally by addition of Asunaprevir (ASV), the chemical inducer of the SMASh degradation tag. Another type of (internal) perturbation is introduced by adding a negative feedback in the controlled network. In particular, a negative feedback loop from tTA-mCitrine to its own production was added by expressing the RNA-binding protein L7Ae under the control of a tTA-responsive TRE promoter. This protein binds to the kink-turn hairpin on the sense mRNA to inhibit the translation of tTA. (b) Steady-state levels of the output (mCitrine) for increasing plasmid ratios. The genetic implementation of the closed-loop circuit as shown in panel (a) was transiently transfected at different molar ratios (setpoint $:=$ *activator* / *antisense*) by varying the concentration of the activator plasmid while keeping the concentration of the antisense plasmid constant. The data is normalized to the lowest setpoint (1/16). This shows that increasing the plasmid ratio increases the steady-state output level. (c) Steady-state response of the open-loop and closed-loop implementations to induced degradation by ASV. The genetic implementation of the open- and closed-loop circuit as shown in panel (a) was transiently transfected at different molar ratios and perturbed with 30 nM of ASV. The data normalized to the unperturbed conditions for each setpoint separately. (d) The closed-loop circuit is not affected by the topology of the regulated network. The closed- and open-loop circuits were perturbed by co-transfecting the *network perturbation* plasmid and by adding 30 nM of ASV. This was done at two setpoints 1/4 and 1/2 (setpoint $:=$ *activator* / *antisense*). The data is normalized to the unperturbed network and no ASV conditionj. For all the data, the HEK293T cells were measured using flow cytometry 48 hours after transfection and the normalized data are shown as mean ± standard error for n = 3 replicates. The unnormalized data is shown in Supplementary Figures B.5 and B.6 and are provided in separate files.

troller, the expression of tTA-mCitrine is ensured to be robust to the induced drug disturbance as depicted in Figure 3.3. This demonstrates that the additional proportional feedback indeed does not break the adaptation property of the antithetic integral controller, as predicted by control theory.

Mathematical modeling

To demonstrate that the circuits in Figures 3.2(a) and 3.3(a) are consistent with our understanding of the regulatory topologies, we first derive detailed mechanistic models of these topologies, starting from basic principles of mass-action kinetics. Next, a model reduction technique is carried out based on a quasi-steady-state approximation that exploits the time-scale separation imposed by the various fast binding/unbinding reactions in the network. The mathematical details can be found in Supplementary Information B.1, B.2 and B.3, where

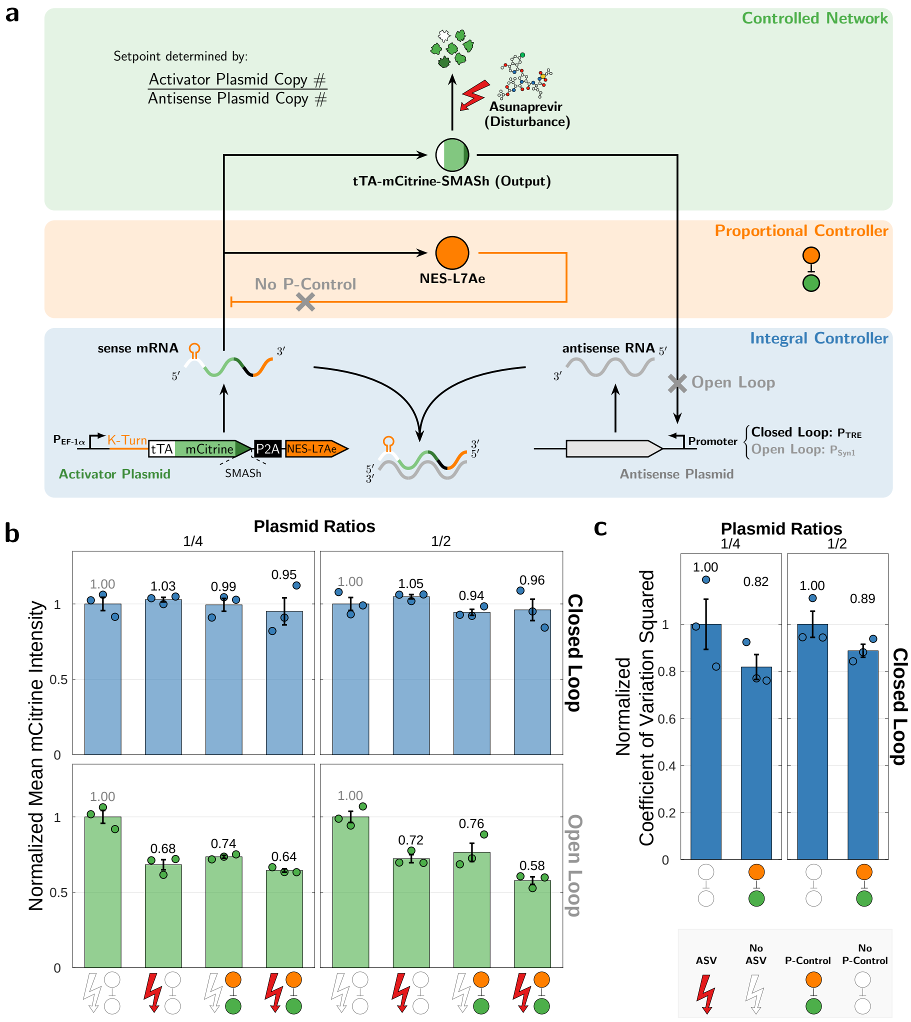


Figure 3.3: See caption on the next page.

each circuit is mathematically treated separately. The resulting reduced models are all compactly presented in a single reaction network depicted in Figure 3.4(a). The overall network can be divided into two biomolecular controller sub-networks – the integral and proportional controllers – that are connected in feedback with another sub-network to be controlled. This is illustrated schematically in Figure 3.4(a) and mathematically as a set of Ordinary Differential Equations (ODEs) in Figures B.1(b), B.3(b) and B.4(b). The reduced models capture the expression dynamics of the three genes, denoted by $\mathbf{G_1}$, $\mathbf{G_2}$ and $\mathbf{G_2'}$, that are encoded in the activator, antisense and network perturbation plasmids, respectively. Gene $\mathbf{G_1}$ is constitutively expressed at a rate $\mu(G_1)$, while the other two genes $\mathbf{G_2}$ and $\mathbf{G_2'}$ are activated by the (dimer) transcription

Figure 3.3: A Proportional-Integral Controller. (a) Genetic implementation of a Proportional-Integral (PI) controller. A negative feedback loop from the RNA-binding protein L7Ae (which is proxy to tTA-mCitrine since it is simultaneously produced from the same mRNA) is added to the antithetic motif. This protein binds in the 5' untranslated region of the sense mRNA species to inhibit the translation of tTA and itself simultaneously. Stronger proportional feedback is realized by adding additional L7Ae binding hairpins. (b) A PI controller does not break the adaptation property. The P and PI circuits were implemented by adding a negative feedback loop from L7Ae to the open- and closed-loop circuits. All circuits were perturbed by adding 30 nM of ASV. The HEK293T cells were measured using flow cytometry 48 hours after transfection and the data is shown as mean per condition normalized to the unpertubed (no ASV) condition ± standard error for n = 3 replicates. (c) Proportional-integral control reduces the steady-state variance. Computing the normalized coefficient of variation squared on the steady-state flow cytometry distributions, reveals a reduction in variation in the presence of proportional feedback. The coefficients of variation squared were normalized to the No P-Control condition for both setpoints and is shown ± standard error for n = 3 replicates. The unnormalized data is shown in Supplementary Figure B.7 and B.8 and is provided in a separate file.

factor **A** at rates $\theta(A; G_2)$ and $\theta_p(A; G_2')$, respectively. The derived mathematical expressions of the functions μ, θ, θ_p and the active degradation propensity, λ, are all given in Figures B.1(b) and B.4(b). Note that the model for the circuit of Figure 3.2(a) without (resp. with) network perturbation can be obtained by setting $G_2' = \tau = 0$ (resp. $\tau = 0$); whereas, the model for the circuit of Figure 3.3(a) can be obtained by setting $G_2' = 0$ and $\tau = 1$.

Next, we calibrate the derived mathematical models to the experimental measurements that were collected at steady state. The measured fluorescence, denoted by M, represents all the molecules involving mCitrine: $\mathbf{X_1}$ $\mathbf{X_2}$, and **A**. It is shown in Supplementary Information B.1 that M can be expressed solely in terms of the concentration of the regulated output **A**, as shown in the bottom of Figure 3.4(a), where c_x is an instrument-related proportionality constant that maps concentrations in nM to fluorescence in a.u., and κ is the dimerization dissociation constant of **A**. Of course, steady-state measurements alone cannot uniquely estimate all parameters in the model. However, by carrying out a steady-state analysis of the underlying differential equations, we can identify a set of parameter groups (or aggregated parameters) that can be uniquely estimated based on the collected data. The detailed mathematical analyses, showing the aggregated parameter groups and their calibrated values are reported for each circuit separately in Supplementary Information B.1, B.2 and B.3.

In the ideal closed-loop scenario where the dilution/degradation rate δ is zero, the steady-state analyses are fairly straight forward and are shown in the bottom of Figures B.1(b), B.3(b) and B.4(b) for each circuit. These analyses show that the steady-state concentration of the regulated output, denoted by $\bar{A}$, is the same for all the circuits and is given by

$$\bar{A} = \kappa_2 \frac{r - k_0/k_2}{1 - r}, \qquad \text{with} \qquad r := \frac{k_1 G_1}{k_2 G_2}. \tag{3.1}$$

Observe that $\bar{A}$ is a monotonically increasing function of the plasmid ratio G_1/G_2, and is independent of the various controlled network parameters, particularly the disturbance D and the plasmid concentration G_2'. As a result, robust perfect adaptation is exactly achieved since the ASV disturbance and the network perturbation have absolutely no effect on the steady-state concentration of the regulated output **A**.

In practice, the dilution/degradation rate δ is never exactly zero, which makes the integrator 'leaky'. In this case, the steady-state analysis becomes more involved, and one cannot obtain an explicit formula for $\bar{A}$ as in the ideal situation. However, implicit (polynomial) formulae can be obtained and are used here to fit the mathematical models to the data. It should be pointed out that when δ is sufficiently small relative to other controller rate parameters (as can be achieved with slowly growing cells and fairly stable sense/antisense RNA) the integrator leakiness will be negligibly small, and perfect adaptation can still be achieved for all practical purposes Aoki et al., 2019; Qian and Del Vecchio, 2018. This is verified experimentally in Figures 3.2(c), (d) and 3.3(b). The model fits for the integral circuit of Figure 3.2(a), shown in Figure 3.4(b), are carried out sequentially for the open-loop circuit first (with and without disturbance), then for the closed-loop circuit (without disturbance). This sequential procedure avoids over-fitting the model to the data. Finally, the closed-loop circuit with disturbance was left for model prediction to assess the calibration accuracy. As shown in the plots of Figure 3.4(b), the model fits the data very well, and is also capable of predicting the experimentally observed disturbance rejection feature of the antithetic integral controller (dashed red curve in the right plot). Similar model calibration procedures were also carried out for the circuits of Figure B.3(a) and B.4(a), and the model fits and predictions are reported in Figures 3.4(c) and B.4(c), respectively. Clearly, the models fit the data quite well, and are also capable of predicting another experimentally observed feature of the antithetic integral controller: robustness to network perturbations. The models also show that appending the proportional controller to the integral controller does not affect the steady state of the measured output, but it is capable of reducing the stationary variance (equivalently the coefficient of variation), as demonstrated experimentally in Figure 3.3(c) (Supplementary Figure B.8(b)) and theoretically through the stochastic simulations depicted in Figure 3.4(d).

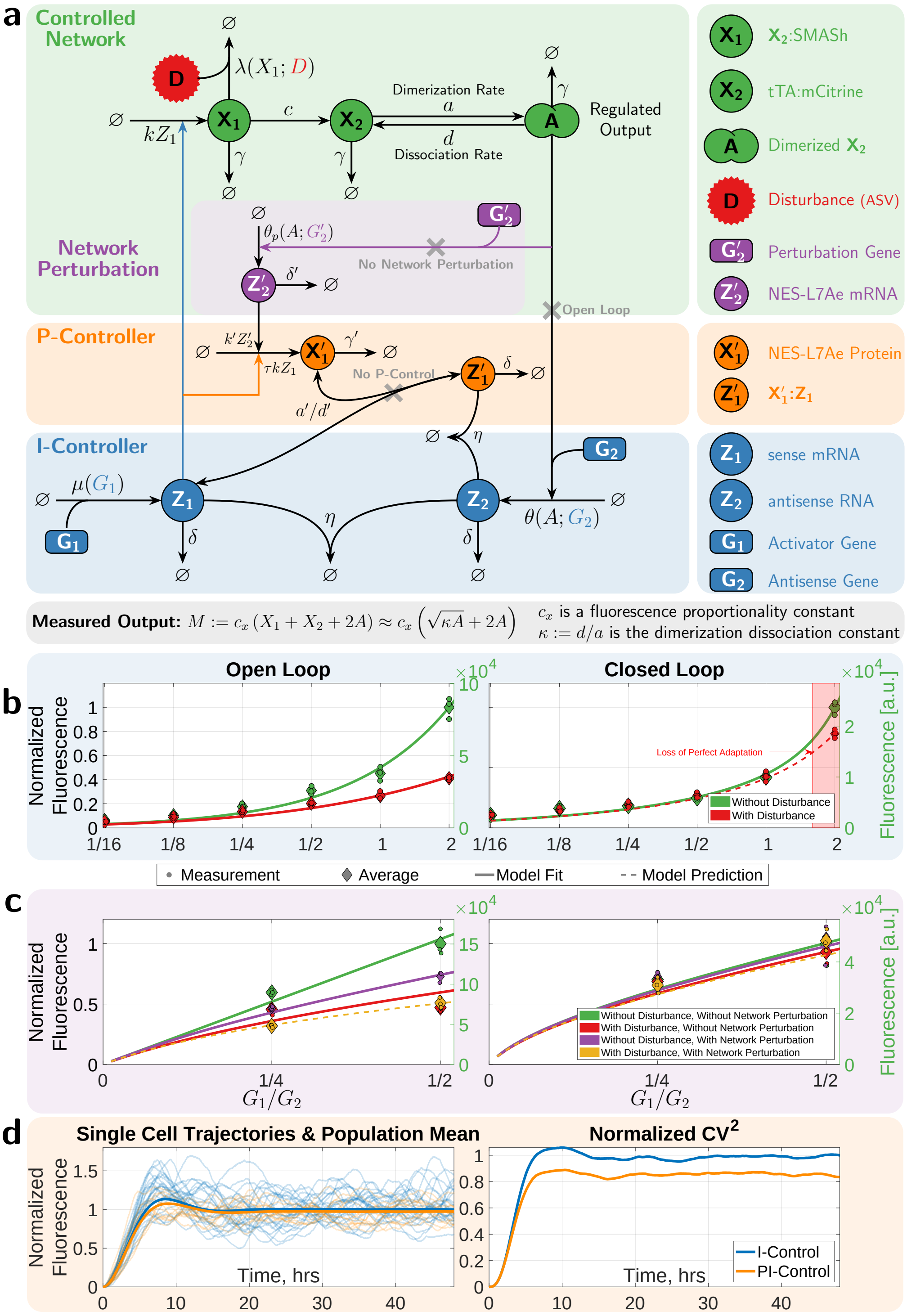


Figure 3.4: See caption on the next page.

Figure 3.4: Mathematical Modeling of the Various Circuits. (a) A Chemical Reaction Network Compactly Modeling the Various Circuits presented in Figures 3.2 and 3.3. The sense mRNA, $\mathbf{Z_1}$, is constitutively produced at a rate $\mu(G_1)$ that depends on the gene (plasmid) concentration, G_1. Then, $\mathbf{Z_1}$ is translated into a fusion of a synthetic transcription factor, fluorescent protein and inducible-degradation tag, referred to as $\mathbf{X_1}$, at a rate k. $\mathbf{X_1}$ is either actively degraded by the ASV disturbance $\mathbf{D}$ at a rate $\lambda(X_1; D)$ or converted to $\mathbf{X_2}$ at a rate c by releasing the SMASh tag. The protein $\mathbf{X_2}$ dimerizes to form $\mathbf{A}$ which activates the transcription of the antisense RNA, $\mathbf{Z_2}$. The transcription rate, denoted by θ, is a function of A and the gene concentration G_2. The antithetic integral control, shown in the blue box, is modeled by the sequestration of $\mathbf{Z_1}$ and $\mathbf{Z_2}$ at a rate η. Note that the open-loop circuit is obtained by removing the feedback from the regulated output $\mathbf{A}$. The proportional controller (orange box) is modeled by producing the protein $\mathbf{X_1}'$, also at a rate k, in parallel with $\mathbf{X_1}$ to serve as its proxy. A negative feedback is then achieved by the (un)binding reaction between the proxy $\mathbf{X_1}'$ and $\mathbf{Z_1}$. Finally, the network perturbation (purple box) is modeled by introducing an additional gene $\mathbf{G_2}'$. This gene is activated by $\mathbf{A}$ to transcribe the mRNA $\mathbf{Z_2}'$ at a rate θ_p which is a function of A and G_2'. $\mathbf{Z_2}'$ is then translated into the protein $\mathbf{X_1}'$ that has, once again, a negative feedback on the production of $\mathbf{X_1}$ by binding to $\mathbf{Z_1}$. See Figures B.1, B.3 and B.4 for a detailed mathematical explanation for each separate circuit. (b) and (c) Model Calibrations to Experimental Data. The left plots show the model fits for the open-loop circuits with/without disturbance in (b) and with/without network perturbation in (c). The right plots similarly show the model fits for the closed-loop circuits. The model fits for proportional control are reported in Figure B.4(c). The solid lines denote model fits, while dashed lines denote model predictions. The model fits and predictions show a very good agreement with the experiments over a wide range of plasmid ratios (setpoints) G_1/G_2 for all scenarios. (d) Stochastic Simulations Demonstrating the Variance Reduction property of the Proportional Controller. The calibrated steady-state parameter groups of the PI closed-loop circuit, given in (B.42), are fixed, while the time-related parameters are set as follows: $\gamma = \gamma', k = c = d = 1\,\text{min}^{-1}$ to demonstrate the variance reduction property that is achieved when a proportional controller is appended to the antithetic integral motif. Note that $G_1 = 0.002\,\text{pmol}$ and $G_2 = 0.004\,\text{pmol}$.

Gene expression burden mitigation

To demonstrate the antithetic integral and proportional-integral controllers in a more practical setting, we apply the circuits introduced in Figure 3.2 and Figure 3.3 to decouple the expression of the transcription factor tTA-mCitrine-SMASh from the expression of other genes when they are competing for finite pools of shared resources. This effect was first described in bacteria (Ceroni et al., 2015) and later also characterized in mammalian cells (Frei et al., 2020; Jones et al., 2020). The effective consequence of this is that changes in the expression of one gene inversely affects the expression of all other genes that share a pool of resources with it. In the context of feedback control, the aforementioned changes in gene expression can be seen as disturbances to the controlled network (Figure 3.5(a)). To experimentally introduce this perturbation, we co-transfected varying amounts of an additional *disturbance plasmid* that constitutively expresses the fluorescent protein miRFP670. Previously, it has been observed that the expression of transiently transfected genes is repressed by the presence of double stranded RNA (dsRNA) (Nejepinska et al., 2014). We similarly observed that the double stranded RNA (dsRNA) formed through the hybridization of sense and antisense mRNA inhibits the expression of the additionally transfected miRFP670 (comparing Closed Loop to Syn1 Open Loop in Supplementary Figure B.9). To make the gene expression burden — reflected by miRFP670 expression levels — comparable between the closed-loop and open-loop conditions, we replaced the inactive Syn1 promoter with a constitutively active EF1α promoter and tuned the plasmid ratio such that the expression of miRFP670 matches the closed-loop expression (Low EF1α Antisense condition in Supplementary Figure B.9). As was already done in Figure 3.3, we now compare the responses of the open-loop (No Control), proportional feedback (P-Control), integral feedback (I-Control) and proportional-integral (PI-Control) variant to this new disturbance. As can be seen in Figure 3.5(b) (Supplementary Figure B.10(a)), a setpoint of $1/2$ is maintained within 10 % up to a disturbance strength of 2.3 for I-Control and for all disturbance strengths for PI-Control (Figure 3.5(b) and Supplementary Figure B.10(a)). This is not the case for the No Control and P-Control configurations, where the steady-state error steadily increases with the increasing strength of the disturbance (Figure 3.5(b) and Supplementary Figure B.10(a)). In all cases, the disturbance is similar in relative extent (Figure 3.5(b) top and Supplementary Figure B.10(a) top). In addition to providing perfect adaptation, PI-Control improves regulation over I-control by further reducing the steady-state cell-to-cell variability (Figure 3.5(c) and Supplementary Figure B.10(b)).

3.5 Discussion

This study presents the first implementation of integral and proportional-integral feedback in mammalian cells. With our proof-of-principle circuit we lay the foundation for robust and predictable control systems engineering in mammalian biology. We believe proportional-integral feedback systems will have a transformative effect on the field of synthetic biology just like they have had on other engineering disciplines.

Based on the antithetic motif (Figure 3.1(a)), we designed and built a proof-of-concept circuit capable of perfect adaptation. This was achieved by exploiting the hybridization of mRNA molecules to complementary antisense RNAs. The resulting inhibition of translation realized the central sequestration mechanism. Specifically, we expressed an antisense RNA through a promoter that was activated by the transcription factor tTA. This antisense RNA was complementary to and bound with the mRNA of tTA to close the negative feedback

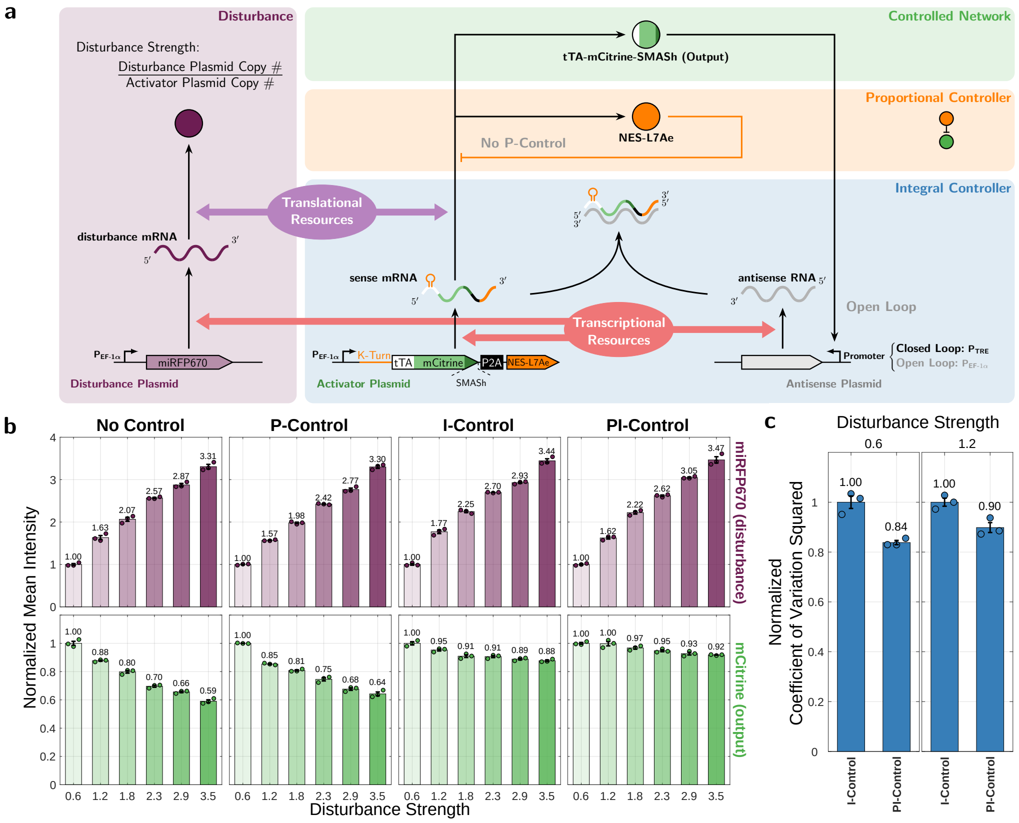


Figure 3.5: Mitigating Competition for Shared Limited Resources with Antithetic Integral and Proportional-Integral Feedback. (a) A genetic implementation of an antithetic integral and proportional-integral feedback circuit for mitigating the effects of limited shared resources. The antithetic integral and proportional-integral feedback circuit characterized in Figure 3.2 and Figure 3.3 are re-purposed to mitigate the coupling of gene expression induced by shared pools of finite resources. Varying the amounts of an additional *disturbance plasmid* that constitutively expresses the fluorescent protein miRFP670 introduces a disturbance to the amount of available resources which indirectly affects the expression levels of tTA-mCitrine-SMASh. (b) Steady-state rejection of disturbances to available limited shared resources. The *activator plasmid* and *antisense plasmid* for all conditions were transiently transfected at a setpoint ratio of 1/2 together with disturbance strengths varying from 0.6 to 3.5. The disturbance strength describes the amount of disturbance plasmid relative to the activator plasmid. (c) Reduction in cell-to-cell variability as a result of proportional-integral feedback control. The coefficient of variation squared was computed for the first two disturbance strengths and normalized two the I-Control condition. The data is shown as the mean ± standard error for $n = 3$ replicates per condition. The unnormalized data is shown in Supplementary Figure B.10 and is provided in a separate file.

loop (Figure 3.2(a)). We further highlighted the properties of integral feedback control by showing that our circuit permits different setpoints. By applying a disturbance to the regulated species we showed that the closed-loop circuit achieved adaptation and provided superior robustness compared to an analogous open-loop circuit (3.2(c)). Further, we showed that adaptation was also achieved when the setpoint of the circuit was changed. An earlier implementation of the antithetic integral feedback motif in bacteria (Aoki et al., 2019) used a σ and anti-σ factor pair to realize the sequestration reaction. Due to the requirement of factors native to the bacterial cell for σ factors to activate transcription, this approach is not directly applicable to mammalian cells. Conversely, the sense and antisense RNA approach utilized in this study is likely to be more difficult to realize in bacterial cells due to rapid mRNA turnover.

Moreover, we demonstrated that our realization of the antithetic integral feedback motif is agnostic to the network structure of the regulated species. This was achieved by introducing a perturbation to the controlled network itself (Figure 3.2(d) and Supplementary Figure B.6). Furthermore, we also demonstrated that the closed-loop circuit still rejected disturbances even in the presence of this extra perturbation to the network. In the open-loop circuit, the disturbance, perturbation, and perturbation with disturbance all led to a strong decrease in tTA-mCitrine expression.

Next, we used the perturbation to the controlled network to incorporate proportional feedback into our integral control circuit directly. We then showed that this proportional-integral feedback controller maintained the same setpoint as the integral controller, even when challenged with induced degradation of the controlled species. To demonstrate that this new controller did utilize proportional feedback, we showed a reduction in the cell-to-cell variability by computing the coefficient of variation squared on the measured fluorescence distributions.

To test our understanding of the mechanistic interactions within our circuits, we derived mechanistic mathematical models for the circuits, starting from basic mass-action kinetics, and showed that the obtained models were capable of fitting the experimental measurements. We also showed that the models were capable of predicting key features of our implementation of the antithetic proportional-integral controller: disturbance rejection and robustness to network perturbations.

Finally, we employed our integral and antithetic-integral feedback circuits to perfectly mitigate gene expression burden on the controlled species caused by introducing an additional, constructively expressed fluorescent protein at varying levels. In light of recent studies on the effects of shared cellular resources in mammalian cells (Frei et al., 2020; Jones et al., 2020), it is important to point out that the dependence of the production of the two controller species on the same resource pool (e.g. transcriptional resources for sense/antisense RNAs) was crucial for maintaining the setpoint despite variations in resource availability. This derives from the fact that the setpoint is a function of the ratio of the production rates of the two controller species (ratio r in 3.1). Whenever both rates depend similarly on the same resource pool, the effect of this dependence cancels out. When the production rates depend on different resource pools, they do not cancel out and the setpoint becomes sensitive to resource allocation.

Aside from realizing integral feedback control, the sense and antisense RNA implementation is very simple to adapt and is versatile. Indeed both sense and antisense are fully programmable, with the only requirement that they share sufficient sequence homology to hybridize and inhibit translation. Due to this fact, mRNAs of endogenous transcription factors may easily be converted into the antithetic motif simply by expressing their antisense RNA from a promoter activated by the transcription factor. However, one should note that in this case the setpoint to the transcription factor will be lower than without the antisense RNA due to the negative feedback. Furthermore, if the mRNA of the endogenous transcription factor is not very stable, the integrator is expected to not perform perfectly.

Genetically engineered controllers have desirable properties as treatment strategies for homeostasis-related pathologies. Previously, it has been demonstrated that when encapsulated insulin-producing designer cells were implanted in diabetic mice they alleviated the effects of type 1 diabetes mellitus (T1DM) by secreting insulin in response to low blood pH mediated by diabetic ketoacidosis (Ausländer et al., 2014) or, alternatively, in response to sensed glucose (Xie et al., 2016). This pioneering work provided a proof-of-concept for the practical feasibility of this approach. In this previous work, however, the designed feedback controller is similar to standalone proportional controller, and therefore cannot exhibit the property of robust perfect adaptation that is characteristic of integral feedback. We next exploit our antithetic proportional-integral controller implementation to carry out a simulation study that demonstrates the achievable robust precision and accuracy of the glucose response in modeled diabetic patients. To illustrate the clinical translatability of our proposed controller topologies, we employed disease models for diabetes mellitus (DM) and interfaced them with the different controller circuits (Figure 3.6). The ability of pancreatic β-cells to synthesize and release insulin determines the classification of DM into two main categories: type 1 DM of autoimmune etiology and type 2 DM (T2DM). As a result, we utilized mathematical models for both T1DM (Man et al., 2014; Dalla Man, Raimondo, et al., 2007) and T2DM (Dalla Man, Rizza, et al., 2007), which originated from the UVA-PADOVA Type 1 Diabetes Simulator (S2008) and its updated versions (Man et al., 2014; Kovatchev et al., 2009).

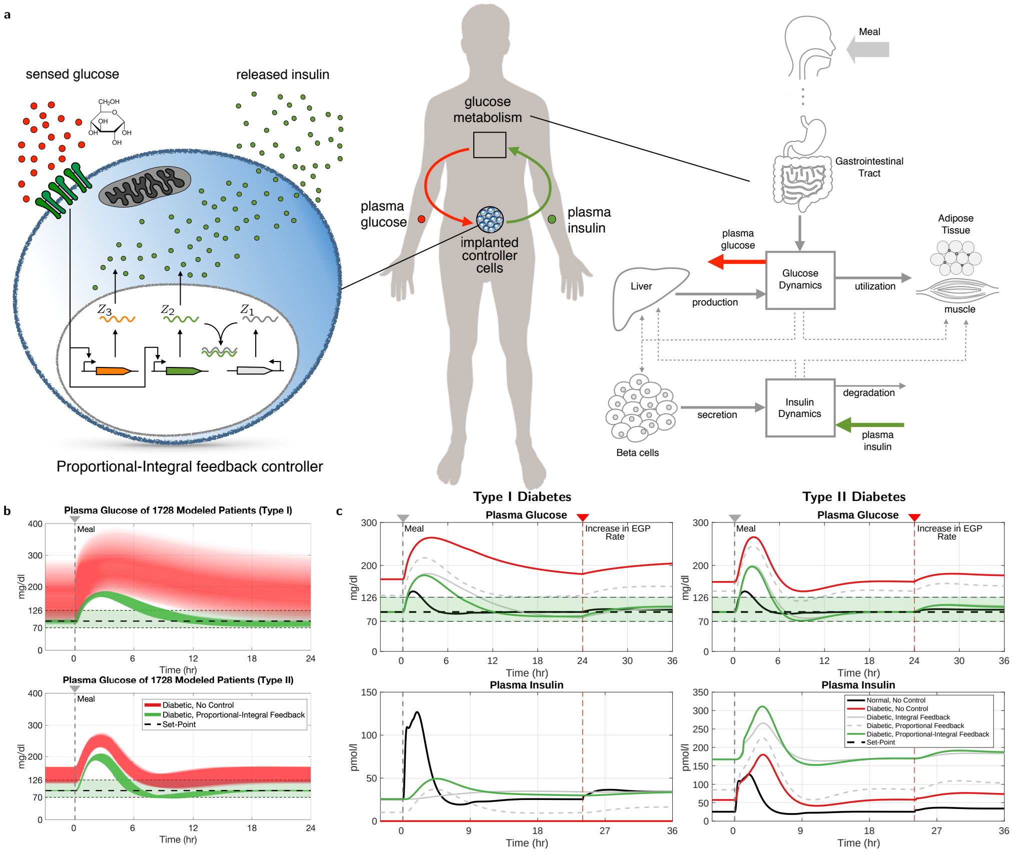


Figure 3.6: See caption on the next page.

This simulator constitutes the first computer model approved by the FDA as an alternative to preclinical trials and animal testing.

Over the past decades the prevalence of DM has increased exponentially, and DM is now considered the most common endocrine disease, affecting approximately 1 in 11 adults globally (Zheng et al., 2018). Our results propose a closed-loop alternative to open-loop replacement therapy with exogenous insulin, which in the case of T1DM is prescribed for life. They also offer a potentially more manageable approach to the combination of lifestyle changes and pharmacological interventions that is recommended for addressing T2DM management (Stumvoll et al., 2005). Moreover, we showed that the simulated glucose control is robust to inter-patient variability (see Figure 3.6(b)), for example due to differences in endogenous glucose production by the liver (clinically found under stress conditions or in critically ill patients (McCowen et al., 2001)), or to changes in renal function, such as physiological or pathological (e.g. diuretic administration, chronic kidney disease) variations in glomerular filtration rate. It was also shown in Figure 3.6(c) that the antithetic integral

Figure 3.6: Simulation of Glucose Regulation in the Blood with Antithetic Proportional-Integral Control. (a) A schematic representation describing the mathematical model of the closed-loop network. The diagram to the right provides a high-level description of the modeled glucose and insulin dynamics based on Dalla Man, Rizza, et al., 2007. This diagram represents the controlled network, where the output of interest (to be controlled) is the glucose concentration ($\mathrm{mg\,dL^{-1}}$) in the plasma; whereas, the input that actuates this network is the insulin concentration ($\mathrm{pmol\,L^{-1}}$) in the plasma. Note that, unlike the controlled network in the previous figures, this network has a negative gain: increasing the input (insulin) decreases the output (glucose). Hence, to ensure an overall negative feedback, a P-Type controller (with positive gain) is adopted here and shown in the schematic to the left which models a genetically embedded antithetic proportional-integral controller. The P-Type property of the integrator is achieved by switching $\mathbf{Z_1}$ with $\mathbf{Z_2}$, that is, the antisense RNA is now constitutively produced while the sense mRNA "senses" the output (glucose) and actuates the input (insulin). The P-Type property of the proportional controller is achieved by using an activation reaction (instead of an inhibition reaction as in Figure 3.3(a)) where glucose activates a gene (in orange) to produce insulin. (b) Robustness to inter-patient variability. To demonstrate the robustness of our proportional-integral controllers, three parameters $k_{p_1} \in [2.4, 3], V_{mx} \in [0.024, 0.071]$ and $k_{e_1} \in [0.0003, 0.0008]$ (see Dalla Man, Rizza, et al., 2007) in the controlled network are varied, while the controller parameters are fixed. Changes of k_{p_1} depict alterations in endogenous glucose production (e.g. in various catabolic or stress states McCowen et al., 2001), V_{mx} is used to simulate variations in the insulin-dependent glucose utilization (U_{id} in Dalla Man, Rizza, et al., 2007) in the peripheral tissues (e.g. by physiological or pathological changes in GLUT4 translocation), while k_{e_1} is the glomerular filtration rate. The responses are shown for a meal of 40 g of glucose at $t = 0$. Adaptation is achieved for all these parameters and for both Type I and II diabetic subjects. (c) Response to 40 g of glucose at time $t = 0$ and a disturbance in endogenous glucose production (EGP) rate at $t = 24\,\mathrm{h}$. A single meal comprised of 40 g of glucose and an increase of endogenous glucose production rate from $k_{p_1} = 2.7\,\mathrm{mg\,min^{-1}} \rightarrow 3\,\mathrm{mg\,min^{-1}}$ (see Dalla Man, Rizza, et al., 2007) are applied to the models of healthy and diabetic subjects at $t = 0$ and $t = 24\,\mathrm{h}$, respectively. The top (resp. bottom) plots depicts the response of glucose (resp. insulin) concentration; whereas the left (resp. right) plots correspond to a Type I (resp. II) diabetic subject. The black curves correspond to a healthy subject whose glucose levels quickly returns back to the glycemic target range (for adults with diabetes) $[80, 130]\,\mathrm{mg\,dL^{-1}}$ Assessment, 2021 after the meal due the naturally secreted insulin. In contrast, the red curves correspond to uncontrolled diabetic patients whose glucose levels are incapable of returning back to the healthy range due to lack of insulin (Type I) or low insulin sensitivity (Type II). Finally, the solid gray, dashed gray and green curves correspond to diabetic patients whose glucose levels are controlled by our integral, proportional and proportional-integral controllers, respectively. Both integral and proportional-integral controllers are capable of restoring a healthy level of glucose concentration by tuning the set-point to a desired value ($100\,\mathrm{mg\,dL^{-1}}$); whereas, the proportional controller alone is neither capable of returning to the desired set-point nor rejecting the disturbance. Furthermore, the proportional-integral controller outperforms the standalone integral controller by speeding up the convergence to the set-point, especially for type I diabetes.

and proportional-integral controllers were capable of achieving robust adaptation. In contrast, a standalone proportional controller did not meet the desired set-point, nor could it reject disturbances such as an increase in endogenous glucose production rate (k_{p_1} in Dalla Man, Rizza, et al., 2007). Note that dissimilarities in the response of the healthy patient and that of the PI-controller-treated patient are, for the most part, not due to any differences between the two regulation strategies (natural vs synthetic). Rather they are mostly attributed to the fact that for the treated patient the insulin was modeled to be synthesized *de novo* from a genetically engineered synthetic insulin gene, leading to inevitable gene expression delay. In comparison, for healthy patients insulin is stored in vesicles for quick release, which ensures a more rapid response — a fact that was also accounted for in the model of the healthy patient. Nevertheless, the response of the PI-controller-treated patient in Figure 3.6(c) meets all the preprandial and peak postprandial plasma glucose guidelines of the American Diabetes Association (Assessment, 2021), and hence offers a potentially effective treatment strategy. Interestingly, the same controller for the single T1DM patient of Figure 3.6(c) (left) was capable of meeting the guidelines for *all* 1728 patients in Figure 3.6(b) (left) without requiring re-tuning for different patients — a clear demonstration of robust adaptation. A similar robust adaptation was seen in T2DM, where a single controller met the guidelines for the majority of patients. For those patients for whom the guidelines were not met, the violation was slight (glucose levels exceeded $180\,\mathrm{mg\,dL^{-1}}$ only briefly beyond the maximum of two hours Figure 3.6(b) (right)). This, however, can be remedied by slightly re-tuning the controller for these patients if necessary. The details of the mathematical modeling can be found in Supplementary Information B.4.

We believe that the ability to precisely and robustly regulate gene expression in mammalian cells will find many applications in industrial biotechnology and biomedicine. In the area of biomedicine, these robust perfectly adapting controllers can be used to restore homeostasis in the treatment of metabolic diseases, as well as for applications in immunotherapy and precise drug delivery.

3.6 Bibliography

Frei, T., Chang, C.-H., Filo, M., Arampatzis, A., & Khammash, M. (2021). Genetically engineered proportional-integral feedback controllers for robust perfect adaptation in mammalian cells. *bioRxiv*. https://doi.org/10.1101/2020.12.06.412304

Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Reece, J. B. (2017). *Campbell biology*. Pearson Education, Incorporated.

Kotas, M. E., & Medzhitov, R. (2015). Homeostasis, inflammation, and disease susceptibility. *Cell*, *160*(5), 816–827. https://doi.org/10.1016/j.cell.2015.02.010

Mullur, R., Liu, Y.-Y., & Brent, G. A. (2014). Thyroid hormone regulation of metabolism. *Physiological reviews*, *94*(2), 355–382.

Yu, F.-X., Zhao, B., & Guan, K.-L. (2015). Hippo pathway in organ size control, tissue homeostasis, and cancer. *Cell*, *163*(4), 811–828.

Åström, K. J., & Murray, R. M. (2010). *Feedback systems: An introduction for scientists and engineers*. Princeton university press.

Yi, T.-M., Huang, Y., Simon, M. I., & Doyle, J. (2000). Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences*, *97*(9), 4649–4653. https://doi.org/10.1073/pnas.97.9.4649

El-Samad, H., Goff, J., & Khammash, M. (2002). Calcium homeostasis and parturient hypocalcemia: An integral feedback perspective. *Journal of Theoretical Biology*, *214*(1), 17–29. https://doi.org/10.1006/jtbi.2001.2422

Miller, P., & Wang, X.-J. (2006). Inhibitory control by an integral feedback signal in prefrontal cortex: A model of discrimination between sequential stimuli. *Proceedings of the National Academy of Sciences*, *103*(1), 201–206.

Muzzey, D., Gómez-Uribe, C. A., Mettetal, J. T., & van Oudenaarden, A. (2009). A systems-level analysis of perfect adaptation in yeast osmoregulation. *Cell*, *138*(1), 160–171.

Ben-Zvi, D., & Barkai, N. (2010). Scaling of morphogen gradients by an expansion-repression integral feedback control. *Proceedings of the National Academy of Sciences*, *107*(15), 6924–6929.

Briat, C., Gupta, A., & Khammash, M. (2018). Antithetic proportional-integral feedback for reduced variance and improved control performance of stochastic reaction networks. *Journal of The Royal Society Interface*, *15*(143), 20180079.

Filo, M., & Khammash, M. (2021). A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance. *bioRxiv*. https://doi.org/10.1101/2021.03.21.436342

Kemmer, C., Gitzinger, M., Daoud-El Baba, M., Djonov, V., Stelling, J., & Fussenegger, M. (2010). Self-sufficient control of urate homeostasis in mice by a synthetic circuit. *Nature biotechnology*, *28*(4), 355.

Wei, P., Wong, W. W., Park, J. S., Corcoran, E. E., Peisajovich, S. G., Onuffer, J. J., Weiss, A., & Lim, W. A. (2012). Bacterial virulence proteins as tools to rewire kinase pathways in yeast and immune cells. *Nature*, *488*(7411), 384–388.

Rössger, K., Charpin-El-Hamri, G., & Fussenegger, M. (2013). A closed-loop synthetic gene circuit for the treatment of diet-induced obesity in mice. *Nature communications*, *4*(1), 1–9.

Xie, M., Ye, H., Wang, H., Charpin-El Hamri, G., Lormeau, C., Saxena, P., Stelling, J., & Fussenegger, M. (2016). *β*-cell–mimetic designer cells provide closed-loop glycemic control. *Science*, *354*(6317), 1296–1301.

Briat, C., Gupta, A., & Khammash, M. (2016). Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Systems*, *2*(1), 15–26. https://doi.org/10.1016/j.cels.2016.01.004

Lillacci, G., Aoki, S., Schweingruber, D., & Khammash, M. (2017). A synthetic integral feedback controller for robust tunable regulation in bacteria. *Biorxiv*, 170951.

Aoki, S. K., Lillacci, G., Gupta, A., Baumschlager, A., Schweingruber, D., & Khammash, M. (2019). A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature*, *570*(7762), 533–537. https://doi.org/10.1038/s41586-019-1321-1

Agrawal, D. K., Marshall, R., Noireaux, V., & Sontag, E. D. (2019). In vitro implementation of robust gene regulation in a synthetic biomolecular integral controller. *Nature communications*, *10*(1), 1–12.

Huang, H.-H., Qian, Y., & Vecchio, D. D. (2018). A quasi-integral controller for adaptation of genetic modules to variable ribosome demand. *Nature Communications*, *9*(1), 5415. https://doi.org/10.1038/s41467-018-07899-z

Milo, R., & Phillips, R. (2015). *Cell biology by the numbers*. Garland Science.

Pérez-Ortín, J. E., Alepuz, P., Chávez, S., & Choder, M. (2013). Eukaryotic mrna decay: Methodologies, pathways, and links to other stages of gene expression. *Journal of molecular biology*, *425*(20), 3750–3775.

Yang, E., van Nimwegen, E., Zavolan, M., Rajewsky, N., Schroeder, M., Magnasco, M., & Darnell, J. E. (2003). Decay rates of human mrnas: Correlation with functional characteristics and sequence attributes. *Genome research*, *13*(8), 1863–1872.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*(7347), 337–342.

Bayer, T. S., & Smolke, C. D. (2005). Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nature biotechnology*, *23*(3), 337–343.

Tigges, M., Marquez-Lago, T. T., Stelling, J., & Fussenegger, M. (2009). A tunable synthetic mammalian oscillator. *Nature*, *457*(7227), 309–312. https://doi.org/10.1038/nature07616

Haberman, R. P., Criswell, H. E., Snowdy, S., Ming, Z., Breese, G. R., Samulski, R. J., & McCown, T. J. (2002). Therapeutic liabilities of in vivo viral vector tropism: Adeno-associated virus vectors, nmdar1 antisense, and focal seizure sensitivity. *Molecular Therapy*, *6*(4), 495–500.

Fuchs, B. C., Perez, J. C., Suetterlin, J. E., Chaudhry, S. B., & Bode, B. P. (2004). Inducible antisense rna targeting amino acid transporter atb0/asct2 elicits apoptosis in human hepatoma cells. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *286*(3), G467–G478.

Tebas, P., Stein, D., Binder-Scholl, G., Mukherjee, R., Brady, T., Rebello, T., Humeau, L., Kalos, M., Papasavvas, E., Montaner, L. J., et al. (2013). Antiviral effects of autologous cd4 t cells genetically modified with a conditionally replicating lentiviral vector expressing long antisense to hiv. *Blood, The Journal of the American Society of Hematology*, *121*(9), 1524–1533.

Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, *89*(12), 5547–5551.

Chung, H. K., Jacobs, C. L., Huo, Y., Yang, J., Krumm, S. A., Plemper, R. K., Tsien, R. Y., & Lin, M. Z. (2015). Tunable and reversible drug control of protein production via a self-excising degron. *Nature chemical biology*, *11*(9), 713–720.

Saito, H., Kobayashi, T., Hara, T., Fujita, Y., Hayashi, K., Furushima, R., & Inoue, T. (2010). Synthetic translational regulation by an l7ae–kink-turn rnp switch. *Nature chemical biology*, *6*(1), 71–78.

Chevalier, M., Gómez-Schiavon, M., Ng, A. H., & El-Samad, H. (2019). Design and analysis of a proportional-integral-derivative controller with biological molecules. *Cell systems*, *9*(4), 338–353.

Qian, Y., & Del Vecchio, D. (2018). Realizing ‘integral control’in living cells: How to overcome leaky integration due to dilution? *Journal of The Royal Society Interface*, *15*(139), 20170902.

Ceroni, F., Algar, R., Stan, G.-B., & Ellis, T. (2015). Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nature methods*, *12*(5), 415–418.

Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., & Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, *11*(1), 1–14.

Jones, R. D., Qian, Y., Siciliano, V., DiAndreth, B., Huh, J., Weiss, R., & Del Vecchio, D. (2020). An endoribonuclease-based feedforward controller for decoupling resource-limited genetic modules in mammalian cells. *Nature communications*, *11*(1), 1–16.

Nejepinska, J., Malik, R., Wagner, S., & Svoboda, P. (2014). Reporters transiently transfected into mammalian cells are highly sensitive to translational repression induced by dsrna expression. *PloS one*, *9*(1), e87517.

Dalla Man, C., Rizza, R. A., & Cobelli, C. (2007). Meal simulation model of the glucose-insulin system. *IEEE Transactions on biomedical engineering*, *54*(10), 1740–1749.

McCowen, K. C., Malhotra, A., & Bistrian, B. R. (2001). Stress-induced hyperglycemia. *Critical care clinics*, *17*(1), 107–124.

Assessment, G. (2021). 6. glycemic targets: Standards of medical care in diabetesd2021. *Diabetes Care*, *44*, S73.

Ausländer, D., Ausländer, S., Charpin-El Hamri, G., Sedlmayer, F., Müller, M., Frey, O., Hierlemann, A., Stelling, J., & Fussenegger, M. (2014). A synthetic multifunctional mammalian ph sensor and co2 transgene-control device. *Molecular cell*, *55*(3), 397–408.

Man, C. D., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., & Cobelli, C. (2014). The uva/padova type 1 diabetes simulator: New features. *Journal of diabetes science and technology*, *8*(1), 26–34.

Dalla Man, C., Raimondo, D. M., Rizza, R. A., & Cobelli, C. (2007). Gim, simulation software of meal glucose—insulin model.

Kovatchev, B. P., Breton, M., Dalla Man, C., & Cobelli, C. (2009). In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes.

Zheng, Y., Ley, S. H., & Hu, F. B. (2018). Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nature Reviews Endocrinology*, *14*(2), 88–98.

Stumvoll, M., Goldstein, B. J., & Van Haeften, T. W. (2005). Type 2 diabetes: Principles of pathogenesis and therapy. *The Lancet*, *365*(9467), 1333–1346.

Chapter Four

Conclusion and outlook

Frei, T., & Khammash, M. (2021). Adaptive circuits in synthetic biology.
Current Opinion in Systems Biology, *28*, 100399

In the present thesis we have described how gene expression burden affects transiently transfected genes in mammalian cells and how these effects can be mitigated using microRNA-based incoherent feedforward loops. Additionally, a mammalian implementation of the antithetic integral feedback motif has been shown to produce robust perfect adaptation and was further extended to realize proportional-integral feedback which further also reduces the variability of the controlled species. Furthermore, it was shown that both integral and proportional-integral feedback mitigate the effects of the competition for shared limited resources on the controlled species.

4.1 Open questions regarding gene expression burden in mammalian cells

In Chapter 2 (Frei et al., 2020) the particular case of gene expression burden in transient transfection was investigated. While, transient transfection serves as a rapid prototyping paradigm for genetic circuit engineering, in most cases the desired end product is a genetically engineered cell line that has the genetic circuit stably integrated into the genome. For this reason, a natural continuation of our work would entail the characterization of gene expression burden to genomically integrated genes. To investigate gene expression burden for genomically integrated genes, one can no longer make use of varying the copy number of an X-tra gene as was done by changing the amount of transfected plasmid. However, one can employ an inducible gene expression system with externally tunable induction as was shown in Fig. 2.2b and c. Moreover, due to the stable copy number of the genes it would also be feasible to perform a characterization of the gene expression dynamics with regard to effects of resource sharing. For this one might also consider using an light-inducible gene expression system such that precise dynamic control of X-tra expression may be achieved. While the simplest methods for genomic integration such as transposase-mediated integration will likely compare more directly with the observations obatined by transient transfection due to their multi-copy integration nature, it will also be of great interest to investigate if and how gene expression burden may also be observed for site-specific single copy integrations.

Another open question in gene expression burden is the exact nature of the limited resources. So far, as was shown in Section 2.3 of Chapter 2 two pools of limiting resources have been distinguished. However, the two identified pools may be further separable into more specific resource pools. Ideally, the limiting resource species would be identified such that one could attempt to directly regulate their amounts to further mitigate resource constraints. In bacteria, ribosomes have been identified as one of the limiting resources for gene expression and coupling of co-expressed genes has been mitigated by introducing orthogonal ribosome species (Darlington et al., 2018). Realizing similar solutions for mammalian cells remains an open challenge.

What mechanisms are capable of maintaining conserved resource levels in a cell remains to be elucidated as well. Given that all biological systems constantly consume and dissipate energy, the conservation might not follow from fundamental conservation laws known in physics but may be conserved through emergent conservation laws (Baez et al., 2018).

4.2 Extensions to more sophisticated feedback controller architectures

While integral feedback can guarantee robust perfect adaptation, it does not ensure safe transient responses. More so, it is known that integral feedback has a tendency to destabilize a system. To additionally ensure safer transient dynamics, the integral feedback may be augmented with additional feedback paradigms such as proportional and derivative feedback. The proportional-integral feedback circuit introduced in Chapter 3 (Frei et al., 2021) represent a first step towards this goal. However, an assessment of the transient response of the circuit was unfortunately outside of the scope of this thesis.

In particular, to reliably record transient responses we require that the circuits we are interested in have been stably integrated into the genome of the host cell line and a monoclonal population was expanded from a single cell. Additionally, the reporters in the circuit which are used as readouts of the behavior of the circuit should be sufficiently destabilized so that they faithfully follow the dynamics they should be reporting on.

Once transient responses can be recorded reliably, one can explore more advanced feedback topologies such as proportional-integral-derivative feedback (Chevalier et al., 2019; Filo et al., 2021). Furthermore, it will also become possible to directly observe the effects of feedback gain tuning on the transient response. This will not only allow the controller architectures to adapt perfectly but also do so quickly and safely.

4.3 Mutual removal as a generalization of the antithetic motif

If one takes the antithetic integral feedback motif and relaxes the assumption that the species Z_1 and Z_2 annihilate each other at the same rate, one arrives at a slight modification of the motif where the interaction of the two controller species can be interpreted as mutually degrading or removing each other. Mathematically, this corresponds to substituting the rate η with two distinct rates η_1 and η_2 in the rate equations for the species Z_1 and Z_2, as shown in Eq. (4.1). In this regard, the canonical antithetic integral feedback motif is a special case of the mutual degradation interpretation, where $\eta_1 = \eta_2 = \eta$.

$$\begin{aligned}\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu - \eta_1 Z_1(t) Z_2(t) \\ \frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \theta f(X_n(t)) - \eta_2 Z_1(t) Z_2(t)\end{aligned} \tag{4.1}$$

By setting the derivatives of $Z_1(t)$ and $Z_2(t)$ to zero and solving for the product $\overline{Z}_1\overline{Z}_2$ in one of the two equations, substituting this expression into the other and solving for $\overline{X}_n$ we find that

$$\overline{X}_n = f^{-1}\left(\frac{\eta_1}{\eta_2}\frac{\mu}{\theta}\right). \tag{4.2}$$

Compared to the steady state of the minimal antithetic motif $\overline{X}_n = f^{-1}\left(\frac{\mu}{\theta}\right)$, the steady state of the mutual removal model is additionally scaled by $\frac{\eta_1}{\eta_2}$. This

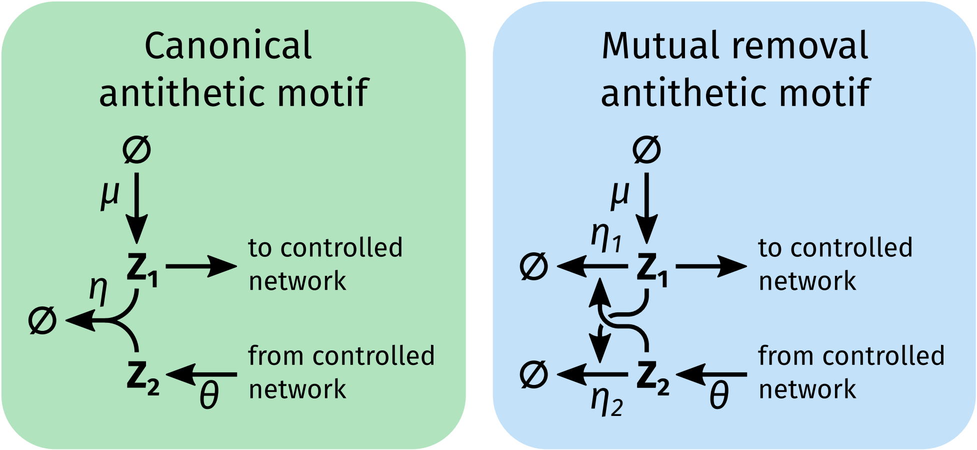


Figure 4.1: Mutual removal interpretation of the antithetic integral feedback motif.

provides additional parameters through which the set point value may be adjusted. However, one may also view this as an increase in the fragility of the set point, as it will now be sensitive to changes in two additional parameters.

The interpretation of the antithetic motif as two separate removal reactions extends the space of possible biomolecular realizations by relaxing the requirement of the strict one-to-one stoichiometry of the sequestration reaction. Furthermore, it also provides a framework in which imperfect sequestration can also to be shown to produce perfect adaptation. For example, even when most interactions between the controller species Z_1 and Z_2 lead to catalytic degradation of either one of them and only a few of these interactions yield proper sequestration perfect adaptation may be achieved. To see how the mutual removal interpretation encompasses this case, consider the case where Z_1 is also catalytically degraded by Z_2.

$$\begin{aligned}\frac{\mathrm{d}Z_1(t)}{\mathrm{d}t} &= \mu - (\eta + \eta^{cat})Z1(t)Z2(t)\\ \frac{\mathrm{d}Z_2(t)}{\mathrm{d}t} &= \theta f(X_n(t)) - \eta Z1(t)Z2(t)\end{aligned} \tag{4.3}$$

By setting $\eta_1 := \eta + \eta^{cat}$ and $\eta_2 := \eta$ in Eq. (4.3) we show the connection to the mutual removal model shown in Eq. (4.1). Within this perspective, a much larger class of biomolecular realizations may be capable of implementing integral feedback controllers. For example, the interaction between microRNAs and their target mRNAs haven been described to be both catalytic through either cleaving or deadenylation of the mRNA, or sequestration trough binding of the microRNA complex to the mRNA and blocking translation (Jonas & Izaurralde, 2015).

4.4 Adaptation beyond simple controllers

Phenomena in biology (e.g. homeostasis) have inspired the application of integral feedback to gene regulation to recreate perfect adaptation. However, biological plasticity may provide more sophisticated forms of adaptation that go beyond fixed network topologies and may be better understood as forms of learning. For example, the *good regulator* theorem from cybernetics (Conant & Ross Ashby, 1970) states that *every good regulator of a system must be a model of that system.* In control theory, a more specific result is known as the *internal model principle* (Francis & Wonham, 1976). It states that integral control provides an internal model for constant disturbances that may also be extended to ramped and parabolic disturbances by sequentially composing integrators. By learning more general internal models of the external world, more sophisticated forms of adaptation may be possible. Specifically, one may envision forms of adaptation that also adaptively change the reference levels such that certain optimality conditions are met. Moreover, learned internal models may hold the potential to generalize beyond the initially considered disturbances. Therefore, understanding the principles that govern learning from a theoretical perspective, such as endeavors in machine learning and artificial intelligence, may lead to both a better understanding of biology as well as enable the discovery of even more sophisticated mechanisms of adaptation.

Some theoretical studies have started considering the realization of artificial neural networks with chemical reactions (Genot et al., 2012; Poole et al., 2017; Moorman et al., 2019; Samaniego et al., 2020; Anderson et al., 2020), a feat that was realized already in DNA computing (Qian et al., 2011; Cherry & Qian, 2018). More generally, more abstract and flexible frameworks capable of capturing aspects of learning might be more fruitful in formalizing learning for reaction networks. For example, theories that originated from evolutionary game theory have been shown to be analogous to bayesian computations (Czégel et al., 2020). More abstractly, a general theory of biological systems as adaptive learners has emerged from theories of the brain. Collectively these theories fall under the umbrella term *free energy principle* and *active inference* (Karl, 2012). The specifics and plausibility of these theories are however still being fiercely debated both from a scientific and philosophical point of view (Bruineberg et al., 2020; Biehl et al., 2021). While these theories are either still very restrictive or not developed for the particular application in chemical reaction networks, they might pose interesting avenues for further exploring adaptive circuits that can learn their internal model.

4.5 Concluding remarks

The thesis presented here demonstrates a step towards the successful integration of control theory concepts into synthetic biology. While there remain many open challenges in this endeavor, we are witnessing and actively contributing to the formation of a subfield in synthetic biology that strives to employ control theoretic principles to engineer reliable and robust biological systems. We firmly believe that further advances in this field will enable the construction of highly sophisticated systems capable of addressing many important challenges.

4.6 Bibliography

Frei, T., & Khammash, M. (2021). Adaptive circuits in synthetic biology. *Current Opinion in Systems Biology*, *28*, 100399.

Frei, T., Cella, F., Tedeschi, F., Gutiérrez, J., Stan, G.-B., Khammash, M., & Siciliano, V. (2020). Characterization and mitigation of gene expression burden in mammalian cells. *Nature communications*, *11*(1), 1–14.

Darlington, A. P., Kim, J., Jiménez, J. I., & Bates, D. G. (2018). Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nature communications*, *9*(1), 1–12.

Baez, J. C., Pollard, B. S., Lorand, J., & Sarazola, M. (2018). Biochemical coupling through emergent conservation laws. *arXiv preprint arXiv:1806.10764*.

Frei, T., Chang, C.-H., Filo, M., Arampatzis, A., & Khammash, M. (2021). Genetically engineered proportional-integral feedback controllers for robust perfect adaptation in mammalian cells. *bioRxiv*. https://doi.org/10.1101/2020.12.06.412304

Chevalier, M., Gómez-Schiavon, M., Ng, A. H., & El-Samad, H. (2019). Design and analysis of a proportional-integral-derivative controller with biological molecules. *Cell systems*, *9*(4), 338–353.

Filo, M., Kumar, S., & Khammash, M. H. (2021). A hierarchy of biomolecular proportional-integral-derivative feedback controllers for robust perfect adaptation and dynamic performance. *bioRxiv*.

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microrna-mediated gene silencing. *Nature reviews genetics*, *16*(7), 421–433.

Conant, R. C., & Ross Ashby, W. (1970). Every good regulator of a system must be a model of that system. *International journal of systems science*, *1*(2), 89–97.

Francis, B. A., & Wonham, W. M. (1976). The internal model principle of control theory. *Automatica*, *12*(5), 457–465.

Genot, A. J., Fujii, T., & Rondelez, Y. (2012). Computing with competition in biochemical networks. *Physical review letters*, *109*(20), 208102.

Poole, W., Ortiz-Munoz, A., Behera, A., Jones, N. S., Ouldridge, T. E., Winfree, E., & Gopalkrishnan, M. (2017). Chemical boltzmann machines. *International Conference on DNA-Based Computers*, 210–231.

Moorman, A., Samaniego, C. C., Maley, C., & Weiss, R. (2019). A dynamical biomolecular neural network. *2019 IEEE 58th Conference on Decision and Control (CDC)*, 1797–1802.

Samaniego, C. C., Moorman, A., Giordano, G., & Franco, E. (2020). Signaling-based neural networks for cellular computation. *bioRxiv*. https://doi.org/10.1101/2020.11.10.377077

Anderson, D. F., Deshpande, A., & Joshi, B. (2020). On reaction network implementations of neural networks. *arXiv preprint arXiv:2010.13290*.

Qian, L., Winfree, E., & Bruck, J. (2011). Neural network computation with dna strand displacement cascades. *Nature*, *475*(7356), 368–372.

Cherry, K. M., & Qian, L. (2018). Scaling up molecular pattern recognition with dna-based winner-take-all neural networks. *Nature*, *559*(7714), 370–376.

Czégel, D., Giaffar, H., Zachar, I., Tenenbaum, J. B., & Szathmáry, E. (2020). Evolutionary implementation of bayesian computations. *BioRxiv*, 685842.

Karl, F. (2012). A free energy principle for biological systems. *Entropy*, *14*(11), 2100–2121.

Bruineberg, J., Dolega, K., Dewhurst, J., & Baltieri, M. (2020). The emperor's new markov blankets. *Behavioral and Brain Sciences*, 1–63.

Biehl, M., Pollock, F. A., & Kanai, R. (2021). A technical critique of some parts of the free energy principle. *Entropy*, *23*(3), 293.

Appendix A

Characterization and mitigation of gene expression burden in mammalian cells

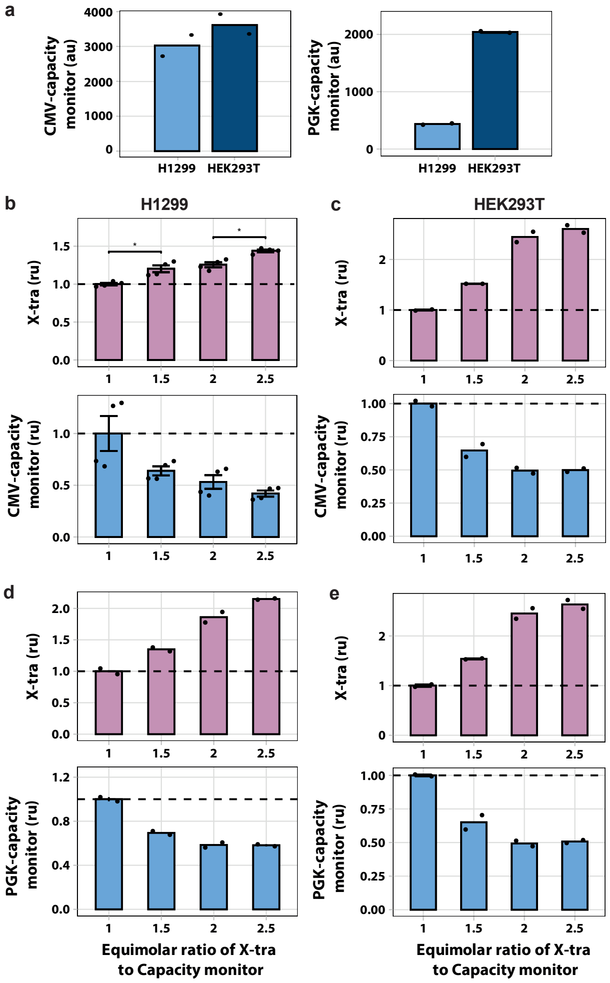


Figure A.1: Relation of X-tra and capacity monitor expression in H1299 and HEK293T cell lines using CMV and PGK promoter. (a) Levels of fluorescence driven by the same promoter (CMV or PGK) differ across cell lines. Data show absolute units of capacity monitor detected by flow cytometry in 1:1 molar ratio transfection. N=2 biological replicates. Source data are provided as a Source Data file. Flow cytometry results of H1299 (N=4 biological replicates). Source data are provided as a Source Data file. (b) and HEK293T (N=2 biological replicates). Source data are provided as a Source Data file. (c) cells co-transfected with fixed amount of CMV-mKate (capacity monitor) PGK-EGFP (X-tra) (molar ratio from 1:1 to 1:2.5). Flow cytometry results of H1299 (N=2 biological replicates) (d) and HEK293T (N=2 biological replicates). Source data are provided as a Source Data file. (e) cells co-transfected with fixed amount of mKate (capacity monitor) under PGK promoter regulation and increasing amount of EGFP (X-tra) under CMV promoter regulation (molar ratio from 1:1 to 1:2.5). N=2 biological replicates. Source data are provided as a Source Data file. Data show the mean fluorescence normalized to its value at a plasmid molar ratio of 1. Error bars represent the standard error, SE. au: arbitrary units. ru: relative units. Unpaired two-sided T-test. p-value: * < 0.05.

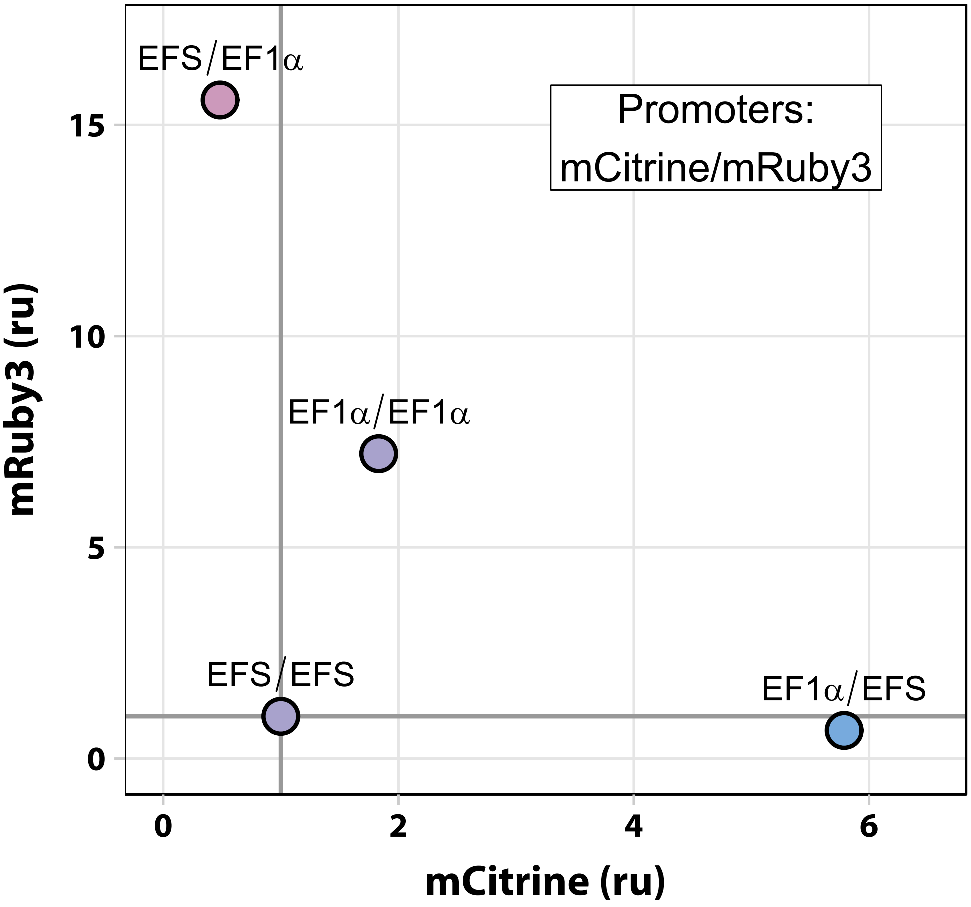


Figure A.2: Promoter expression strength indirectly affects expression of co-transfected genes. Plasmids expressing the fluorescent protein mCitrine and mRuby3 from a strong (EF-1α) or a medium strength EF-1α short (EFS) promoter were co-transfected in several molar ratio combinations. The expression levels for both mCitrine and mRuby3 were normalized by the data obtained from the weakest promoters pair (EFS/EFS). Similar to Fig. 2.2a, there is a negative correlation between the expression strength of one protein and the promoter strength of the other gene. Of note, when strong promoters drive both proteins, the global expression levels drop as already suggested by Fig. 2.2a. Source data are provided as a Source Data file. Data was acquired 48 hours after transfection and is plotted as fluorescence normalized to the EFS/EFS sample. ru: relative units.

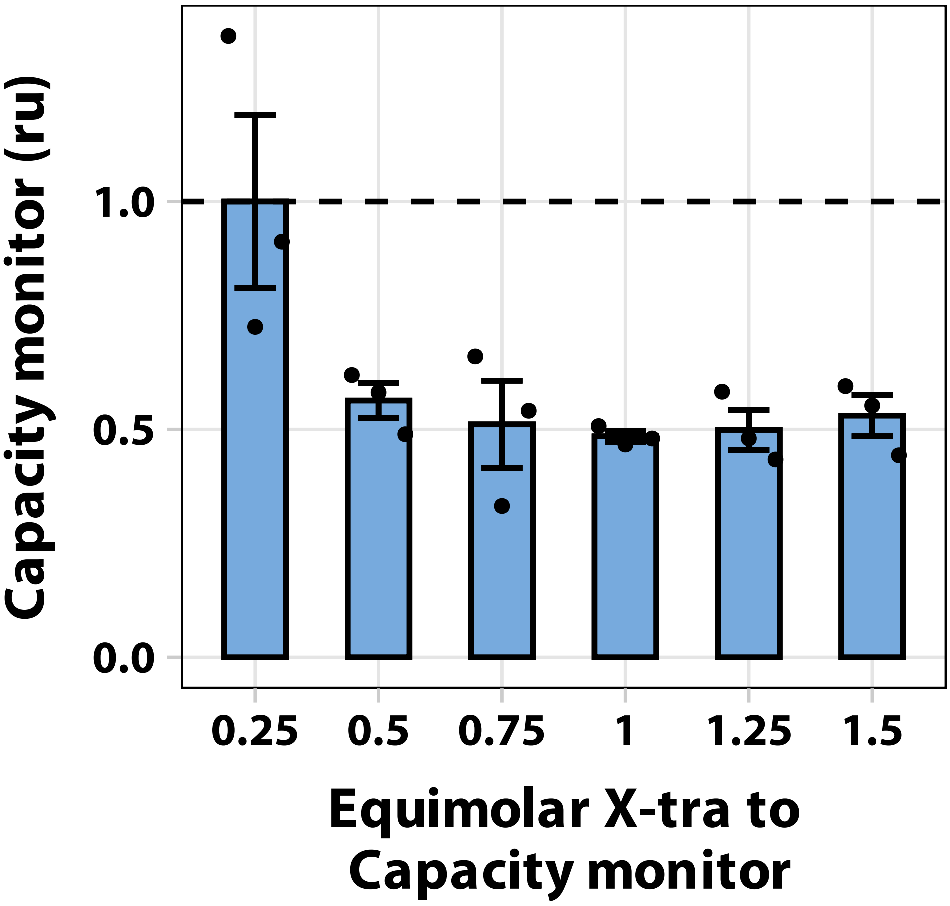


Figure A.3: Non-fluorescent protein genes also compete for cellular resources. A plasmid encoding a human codon optimized variant of the bacterial σ-factor sigW was co-transfected in increasing amounts with a fixed concentration of the mCitrine capacity monitor plasmid. Data were acquired 48 hours after transfection and are plotted as mean fluorescence normalized to the lowest equimolar ratio. Error bars represent the standard error, SE. ru: relative units. N=3 biological replicates. Source data are provided as a Source Data file.

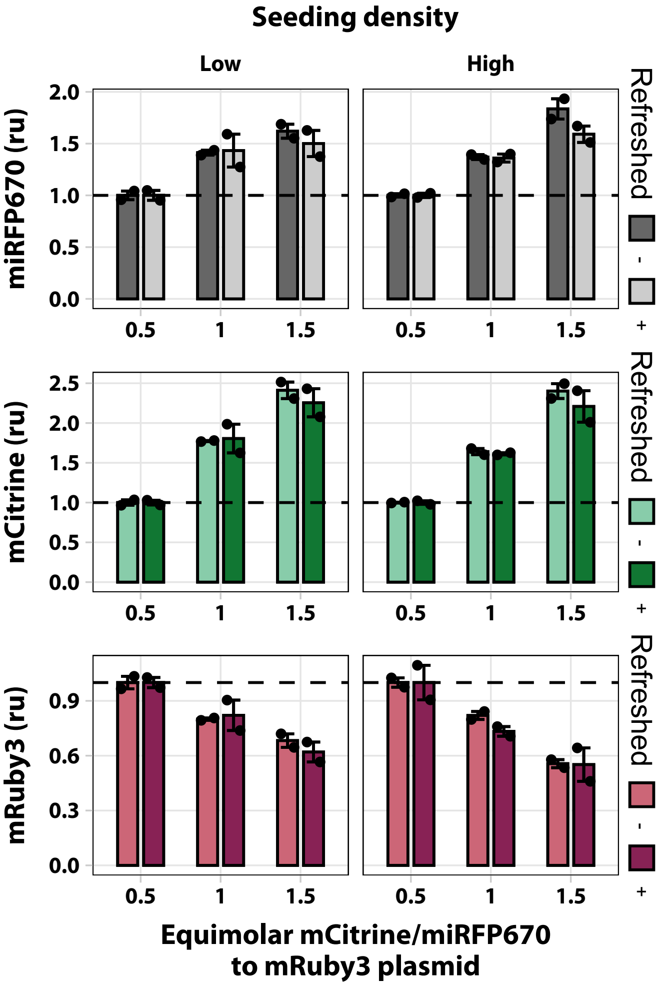


Figure A.4: Assessing nutrient starvation and cell seeding density as potential impacts on limited resources. In this experimental setting, cells were co-transfected with two plasmids. The first plasmid, which was provided at incremental levels, is composed of two transcriptional units (TU), one consisting of a strong promoter driving the expression of mCitrine (hEF1a), the other driving the expression of miRFP670 under a weak promoter (SV40). The second plasmid encodes for mRuby3 under a strong constitutive promoter (hEF1a). HEK293T cells were seeded at 5e^4 (low) and 7.5e^4 (high) cells/well to assess seeding density effects. To investigate the effects of nutrient starvation, we refreshed the medium in two out of four wells per condition. Data were collected 48 hours post transfection and represent the mean fluorescence intensity of the three fluorescent proteins normalized to the 0.5 equimolar ratio condition. Error bars represent the standard error, SE. N=2 biological replicates. Source data are provided as a Source Data file.

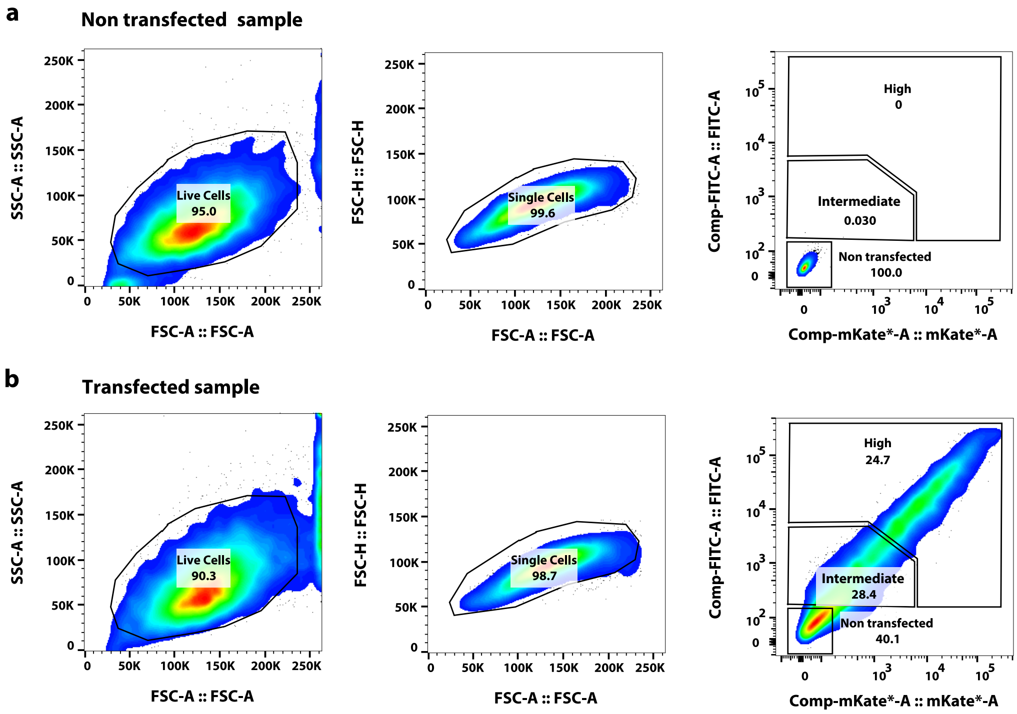


Figure A.5: Sorting strategy. H1299 cells were transfected with a plasmid encoding the fluorescent proteins EGFP and mKate, expressed from a bidirectional promoter. Cells were sorted by fluorescence intensity 48 hours post-transfection to collect non-transfected, intermediate and high transfected cells from the same transfection plate. (a, b) First, gates to select live and single cells were determined (left and middle plots). Then, the threshold for fluorescent intensity was set using a non-transfected sample as reference (a, right). The two additional gates to collect intermediate and high transfected cells were created as shown in the plots (a, b) on the right.

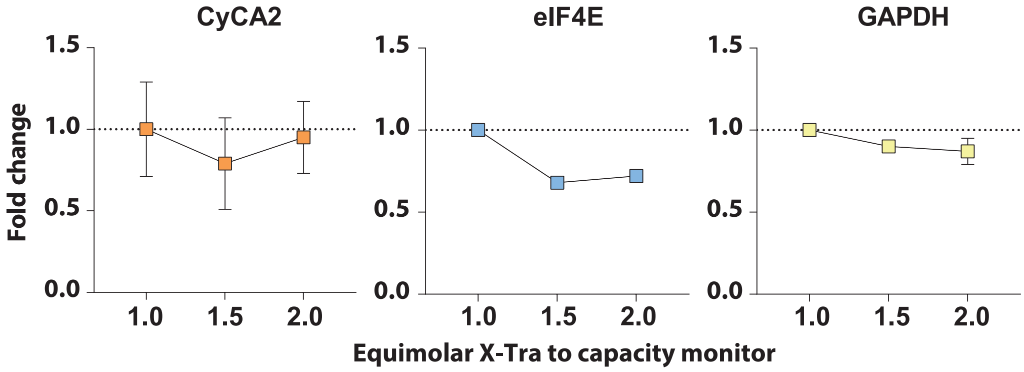


Figure A.6: Effect of X-tra titration on endogenous genes. We measured CyCA2, eIF4E and GAPDH mRNA levels by qPCR in the samples shown in Fig. 2.2c at 1.0, 1.5 and 2.0 molar ratios. Data represent the mean value normalized to the equimolar ratio of 1.0. Error bars represent the standard error, SE. N=4 biological samples for CyCA2 and GAPDH. N=2 biological samples for eIF4E. Source data are provided as a Source Data file.

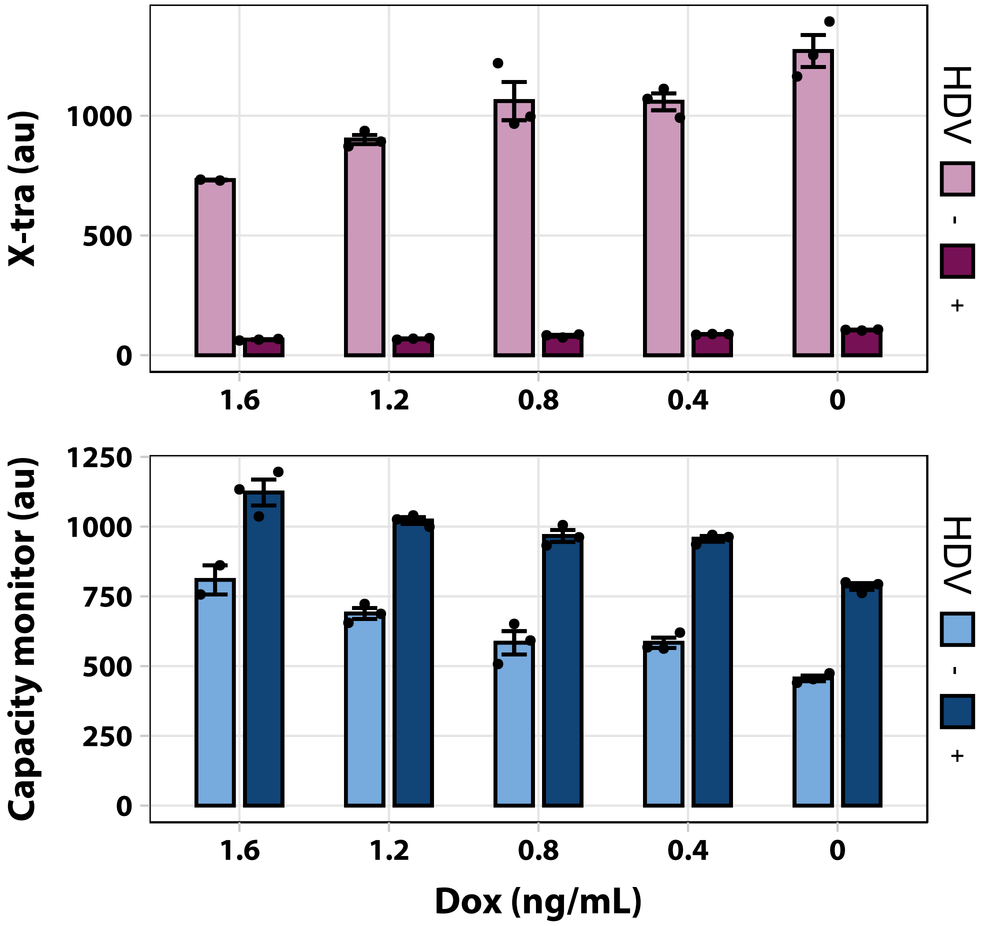


Figure A.7: Fluorescence data of Fig. 2.2e shown in arbitrary units. In this experimental setting Dox represses X-tra transcription. Thus, the lower Dox, the higher the X-tra levels, and as a consequence, the lower the capacity monitor levels. The HDV-dependent mRNA decapping and degradation of X-tra should consume less translational resources, which is consistent with the higher expression of the capacity monitor (dark blue bars) as compared to the inactive mutant (pale blue bars). Data was acquired 48 hours after transfection and is plotted as mean fluorescence intensity +/- SE. SE: standard error. N=3 biological replicates (N = 2 for HDV -, 1.6 ng/µL DOX). Source data for Fig. 2.2e are provided as a Source Data file.

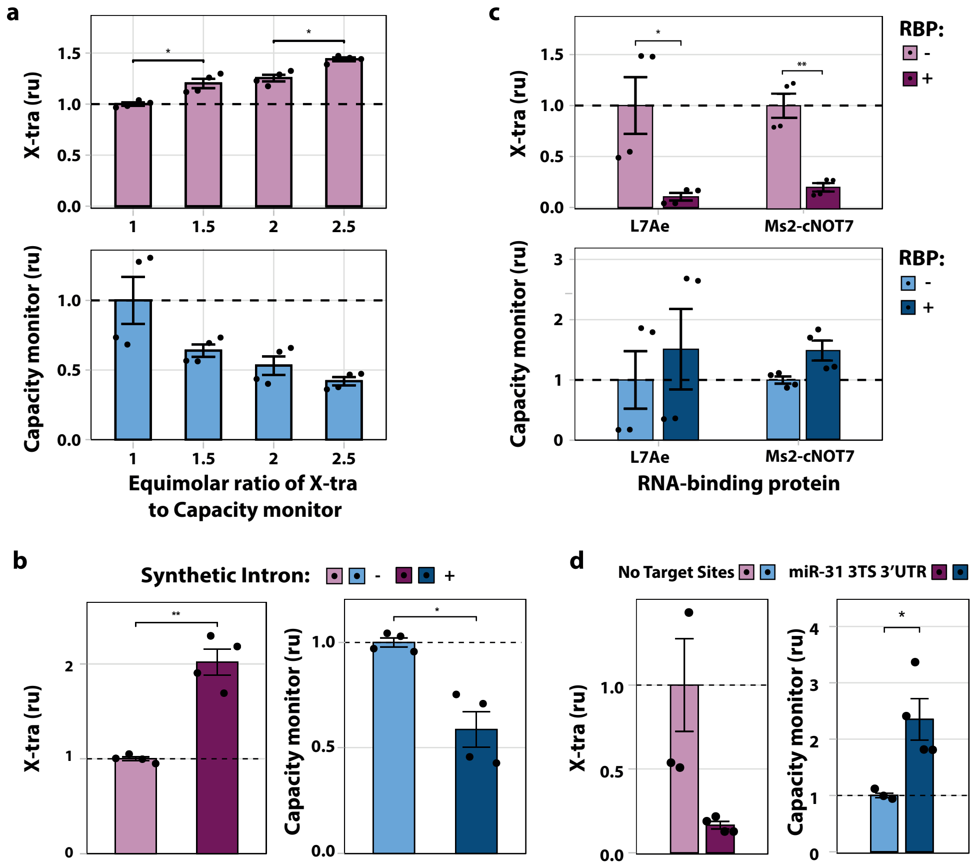


Figure A.8: Gene expression burden in H1299 cells. (a) Flow cytometry results of H1299 cells co-transfected with fixed amount of capacity monitor and increasing amount of X-tra (1:1 to 1:2.5 molar ratio), both under CMV promoter regulation. Data show the mean fluorescence normalized to its value at a plasmid molar ratio of 1. N=4 biological replicates. Source data are provided as a Source Data file. (b) Flow cytometry results of H1299 cells co-transfected with X-tra (mKate) which includes or not a synthetic intron in the 5'UTR, and capacity monitor (EGFP). Data show that when mKate expression is enhanced by the synthetic intron, EGFP levels decrease. Data represent the mean fluorescence normalized to fluorescence values in the absence of the intron. N=4 biological replicates. Source data are provided as a Source Data file. (c) Flow cytometry results of H1299 cells co-transfected with 2kturn-EGFP or EGFP-8xMs2 (X-tra) and mKate (capacity monitor) in presence or absence of L7Ae or Ms2-cNOT7 respectively. Data show that when X-tra is down-regulated, the capacity monitor levels increase. Plot represents mean fluorescence normalization of fluorescence values to the condition without RBP. N=4 biological replicates. Source data are provided as a Source Data file. (d) Flow cytometry results of H1299 cells co-transfected with mKate (X-tra) that includes or not miR-31 target sites in the 5'UTR, and EGFP (capacity monitor). Data show that capacity monitor levels are higher when the X-tra is downregulated by miR-31. N=4 biological replicates (N = 3 for noTS). Source data are provided as a Source Data file. Plot represents normalization of mean fluorescence values to the no target site condition. Data were acquired 48 hours post-transfection. Error bars represent the standard error. ru: relative units. Unpaired two-sided T-test. p-value: **<0.005, *<0.05.

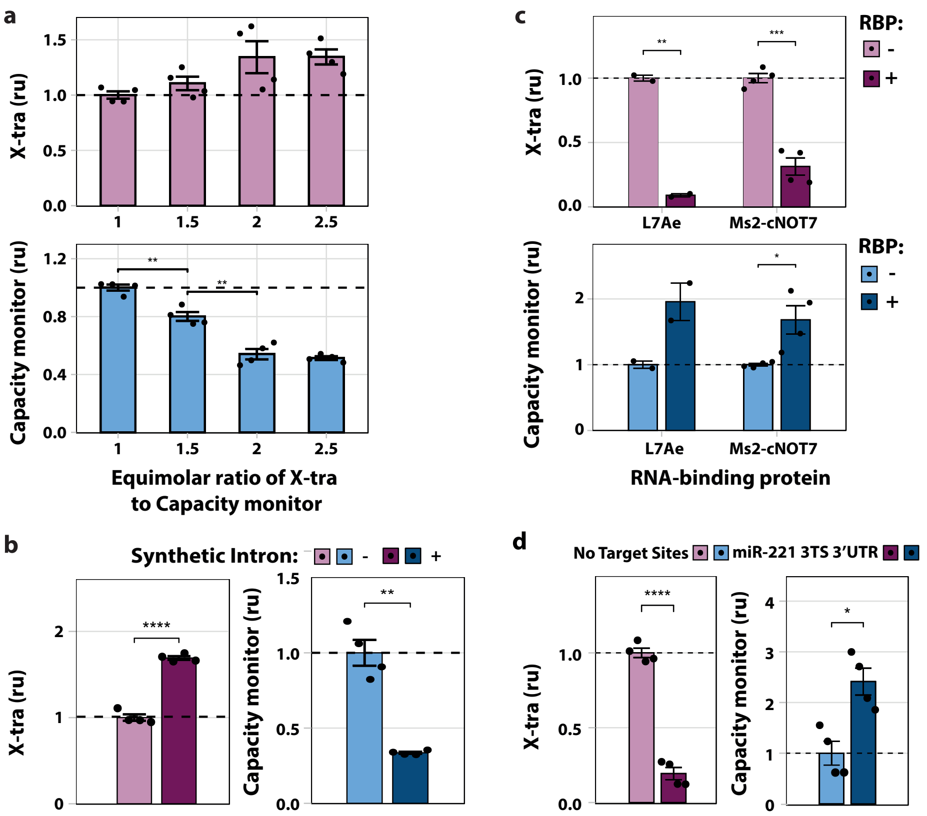


Figure A.9: Gene expression burden in U2OS cells. (a) Flow cytometry results of U2OS cells co-transfected with fixed amount of capacity monitor and increasing amount of X-tra (1:1 to 1:2.5 molar ratio), both under CMV promoter regulation. Data show the mean fluorescence normalized to its value at a plasmid molar ratio of 1. N=4 biological replicates. Source data are provided as a Source Data file. (b) Flow cytometry results of U2OS cells co-transfected with X-tra (mKate) which includes or not a synthetic intron in the 5'UTR, and capacity monitor (EGFP). Data show that when mKate expression is enhanced by the synthetic intron, EGFP levels decrease. Data are the mean fluorescence normalized to fluorescence values in the absence of the intron. N=4 biological replicates. Source data are provided as a Source Data file. (c) Flow cytometry results of U2OS cells co-transfected with 2kturn-EGFP or EGFP-8xMs2 (X-tra) and mKate (capacity monitor) in presence or absence of L7Ae or Ms2-cNOT7 respectively. Data show that when X-tra is down-regulated, the capacity monitor levels increase. Plot represents normalization of mean fluorescence values to the condition without RBP. N=2 biological replicates for L7Ae and N=4 for Ms2-cNOT7. Source data are provided as a Source Data file. (d) Flow cytometry results of U2OS cells co-transfected with mKate (X-tra) that includes or not miR-221 target sites in the 5'UTR, and EGFP (capacity monitor). Data show that capacity monitor levels are higher when the X-tra is downregulated by miR-221. N=4 biological replicates. Source data are provided as a Source Data file. Plot represents normalization of mean fluorescence values to the no target site condition. Data were acquired 48 hours post-transfection. Error bars represent the standard error. ru: relative units. Unpaired two-sided T-test. p-value: **<0.005, *<0.05.

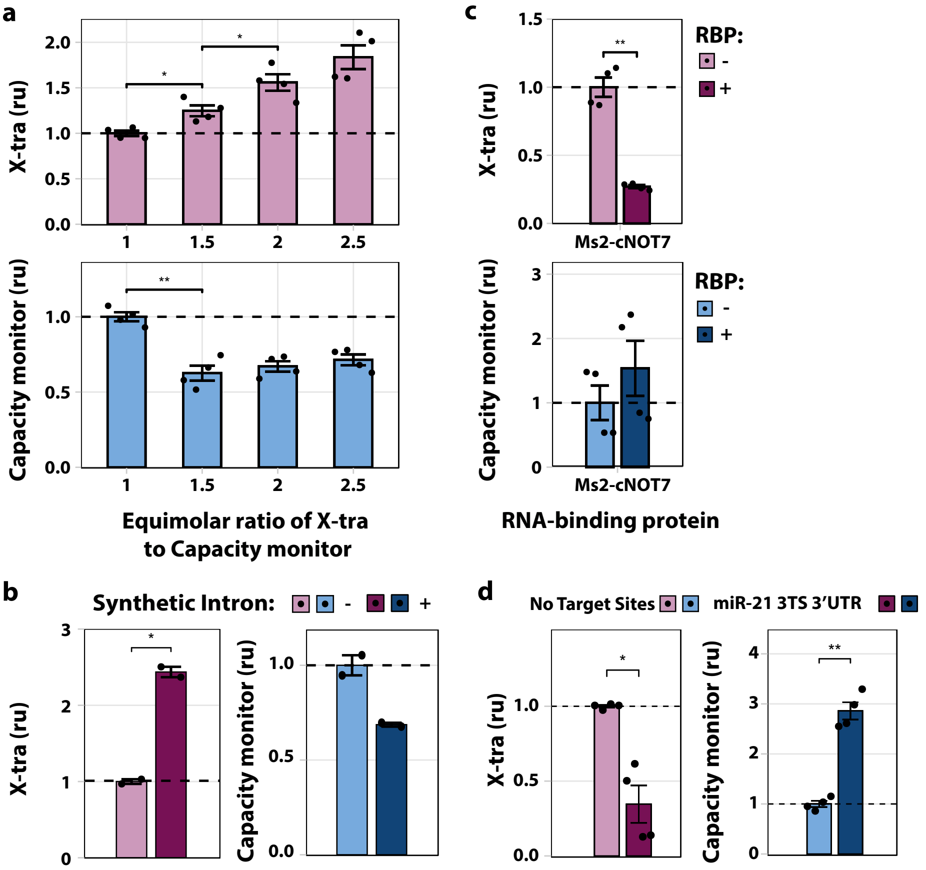


Figure A.10: Gene expression burden in HeLa cells. (a) Flow cytometry results of HeLa cells co-transfected with fixed amount of capacity monitor and increasing amount of X-tra (1:1 to 1:2.5 molar ratio), both under CMV promoter regulation. Data show the mean fluorescence normalized to its value at a plasmid molar ratio of 1. N=4 biological replicates. Source data are provided as a Source Data file. (b) Flow cytometry results of HeLa cells co-transfected with X-tra (mKate) which includes or not a synthetic intron in the 5'UTR, and capacity monitor (EGFP). Data show that when mKate expression is enhanced by the synthetic intron, EGFP levels decrease. Data are the mean fluorescence normalized to fluorescence values in the absence of the intron. N=2 biological replicates. Source data are provided as a Source Data file. (c) Flow cytometry results of HeLa cells co-transfected with 2kturn-EGFP or EGFP-8xMs2 (X-tra) and mKate (capacity monitor) in presence or absence of L7Ae or Ms2-cNOT7 respectively. Data show that when X-tra is down-regulated, the capacity monitor levels increase. Plot represents normalization of mean fluorescence values to the condition without RBP. N=4 biological replicates. Source data are provided as a Source Data file. (d) Flow cytometry results of HeLa cells co-transfected with mKate (X-tra) that includes or not miR-21 target sites in the 5'UTR, and EGFP (capacity monitor). Data show that capacity monitor levels are higher when the X-tra is downregulated by miR-21. N=4 biological replicates. Source data are provided as a Source Data file. Plot represents normalization of mean fluorescence values to the no target site condition. Data were acquired 48 hours post-transfection. Error bars represent the standard error. ru: relative units. Unpaired two-sided T-test. p-value: **<0.005, *<0.05.

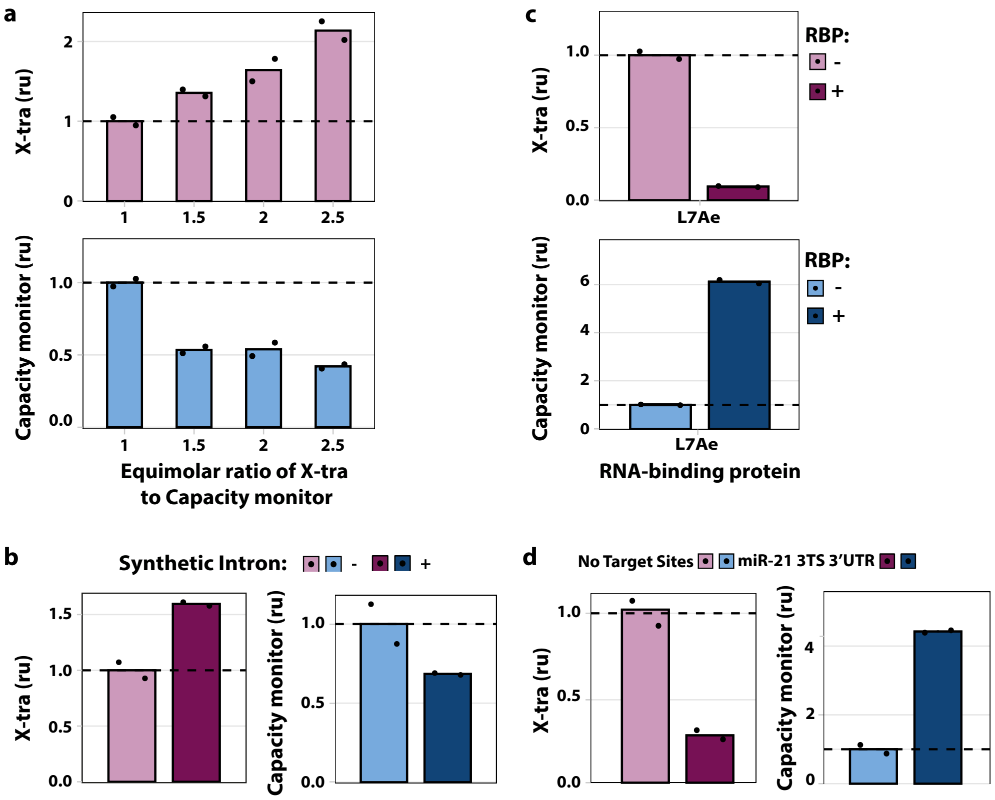


Figure A.11: Gene expression burden in CHO-K1 cells. (a) Flow cytometry results of CHO-K1 cells co-transfected with fixed amount of capacity monitor and increasing amount of X-tra (1:1 to 1:2.5 molar ratio), both under CMV promoter regulation. Data show the mean fluorescence normalized to its value at a plasmid molar ratio of 1. (b) Flow cytometry results of CHO-K1 cells co-transfected with X-tra (mKate) which includes or not a synthetic intron in the 5'UTR, and capacity monitor (EGFP). Data show that when mKate expression is enhanced by the synthetic intron, EGFP levels decrease. Data represent the mean fluorescence normalized to fluorescence values in the absence of the intron. (c) Flow cytometry results of CHO-K1 cells co-transfected with 2kturn-EGFP (X-tra) and mKate (capacity monitor) in presence or absence of L7Ae. Data show that when X-tra is down-regulated, the capacity monitor levels increase. Plot represents mean normalization of fluorescence values to the condition without L7Ae. (d) Flow cytometry results of CHO-K1 cells co-transfected with mKate (X-tra) that includes or not miR-21 target sites in the 5'UTR, and EGFP (capacity monitor). Data show that capacity monitor levels are higher when the X-tra is downregulated by miR-21. Plot represents normalization of mean fluorescence values to the no target site condition. Data were acquired 48 hours post-transfection. Error bars represent the standard error. ru: relative units. N=2 biological replicates. Source data are provided as a Source Data file.

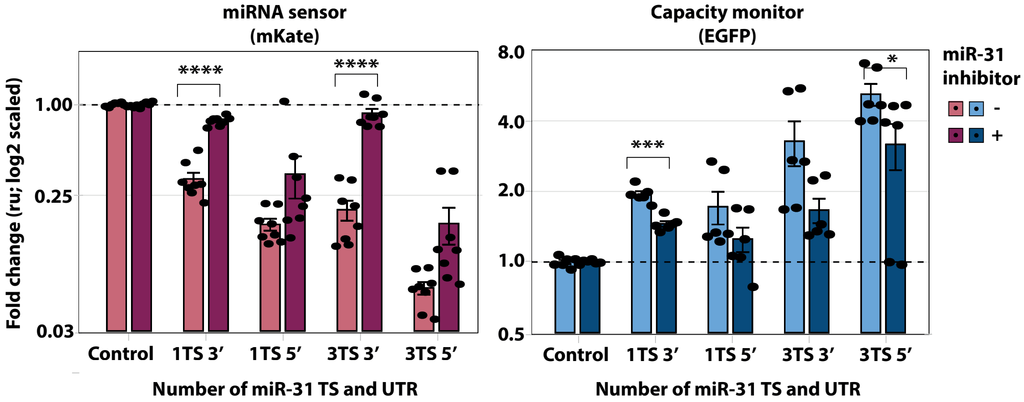


Figure A.12: Inhibition of miR-31 in H1299 cells. miR-31 activity impaired by a miR-31 inhibitor, leads to the rescue of miRNA sensor (mKate) expression in transfected H1299 cells. As a consequence, capacity monitor (EGFP) levels decrease. Both fluorescent proteins do not vary in the control. Data are expressed in logarithmic base 2 scale. Flow cytometry data were acquired 48 hours post-transfection and are plotted as mean fluorescence normalized on the control +/- SE. SE: standard error. ru: relative units. N=6 biological replicates. Source data are provided as a Source Data file. Unpaired two-sided T-test. p-value: ****<0.0001, ***<0.0005, *<0.05.

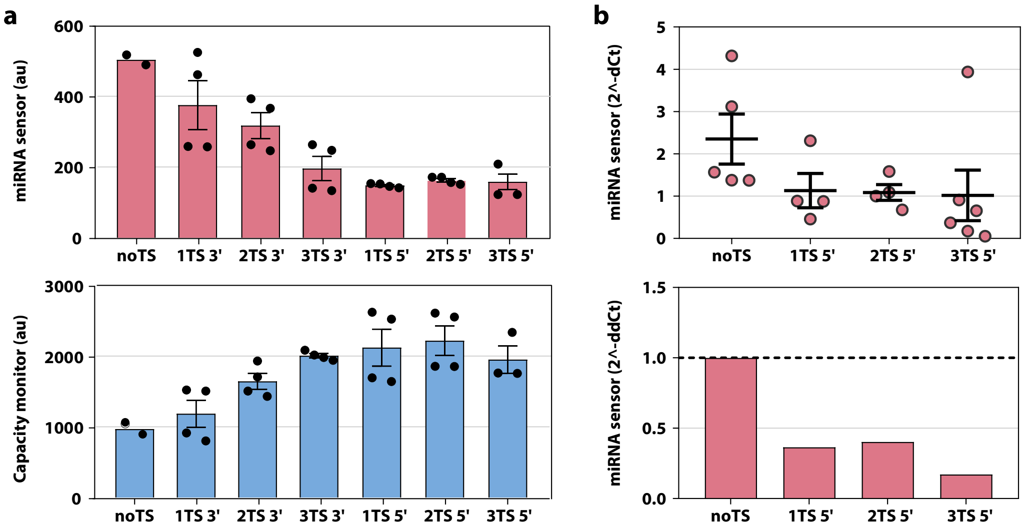


Figure A.13: miR-31 sensor in H1299. (a) Flow cytometry results of mKate-miR31-TS (miRNA sensor) co-transfected with EGFP (capacity monitor) in H1299 cells show that downregulation of miRNA sensor expression leads to an increase in capacity monitor levels. Data were acquired 48 h post transfection and are plotted as mean fluorescence +/- SE. SE: standard error. au: arbitrary units. N=2 biological replicates (N=2 in noTS sample; N=3 in 3TS 5' sample). Source data are provided as a Source Data file. (b) qPCR measurement confirms lower mRNA levels of miRNA sensor. Top, scattered dot plot of $2^{-\mathrm{dCt}}$ values. Data are plotted +/- SE. SE: standard error. Bottom, bar plot of the fold change measured with the $2^{-\mathrm{ddCt}}$ method8. Data were acquired 48 h post transfection. N=4 biological replicates. Source data are provided as a Source Data file.

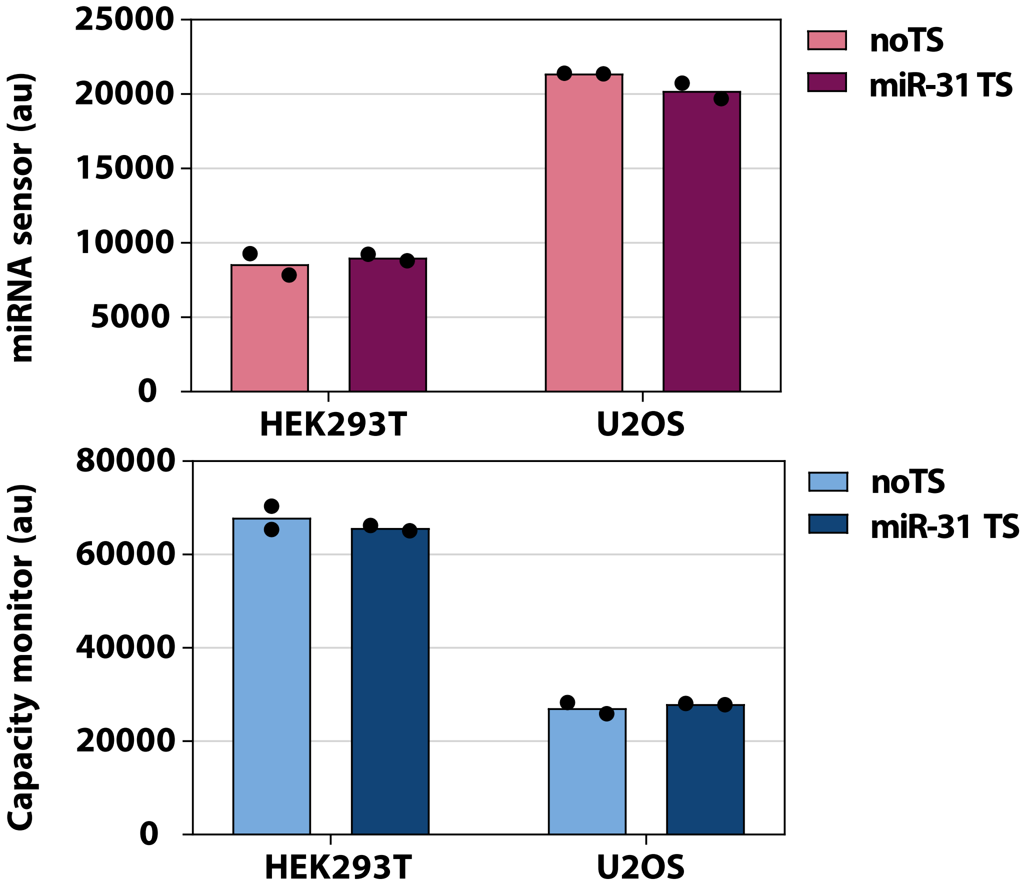


Figure A.14: The increase of capacity monitor levels is a consequence of miRNA regulation. U2OS and HEK293T cells were co-transfected with the 4-TS-3'UTR miR-31 sensor (miRNA sensor) and EGFP (capacity monitor). Both cell lines do not exhibit high expression of miR31, therefore miRNA sensor levels should not change. Data show that both miRNA sensor and the capacity monitor levels are comparable with and without miR-31 TS, indicating that the higher capacity monitor levels are indeed a consequence of miRNA activity. Data were acquired 48 h post-transfection +/- SE. SE: standard error. au: arbitrary units. N=2 biological replicates. Source data are provided as a Source Data file.

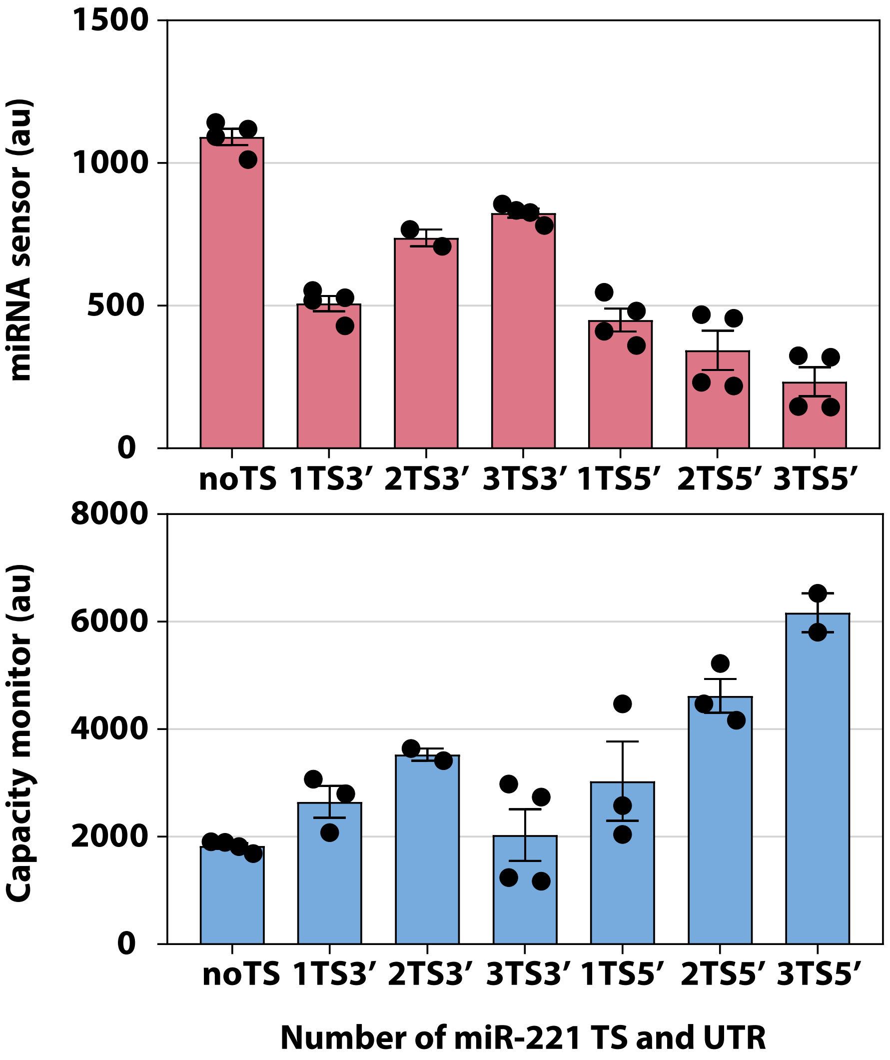


Figure A.15: miRNA-mediated resource re-allocation in the U2OS cell line. The miRNA sensor responds to miR-221, which is highly expressed in U2OS cells. Flow cytometry results of a co-transfection of mKate-miR221-TS (miRNA sensor) and EGFP (capacity monitor) in U2OS cells show the negative correlation of the two genes. Data were acquired 48 h post-transfection and are plotted as mean fluorescence +/- SE. SE: standard error. au: arbitrary units. N=4 biological replicates (N=2 in 2TS 3' sample). Source data are provided as a Source Data file.

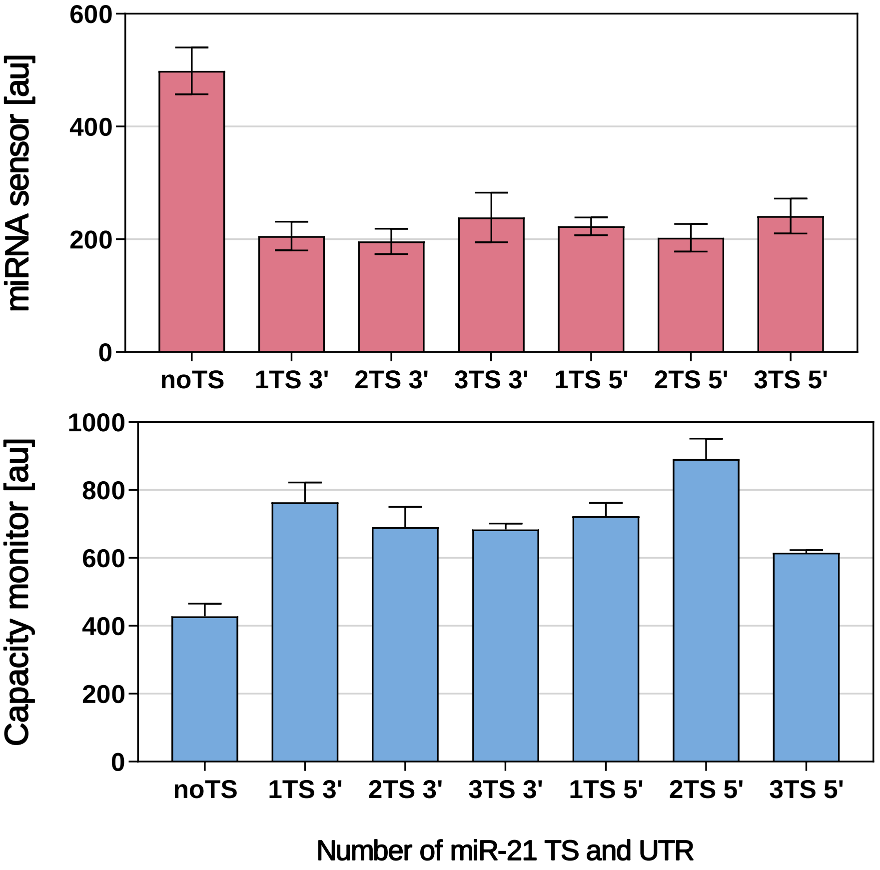


Figure A.16: miRNA-mediated resource re-allocation in the HeLa cell line. The miRNA sensor gene was designed for miR-21, which is highly expressed in HeLa cells. Flow cytometry results from a co-transfection of mKate-miR21-TS (miRNA sensor) and EGFP (capacity monitor) in HeLa cells. Interestingly, mKate downregulation seems to saturate already at 1TS3'. This may be due to the absolute levels of miR21 in this cell line. Data were acquired 48 h post transfection and are plotted as mean fluorescence +/- SE. SE: standard error. au: arbitrary units. N=2 biological replicates for noTS, 3TS 3', 3TS 5' and N=4 for 1TS 3', 2TS 3', 1TS 5' and 2TS 5'. Source data are provided as a Source Data file.

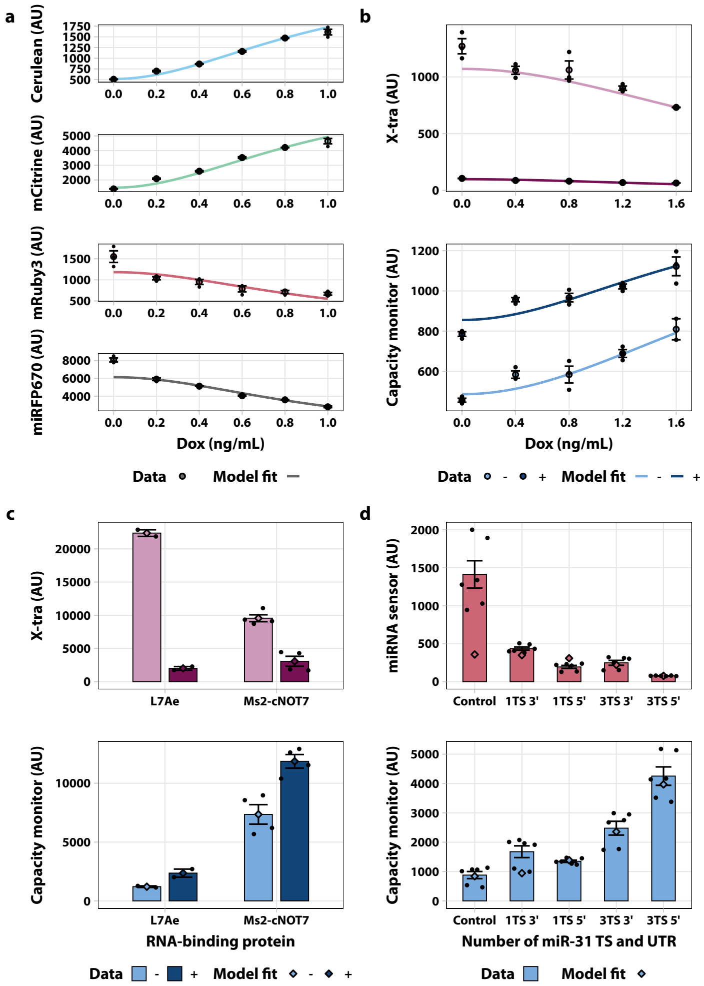


Figure A.17: Model fits for data displayed in Fig. 2.2b,e,g and Fig. 2.3b. The modeling framework described in Fig. 2.4a was applied to models of the genetic circuits used to generate the data in Fig. 2.2b,e,g and Fig. 2.3b. Detailed descriptions of these models can be found in Supplementary Note A.6. The parameters obtained are summarized in Supplementary Table A.34–A.38. Source data for the respective figures are provided as a Source Data file. The data are presented as mean +/- SE. SE: standard error. au: arbitrary units.

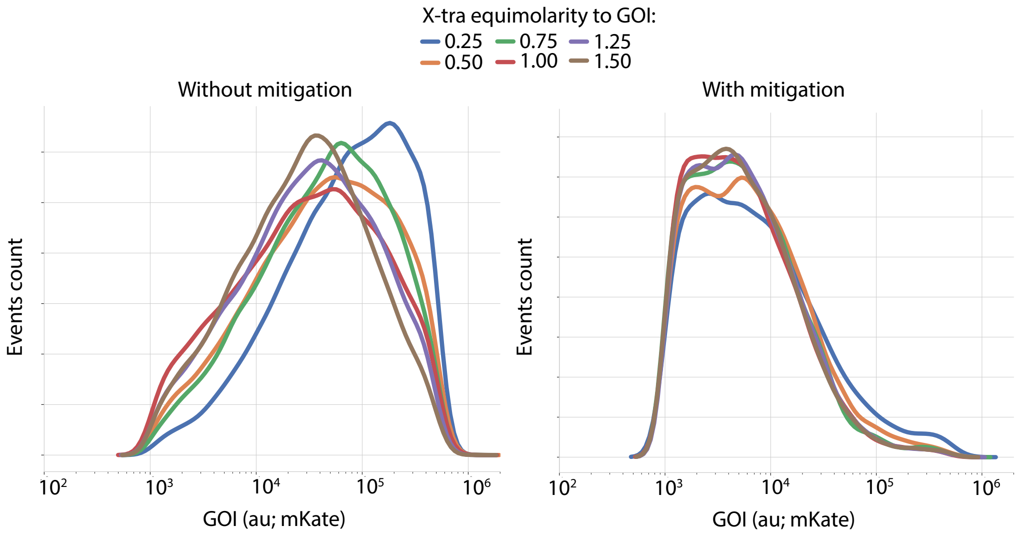


Figure A.18: GOI fluorescence distribution in co-transfected cells population. We compared the tolerance of mKate to increasing levels of X-tra gene in the absence or presence of an iFFL in which mKate includes miR-31 TS in the 5'UTR. The iFFL mitigation of resource competition is reflected by smaller shifts in GOI fluorescence at different equimolarities (right side). Data were acquired 48 h post-transfection. Source data are provided as a Source Data file.

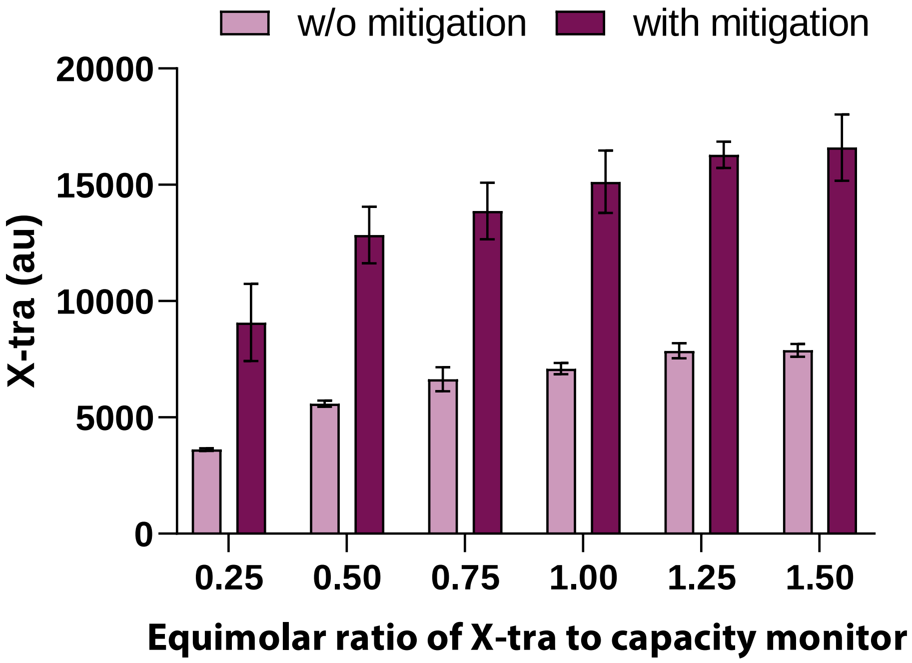


Figure A.19: X-tra absolute fluorescence in presence or absence of iFFL. X-tra expression levels of the plasmid titration experiment in Fig. 2.5c. With miR-31 iFFL mitigation, the absolute X-tra expression increases about 2 fold compared to w/o mitigation. Data is plotted as mean +/- SE. SE: standard error. au: arbitrary units. N=3 biological replicates. Source data are provided as a Source Data file.

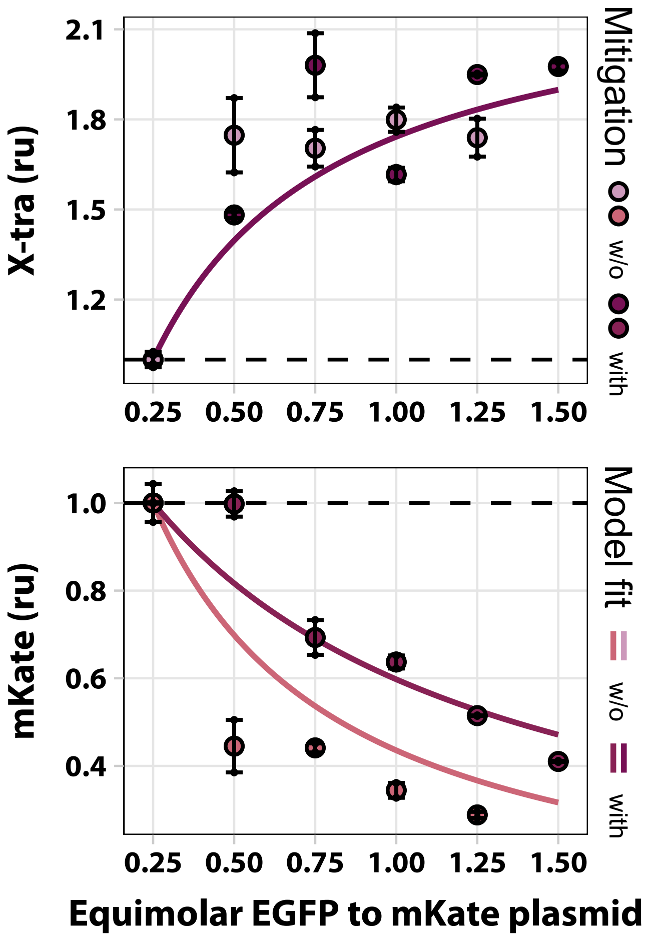


Figure A.20: The miR-221-based iFFL improves tolerance to exogenous gene load in U2OS cells. An iFFL whereby mKate includes miR-221 TS in the 5'UTR is less affected by the increased amount of the X-tra gene, as compared to the expression in the absence of miR-221 regulation. The model was unable to capture the differences in expression between the two conditions in the X-tra response due to the variability in the data. Therefore, the two lines plotted are exactly the same and it appears as if only one was plotted. Experimental data are normalized to the lowest equimolar ratio. The parameter values obtained by fitting are summarized in Supplementary Table A.31. Data were acquired 48 h post-transfection and are plotted +/- SE. SE: standard error. ru: relative units. N=2 biological replicates (N=1 for w/o Mitigation, 1.5 equimolar EGFP to mKate plasmid). Source data are provided as a Source Data file.

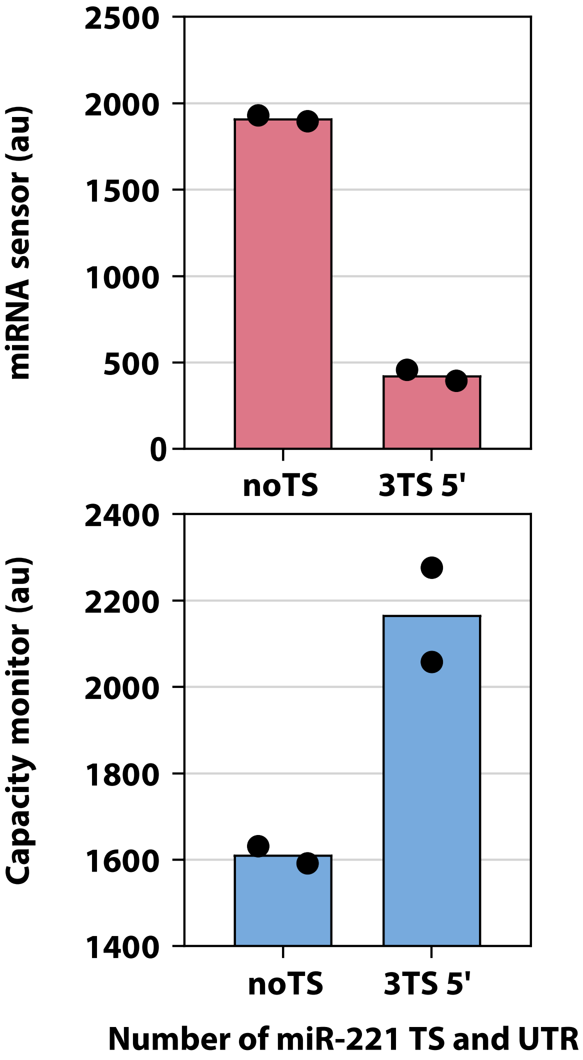


Figure A.21: miRNA-mediated resource re-allocation in the HEK293T cell line. The miRNA sensor gene was designed for miR-221, which is highly expressed in HEK293T cells. Flow cytometry results from a co-transfection of mKate-3xmiR221$_5$'UTR-TS (miRNA sensor) and EGFP (capacity monitor) in HEK293T cells. Data were acquired 48 h post-transfection and are plotted +/- SE. SE: standard error. au: arbitrary units. N=2 biological replicates. Source data are provided as a Source Data file.

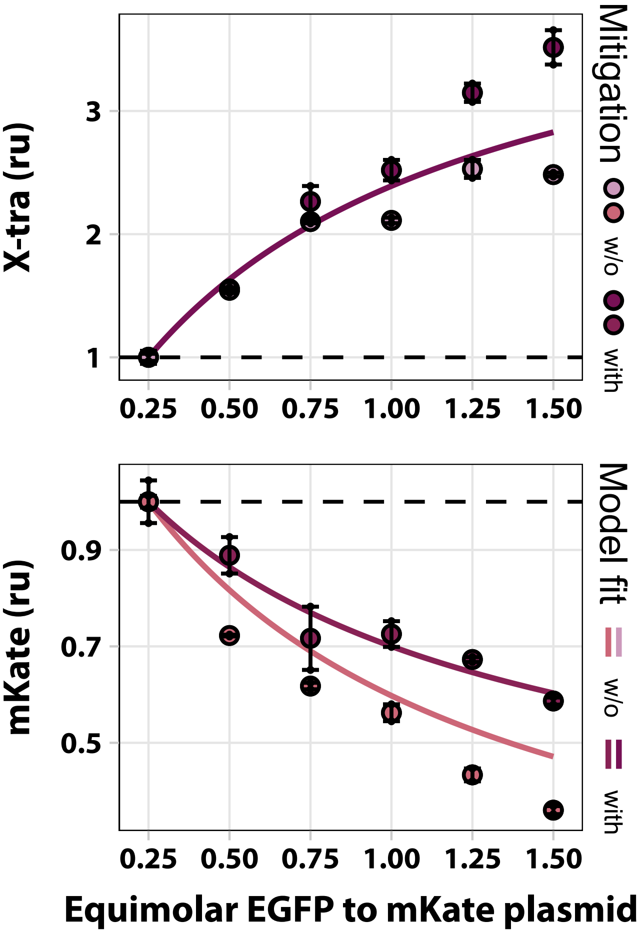


Figure A.22: The miR-221-based iFFL improves tolerance to exogenous gene load in HEK293T cells. We compared the tolerance of mKate to increasing levels of X-tra gene in the absence or presence of an iFFL whereby mKate includes miR-221 TS in the 5'UTR. The iFFL mitigates the effects of resource competition. The model was unable to capture the differences in expression between the two conditions in the X-tra response. Therefore, the two lines plotted are exactly the same and it appears as if only one was plotted. The parameter values obtained by fitting are summarized in Supplementary Table A.32. Experimental data are normalized to the lowest equimolar ratio. Data were acquired 48 h post-transfection and are plotted +/- SE. SE: standard error. ru: relative units. N=2 biological replicates. Source data are provided as a Source Data file.

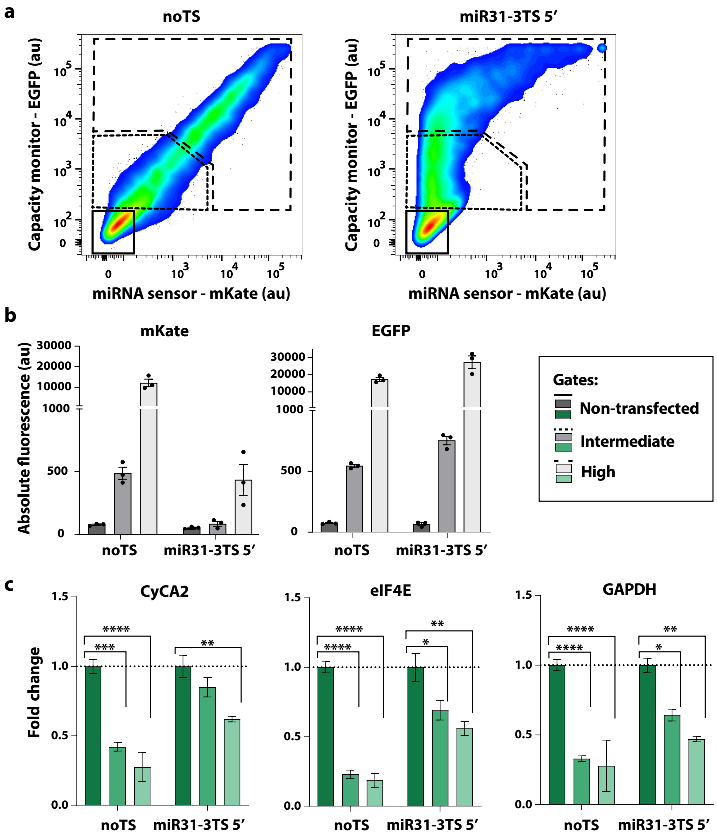


Figure A.23: Impact of transient plasmid transfection on endogenous genes in noTS and miR31-sensor samples. (a) H1299 cells were transfected with a bidirectional promoter plasmid encoding the fluorescent proteins EGFP (capacity monitor) and mKate (miRNA sensor), without (noTS, left) or with TS for miR-31 (miR31-sensor, right). Cells were sorted by fluorescence intensity 48 hours after transfection to collect non-transfected, intermediate transfected and high transfected cells from the same transfection plate. (b) Protein levels of EGFP and mKate in sorted populations of noTS and miR31-sensor transfected cells. Consistent with the gates, fluorescence intensity increases in intermediate and high transfected cells when compared to non-transfected cells. In agreement with data shown in Fig 2h and 3b,c, EGFP fluorescence is higher in miR31-sensor samples, while mKate is lower. Data are the mean fluorescence +/- SE. (c) mRNA levels of CyCA2, eIF4E and GAPDH in sorted samples. All three endogenous genes decrease in intermediate and high transfected cells as compared to non-transfected cells. However, in cells transfected with the miR31-sensor circuit the decrease of expression is lower. mRNA levels are normalized to the non-transfected population. Data were collected 48 hours after transfection and are represented as mean +/- SE. SE: standard error. au: arbitrary units. Unpaired two-sided T-test. p-value: ****<0.0001, ***<0.0005, **<0.005, *<0.05. N=3 biological replicates. Source data are provided as a Source Data file.

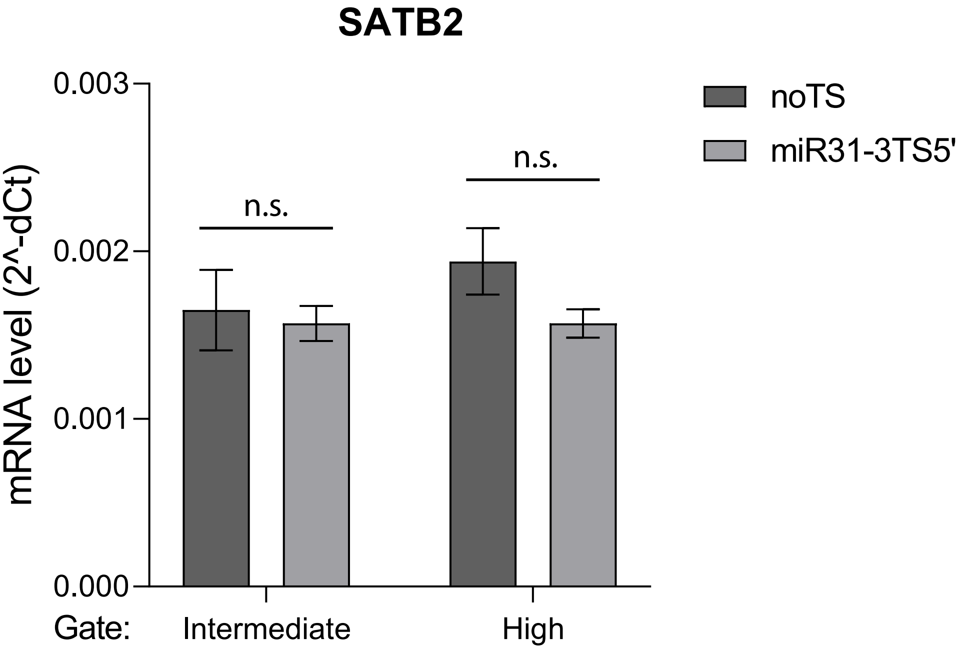


Figure A.24: Impact of noTS control vs miR-31-3TS5' transfections on the native miR-31 target SATB2. mRNA extracted from sorted populations as described in Supplementary Fig. A.22a and analyzed in Fig. 2.2d and Supplementary Fig. A.22c, was used to measure SATB2 levels relative to the internal reference 18S. SATB2 expression does not vary when miR-31 TS are present in our genetic circuit. Data represent the mean mRNA expression +/- SE. SE: standard error. ru: relative units. Unpaired two-sided T-test. N=3 biological replicates. Source data are provided as a Source Data file.

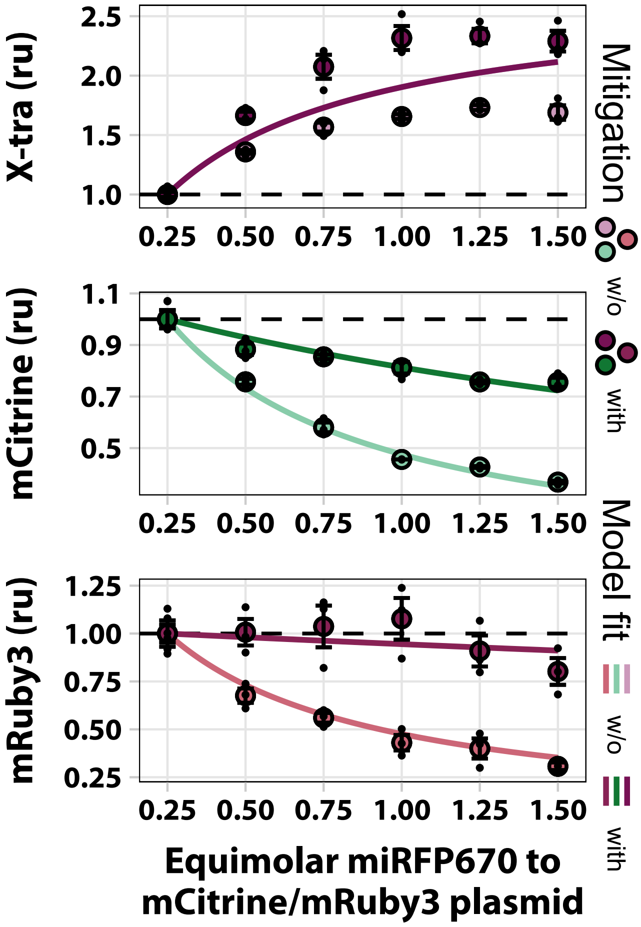


Figure A.25: The iFFL architecture improves tolerance to increase gene load in a 3-output system. Mouse embryonic stem cells were transfected with the miRNA mitigation iFFL shown in Fig. 2.5d. Light and dark colors represent gene expression levels in the absence or presence of mitigation. The solid lines show a model that includes resources, fit to the experimental data. Experimental data are normalized to the lowest equimolar ratio. The parameter values obtained by fitting are summarized in Supplementary Table A.33. All data were acquired 48 h post transfection and are plotted +/- SE. SE: standard error. ru: relative units. N=3 biological replicates. Source data are provided as a Source Data file.

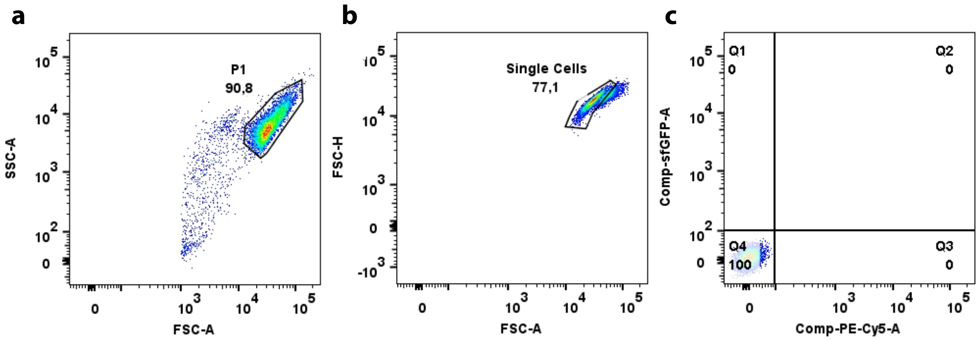


Figure A.26: FACS gating strategy. (a) The recorded events were gated in the FSC-A vs SSC-A channels to select the living cells population (P1). (b) The P1 was then gated in the FSC-A vs FSC-H channels to select the single cell population. (c) For each experiment a sample of non-transfected cells was used to set the positive threshold for each fluorescence. Cells selected following this pipeline were then analyzed.

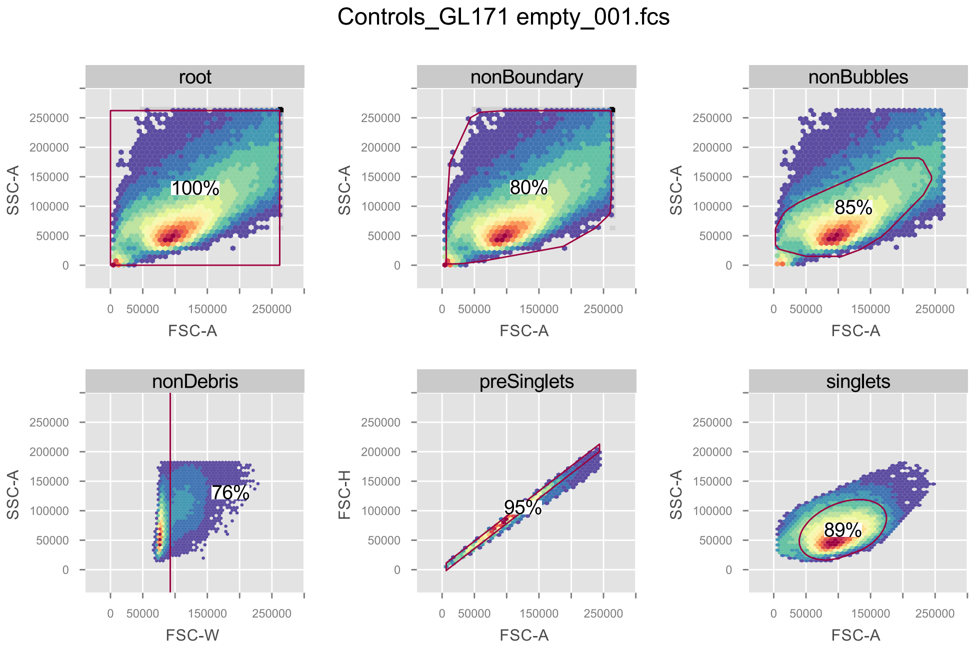


Figure A.27: Alternative FACS gating strategy. The plots show the hierarchical gating strategy implemented in a custom R script. The hierarchy progresses from left to right, top to bottom. Top left: The first gate removes events that potentially lie on the boundary of the detectable values. Top middle: This gate facilitates the subsequent gating by removing potential bubbles that were recorded. Top right: Here, a custom density-based gating strategy is employed to select for the living cell population and remove debris. Bottom left: In this gate the tail of the distribution in the FSC-W channel representing the bulk of the doublet event is removed. Bottom middle: The singlet population is further refined by gating in the FSC-H vs FSC-A channels. Bottom right: The resulting singlets are further refined by applying an ellipse gate around the point of highest density.

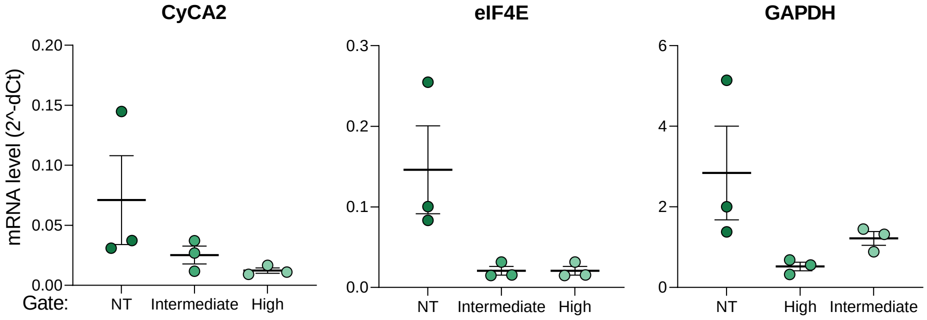


Figure A.28: mRNA level of data show in Fig. 2.2d. Individual measures of mRNA level in the 3 biological replicates. Source data are provided as a Source Data file. mRNA was extracted 48 hours post-transfection from sorted cells according to Supplementary Fig. A.5 gating. Data are plotted +/- SE. NT: Non Transfected. SE: standard error.

A.1 Supplementary Note 1

Derivation of the effective production rate constant $k_i^{eff}(A_1, \ldots, A_n)$. Here we show the derivation of the effective production rate constant as the quasi-steady-state solution for the complexes formed between substrate species A_i and their corresponding resource pool R. The reactions shown in Fig. 2.4a are:

$$\begin{aligned} \mathbf{A}_i + \mathbf{R} &\xrightarrow{\mathbf{k}_i^+} \mathbf{C}_i \\ \mathbf{C}_i &\xrightarrow{\mathbf{k}_i^-} \mathbf{A}_i + \mathbf{R} \\ \mathbf{C}_i &\xrightarrow{\mathbf{k}_i^{\mathbf{cat}}} \mathbf{A}_i + \mathbf{R} + \mathbf{B}_i \end{aligned}$$

In these reactions, B_i is the product formed. We apply the law of mass action to get the following ordinary differential equation for the complex species C_i:

$$\dot{C}_i(t) = k_i^+ R(t) A_i(t) - (k_i^- + k_i^{cat}) C_i(t)$$

Assuming a limited amount of available resources, we can express the free resources as:

$$R(t) = R^{total} - \sum_{j=1}^{n} C_j(t)$$

Since we would like to obtain a quasi-steady-state expression for any n, we cast the equations in matrix from:

$$\dot{\mathbf{C}} = \text{diag}(\mathbf{k}^+)(R^{total} - \mathbf{1}^\top \mathbf{C})\mathbf{A} - \text{diag}(\mathbf{k}^- + \mathbf{k}^{\mathbf{cat}})\mathbf{C}$$

Here we use bold font to denote vectors and diag($\mathbf{x}$) to represent the matrix with the vector $\mathbf{x}$ as its diagonal. When we set the left-hand side to zero, lump the binding, unbinding and production rates into $\mathbf{k_m}$ such that $k_{m_i} := \frac{k_i^- + k_i^{cat}}{k_i^+}$ and solve for $\mathbf{C}^*$ we get the expression:

$$\begin{aligned} \mathbf{C}^* &= R^{total}(I + \text{diag}(\mathbf{k_m^{-1}})\mathbf{A}^* \mathbf{1}^\top)^{-1} \text{diag}(\mathbf{k_m^{-1}})\mathbf{A}^* \\ &= R^{total}(I - \frac{\text{diag}(\mathbf{k_m^{-1}})\mathbf{A}^* \mathbf{1}^\top}{1 + \mathbf{1}^\top \text{diag}(\mathbf{k_m^{-1}}}) \text{diag}(\mathbf{k_m^{-1}})\mathbf{A}^* \\ &= \frac{R^{total} \text{diag}(\mathbf{k_m^{-1}})\mathbf{A}^*}{1 + \mathbf{1}^\top \text{diag}(\mathbf{k_m^{-1}})\mathbf{A}^*} \end{aligned}$$

Where the second step is an application of the Sherman-Morrison formula.

The effective production rate is obtained by multiplying this last expression with diag($\mathbf{k}^{\mathbf{cat}}$). For an individual element this gives:

$$k_i^{eff}(A_1, \ldots, A_n) := k_i^{cat} R^{total} \frac{k_{m_i}^{-1} A_i}{1 + \sum_{j=1}^{n} k_{m_j}^{-1} A_j}$$

A.2 Supplementary Note 2

A simple model of plasmid copy number induced competition for limited resources.

To derive the equations used in Fig. 2.2a of the main text the following model was used:

$$\dot{M}_i = k_{M_i}^{eff}(G_1, \dots, G_n)G_i - \delta_{M_i}M_i$$
$$\dot{P}_i = k_{P_i}^{eff}(M_1, \dots, M_n)M_i - \delta_{P_i}P_i$$

Here, we distinguish between two potential pools of shared limited resources. One for mRNA production and one for protein production. The mRNA species are denoted by the subscript letter M and the protein species are denoted by the subscript letter P. The copy number of the genes are denoted by G_i and are assumed to be constant. The degradation rates of each species is represented by a δ subscripted with the respective species' name. Using the definition:

$$k_{B_i}^{eff}(A_1, \dots, A_n) \ := k_{B_i}^{cat} R^{total} \frac{k_{m_{B_i}}^{-1}}{1 + \sum_{j=1}^{n} k_{m_{B_j}}^{-1} A_j}$$

The equations can be written as:

$$\dot{M}_i = k_{M_i}^{cat} R_M^{total} \frac{k_{m_{M_i}}^{-1}}{1 + \sum_{j=1}^{n} k_{m_{M_j}}^{-1} G_j} G_i - \delta_{M_i} M_i$$
$$\dot{P}_i = k_{P_i}^{cat} R_P^{total} \frac{k_{m_{P_i}}^{-1}}{1 + \sum_{j=1}^{n} k_{m_{P_j}}^{-1} M_j} M_i - \delta_{P_i} P_i$$

We solve for the steady state expressions of each of the species by setting the left-hand side of the equations to zero. After simplifying we obtain the following expressions:

$$M_i^* = \frac{k_{M_i}^{cat} R_M^{total}}{\delta_{M_i}} \frac{k_{m_{M_i}}^{-1} G_i}{1 + \sum_{j=1}^{n} k_{m_{M_j}}^{-1} G_j}$$

$$P_i^* = \frac{k_{M_i}^{cat} R_M^{total}}{\delta_{M_i}} \frac{k_{P_i}^{cat} R_P^{total}}{\delta_{P_i}} \frac{k_{m_{M_i}}^{-1} k_{m_{P_i}}^{-1} G_i}{1 + \sum_{j=1}^{n} k_{m_{M_j}}^{-1} \left(1 + k_{m_{P_j}}^{-1} \frac{k_{M_j}^{cat} R_M^{total}}{\delta_{M_j}}\right) G_j}$$

The normalized expressions as shown in Fig. 2.2a are further defined as $\hat{P}_i^* = \frac{P_i^*}{P_i^*|_{G_k=G_{k_0}}}$ with $k \neq i$, which gives:

$$\hat{P}_i^* = \frac{1 + k_{m_{M_k}}^{-1}\left(1 + k_{m_{P_k}}^{-1}\frac{k_{M_k}^{cat} R_M^{total}}{\delta_{M_k}}\right) G_{k_0} + \sum_{j\neq k}^{n} k_{m_{M_j}}^{-1}\left(1 + k_{m_{P_j}}^{-1}\frac{k_{M_j}^{cat} R_M^{total}}{\delta_{M_j}}\right) G_j}{1 + k_{m_{M_k}}^{-1}\left(1 + k_{m_{P_k}}^{-1}\frac{k_{M_k}^{cat} R_M^{total}}{\delta_{M_k}}\right) G_k + \sum_{j\neq k}^{n} k_{m_{M_j}}^{-1}\left(1 + k_{m_{P_j}}^{-1}\frac{k_{M_j}^{cat} R_M^{total}}{\delta_{M_j}}\right) G_j}$$

Specifically for Fig. 2.2a n was set to 2.

A.3 Supplementary Note 3

Normalized gene expression of low absolute expression levels are more sensitive to reduced availability of resources.

To show that normalized expression is more sensitive to burden at low expression levels we take the general term for the protein levels derived in Supplementary Note A.1.

$$P_i^* = \frac{k_{M_i}^{cat} R_M^{total}}{\delta_{M_i}} \frac{k_{P_i}^{cat} R_P^{total}}{\delta_{P_i}} \frac{k_{m_{M_i}}^{-1} k_{m_{P_i}}^{-1} G_i}{1 + \sum_{j=1}^{n} k_{m_{M_j}}^{-1}\left(1 + k_{m_{P_j}}^{-1}\frac{k_{M_j}^{cat} R_M^{total}}{\delta_{M_j}}\right) G_j}$$

To simplify the above term we lump parameters by setting:

$$\beta_{M_i} := k_{m_{P_i}}^{-1}\frac{k_{M_i}^{cat} R_M^{total}}{\delta_{M_i}}$$

$$\alpha_{P_i} := \frac{k_{P_i}^{cat} R_P^{total}}{\delta_{P_i}}$$

Which gives:

$$P_i^* = \frac{\alpha_{P_i}\beta_{M_i} k_{m_{M_i}}^{-1} G_i}{1 + \sum_{j=1}^{n} k_{m_{M_j}}^{-1}\left(1 + \beta_{M_j}\right) G_j}$$

From this term we can write the expression level normalized with respect to the expression level at G_{k_0} as:

$$\hat{P}_i^* = \frac{P_i^*}{P_i^*|_{G_k=G_{k_0}}} = \frac{1 + k_{m_{M_k}}^{-1}\left(1 + \beta_{M_k}\right) G_{k_0} + \sum_{j\neq k}^{n} k_{m_{M_j}}^{-1}\left(1 + \beta_{M_j}\right) G_j}{1 + k_{m_{M_k}}^{-1}\left(1 + \beta_{M_k}\right) G_k + \sum_{j\neq k}^{n} k_{m_{M_j}}^{-1}\left(1 + \beta_{M_j}\right) G_j}$$

To test whether the normalized expression for a gene expressing at low levels is less sensitive to reduced resource availability than its high expressing counterpart we wish to evaluate the following inequality $\hat{P}_i^* < \hat{P}_i^*|_{G_i=\bar{G}_i}$ where the low expression is given by the assumption that $\bar{G}_i < G_i$. To simplify this inequality we write it as the equivalent inequality given by $\text{sgn}(\hat{P}_i^* - \hat{P}_i^*|_{G_i=\bar{G}_i}) < 0$ where sgn denotes the sign function. Plugging in the expression for $\hat{P}_i^*$ and $\hat{P}_i^*|_{G_i=\bar{G}_i}$ gives:

$$\begin{aligned}
&\text{sgn}\left(\frac{1 + k_{m_{M_k}}^{-1}\left(1+\beta_{M_k}\right)G_{k_0} + \sum_{j\neq k}^{n} k_{m_{M_j}}^{-1}\left(1+\beta_{M_j}\right)G_j}{1 + k_{m_{M_k}}^{-1}\left(1+\beta_{M_k}\right)G_k + \sum_{j\neq k}^{n} k_{m_{M_j}}^{-1}\left(1+\beta_{M_j}\right)G_j}\right.\\
&\left. - \frac{1 + k_{m_{M_i}}^{-1}\left(1+\beta_{M_i}\right)\bar{G}_i + k_{m_{M_k}}^{-1}\left(1+\beta_{M_k}\right)G_{k_0} + \sum_{j\neq\{i,k\}}^{n} k_{m_{M_j}}^{-1}\left(1+\beta_{M_j}\right)G_j}{1 + k_{m_{M_i}}^{-1}\left(1+\beta_{M_i}\right)\bar{G}_i + k_{m_{M_k}}^{-1}\left(1+\beta_{M_k}\right)G_k + \sum_{j\neq\{i,k\}}^{n} k_{m_{M_j}}^{-1}\left(1+\beta_{M_j}\right)G_j}\right) < 0
\end{aligned} \tag{A.1}$$

Given that all parameters are positive and we demand that $G_k > G_{k_0}$ the term on the left hand side of the inequality above reduces to $\text{sgn}(G_i - \bar{G}_i) < 0$. Given our initial assumption that $\bar{G}_i < G_i$ and the requirement that both are positive real numbers we find that the claim stated in the inequality is false because the sign function evaluates to 1. Therefore, the model shows that the normalized expression of a gene expressing at low absolute levels will be more affected by resource availability than its high expressing counterpart.

A.4 Supplementary Note 4

Model for the topologies from Lillacci et al. (Lillacci et al., 2018).

The model for the four topologies can be given by a system of ordinary differential equations, where setting the individual repression rates η_{M_1} and/or η_{M_2} of the microRNA to zero specifies the topology. More specifically, $\eta_{M_1} = 0$ and $\eta_{M_2} = 0$ is the open-loop (OLP) topology, $\eta_{M_1} = 0$ and $\eta_{M_2} > 0$ the incoherent feedforward (IFF) topology, $\eta_{M_1} > 0$ and $\eta_{M_2} = 0$ is the feedback

(FBK) topology and $\eta_{M_1} > 0$ and $\eta_{M_2} > 0$ is the hybrid (HYB) topology.

$$
\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_2 f(P_1), G_3)G_1 - (\delta_{M_1} + \eta_{M_1} m)M_1\\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_2 f(P_1), G_3)G_2 f(P_1) - (\delta_{M_2} + \eta_{M_2} m)M_2\\
\dot{M}_3 &= k_{M_3}^{eff}(G_1, G_2 f(P_1), G_3)G_3 - \delta_{M_3} M_3\\
\dot{m} &= k_{m}^{eff}(G_1, G_2 f(P_1), G_3)G_2 f(P_1) - \delta_m m\\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2, M_3)M_1 - \delta_{P_1} P_1\\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2, M_3)M_2 - \delta_{P_2} P_2\\
\dot{P}_1 &= k_{P_3}^{eff}(M_1, M_2, M_3)M_3 - \delta_{P_3} P_3
\end{aligned}
$$

In this system of equations the species M_1, M_2 and M_3 correspond to tTA-Cer. mRNA, DsRed mRNA and mCitr. mRNA respectively as shown in Fig. 2.4b and c. P_1 denotes the transcriptional activator tTA-Cer., P_2 denotes the fluorescent protein DsRed and P_3 denotes the fluorescent protein mCitrine. m denotes miR-FF4 expressed from the same gene as DsRed. Furthermore, G_1, G_2 and G_3 correspond to the plasmid copy number of tTA-Cer., DsRed and mCitr. respectively. The rates beginning with a δ denote the degradation rates of the species written in the subscript. The rates beginning with η correspond to the repression rates of the microRNA FF4. Lastly, the transcriptional activation was modeled by a hill-type function $f(x) := \frac{x^h}{\kappa^h + x^h}$. The steady states for the protein species used for fitting can be written as:

$$
P_1^* = \frac{\alpha_{P_1^*} \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_1}\right) + \gamma_{M_3}}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_1}\right) + \gamma_{M_3}} + \frac{\beta_{M_2} \gamma_{M_2} f(P_1^*)}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_2}\right) + \gamma_{M_3}} + \frac{\beta_{M_3} \gamma_{M_3}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*) + \gamma_{M_3}}}
$$

$$
P_2^* = \frac{\alpha_{P_2} \frac{\beta_{M_2} \gamma_{M_2} f(P_1^*)}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_2}\right) + \gamma_{M_3}}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_1}\right) + \gamma_{M_3}} + \frac{\beta_{M_2} \gamma_{M_2} f(P_1^*)}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_2}\right) + \gamma_{M_3}} + \frac{\beta_{M_3} \gamma_{M_3}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*) + \gamma_{M_3}}}
$$

$$
P_3^* = \frac{\alpha_{P_3} \frac{\beta_{M_3} \gamma_{M_3}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*) + \gamma_{M_3}}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_1}\right) + \gamma_{M_3}} + \frac{\beta_{M_2} \gamma_{M_2} f(P_1^*)}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*)\left(1+\theta_{M_2}\right) + \gamma_{M_3}} + \frac{\beta_{M_3} \gamma_{M_3}}{1+G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} f(P_1^*) + \gamma_{M_3}}}
$$

Here, α_{P_i} and β_{M_i} are the same as defined in Supplementary Note A.2. Additionally, $\gamma_{M_i} := k_{m_{M_i}}^{-1} G_i$ for $i \in \{2, 3\}$ and $\theta_{M_i} := \frac{\eta_{M_i}}{\delta_{M_i}} \frac{k_{M_2}^{cat} R_M^{total}}{\delta_m}$ were introduced. The equations were fit in the implicit form shown because a closed form solution could not be obtained.

A.5 Supplementary Note 5

Models for endogenous microRNA-based iFFL and synthetic microRNA-based iFFL circuits.

Endogenous microRNA-based iFFL: The system of equations used to derive the steady state expressions is given by:

$$\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_2, G_m)G_1 - \delta_{M_1} M_1 \\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_2, G_m)G_2 - \left(\delta_{M_2} + \eta_{M_2} m\right) M_2 \\
\dot{m} &= k_m^{eff}(G_1, G_2, G_m)G_m - \delta_m m \\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2)M_1 - \delta_{P_1} P_1 \\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2)M_2 - \delta_{P_2} P_2
\end{aligned}$$

Here, M_1 and M_2 represent the mRNA species for the fluorescent proteins EGFP (X-tra) and mKate (GOI) respectively. P_1 and P_2 correspond to the proteins themselves and m denotes the microRNA miR-31. As in Supplementary Notes A.1 and A.3, the degradation rates are shown as δ subscripted with the species they correspond to and the repression rates are shown as η subscripted with the respective species. G_1 and G_2 are the plasmid copy number of the EGFP and mKate plasmids respectively and G_m is the copy number of the microRNA on the genome. The steady states for the protein species can be written as:

$$\begin{aligned}
P_1^* &= \frac{\alpha_{P_1} \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m} + \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}} \\
P_2^* &= \frac{\alpha_{P_2} \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m} + \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}}
\end{aligned}$$

For fitting, the expressions were normalized the same way as introduced in Supplementary Notes A.1 and A.2 which yields the expressions:

$$\begin{aligned}
\hat{P}_1^* &= \frac{G_1 \left(1 + G_{1_0} k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m\right)}{G_{1_0} \left(1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m\right)} \frac{1 + \frac{\beta_{M_1} G_{1_0} k_{m_{M_1}}^{-1}}{1 + G_{1_0} k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m} + \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_{1_0} k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m} + \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}} \\
\hat{P}_2^* &= \frac{1 + G_{1_0} k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)} \frac{1 + \frac{\beta_{M_1} G_{1_0} k_{m_{M_1}}^{-1}}{1 + G_{1_0} k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m} + \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_{1_0} k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}}{1 + \frac{\beta_{M_1} G_1 k_{m_{M_1}}^{-1}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m} + \frac{\beta_{M_2} \gamma_{M_2}}{1 + G_1 k_{m_{M_1}}^{-1} + \gamma_{M_2} + \gamma_m \left(1 + \theta_{M_2}\right)}}
\end{aligned}$$

Synthetic microRNA-based iFFL: The system of equations used to obtain the steady state expressions is given by:

$$
\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_2)G_1 - \delta_{M_1} M_1 \\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_2)G_2 - \left(\delta_{M_2} + \eta_{M_2} m\right) M_2 \\
\dot{M}_3 &= k_{M_3}^{eff}(G_1, G_2)G_2 - \left(\delta_{M_3} + \eta_{M_3} m\right) M_3 \\
\dot{m} &= k_{m}^{eff}(G_1, G_2)G_2 - \delta_m m \\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2, M_3)M_1 - \delta_{P_1} P_1 \\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2, M_3)M_2 - \delta_{P_2} P_2 \\
\dot{P}_1 &= k_{P_3}^{eff}(M_1, M_2, M_3)M_3 - \delta_{P_3} P_3
\end{aligned}
$$

Here, M_1, M_2 and M_3 denote the mRNA species of the fluorescent proteins miRFP670 (X-tra), mCitrine (GOI1) and mRuby3 (GOI2) respectively. Similarly, P_1, P_2 and P_3 denote their protein species and m represents the microRNA FF4. G_1 corresponds to the plasmid which encodes miRFP670 and G_2 corresponds to the plasmid which encodes both the transcriptional unit of mCitrine and the transcriptional unit of mRuby3 and the microRNA. Again, rates beginning with δ describe degradation rates, while rates beginning with η denote the repression by the microRNA. Compared to the endogenous system, the microRNA FF4 is produced from the same gene as mRuby3 and therefore we model their production rates in a similar manner. The steady states for the protein species can be obtained to be:

$$
P_1^* = \frac{\alpha_{P_1}\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}}{1+\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}
$$

$$
P_2^* = \frac{\alpha_{P_2}\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}}{1+\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}
$$

$$
P_3^* = \frac{\alpha_{P_3}\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}{1+\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}
$$

For fitting we again use the expression normalized to the first titration of miRFP670.

$$
\hat{P}_1^* = \frac{G_1\left(1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\right)}{G_{1_0}\left(1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\right)}\frac{1+\frac{\beta_{M_1}G_{1_0}k_{mM_1}^{-1}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}{1+\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}
$$

$$
\hat{P}_2^* = \frac{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}\frac{1+\frac{\beta_{M_1}G_{1_0}k_{mM_1}^{-1}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}{1+\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}
$$

$$
\hat{P}_1^* = \frac{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}\frac{1+\frac{\beta_{M_1}G_{1_0}k_{mM_1}^{-1}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_{1_0}k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}{1+\frac{\beta_{M_1}G_1k_{mM_1}^{-1}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_2}\gamma_{M_2}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_2}\right)}+\frac{\beta_{M_3}\gamma_{M_3}}{1+G_1k_{mM_1}^{-1}+\gamma_{M_2}+\gamma_{M_3}\left(1+\theta_{M_3}\right)}}
$$

A.6 Supplementary Note 6

Models for DOX, HDV, RNA-binding protein and miRNA experiments.

Doxycycline (DOX) titration: The system of ordinary differential equations used to obtain the steady state expressions is:

$$
\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_1, G_2 f(P_1), G_2 f(P_1))G_1 - \delta_{M_1} M_1\\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_1, G_2 f(P_1), G_2 f(P_1))G_1 - \delta_{M_2} M_2\\
\dot{M}_3 &= k_{M_3}^{eff}(G_1, G_1, G_2 f(P_1), G_2 f(P_1))G_2 f(P_1) - \delta_{M_3} M_3\\
\dot{M}_4 &= k_{M_4}^{eff}(G_1, G_1, G_2 f(P_1), G_2 f(P_1))G_2 f(P_1) - \delta_{M_4} M_4\\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2, M_3, M_4)M_1 - \delta_{P_1} P_1\\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2, M_3, M_4)M_2 - \delta_{P_2} P_2\\
\dot{P}_3 &= k_{P_3}^{eff}(M_1, M_2, M_3, M_4)M_3 - \delta_{P_3} P_3\\
\dot{P}_4 &= k_{P_4}^{eff}(M_1, M_2, M_3, M_4)M_4 - \delta_{P_4} P_4
\end{aligned}
$$

Here, the species M_1 and M_2 represent the mRNA species of the constitutively expressed transcriptional activator tagged with a fluorescent protein tTA-Cerulean (capacity monitor; P_1) and the fluorescent protein mCitrine (P_2). The species M_3 and M_4 represent the mRNA species of the two fluorescent proteins mRuby3 (X-tra; P_3) and miRFP670 (P_4). These two genes are expressed from a bidirectional tTA activatible promoter. By setting the left-hand side to zero, we can solve for the steady state of the protein species:

$$
P_1^* = \frac{\alpha_{P_1} \frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}{1 + \frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_3}\gamma_{M_3}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_4}\gamma_{M_4}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}
$$

$$
P_2^* = \frac{\alpha_{P_2} \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}{1 + \frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_3}\gamma_{M_3}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_4}\gamma_{M_4}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}
$$

$$
P_3^* = \frac{\alpha_{P_3} \frac{\beta_{M_3}\gamma_{M_3}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}{1 + \frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_3}\gamma_{M_3}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_4}\gamma_{M_4}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}
$$

$$
P_4^* = \frac{\alpha_{P_4} \frac{\beta_{M_4}\gamma_{M_4}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}{1 + \frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_3}\gamma_{M_3}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)} + \frac{\beta_{M_4}\gamma_{M_4}f(P_1^*)}{1+\gamma_{M_1}+\gamma_{M_2}+(\gamma_{M_3}+\gamma_{M_4})f(P_1^*)}}
$$

These expressions were fit to the data. The parameters were lumped according to:

$$
\begin{aligned}
\alpha_{P_i} &:= \frac{k_{P_i}^{cat} R_P^{total}}{\delta_{P_i}}\\
\beta_{M_i} &:= k_{mP_i}^{-1} \frac{k_{M_i}^{cat} R_M^{total}}{\delta_{M_i}}\\
\gamma_{M_i} &:= k_{mM_i}^{-1} G_i
\end{aligned}
$$

Further, to reflect the inhibiting action of Doxycycline we use the product of an activating and a inhibiting hill-type function:

$$
f(x, DOX) = \frac{x^{h_a}}{\kappa_a^{h_a} + x^{h_a}} \frac{\kappa_i^{h_i}}{\kappa_i^{h_i} + DOX^{h_i}}
$$

HDV: The system of ordinary differential equations used to obtain the steady state expressions is:

$$\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1 f(P_1), G_1) G_1 f(P_2) - \delta_{M_1} M_1 \\
\dot{M}_2 &= k_{M_2}^{eff}(G_1 f(P_1), G_1) G_1 - \delta_{M_2} M_2 \\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2) M_1 - \delta_{P_1} P_1 \\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2) M_2 - \delta_{P_2} P_2
\end{aligned}$$

In these equations, the HDV-tagged mRNA species of X-tra (mCitrine) is denoted as M_1 and X-tra itself is denoted by P_1. Further, the species M_2 and P_2 represent the capacity monitor (tTA-P2A-mRuby3) mRNA and protein, respectively. To model the different expression of P_1 with and without the HDV ribozyme, we introduce an additional parameter ω_{M_1}. This parameter captures the reduction in the production rate of mature mRNA M_1 in the presence of the ribozyme. In the absence of the ribozyme, we set $\omega_{M_1} = 1$.

$$\begin{aligned}
P_1^* &= \frac{\alpha_{P_1} \frac{\frac{\beta_{M_1}}{\omega_{M_1}} \gamma_{M_1} f(P_2^*)}{1+\gamma_{M_1} f(P_2^*)+\gamma_{M_2}}}{1 + \frac{\frac{\beta_{M_1}}{\omega_{M_1}} \gamma_{M_1} f(P_2^*)}{1+\gamma_{M_1} f(P_2^*)+\gamma_{M_2}} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1} f(P_2^*)+\gamma_{M_2}}} \\
P_2^* &= \frac{\alpha_{P_2} \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1} f(P_2^*)+\gamma_{M_2}}}{1 + \frac{\frac{\beta_{M_1}}{\omega_{M_1}} \gamma_{M_1} f(P_2^*)}{1+\gamma_{M_1} f(P_2^*)+\gamma_{M_2}} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1} f(P_2^*)+\gamma_{M_2}}}
\end{aligned}$$

These expressions were fit to the data. The parameters were lumped according to the same definitions as in the Doxycycline titration model above.

RNA-binding proteins:

L7Ae:

The system of ordinary differential equations used to obtain the steady state expressions is:

$$\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_2, G_3) G_1 - (\delta_{M_1} + \eta P_3) M_1 + \nu C_{M_1,P_3} \\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_2, G_3) G_2 - \delta_{M_2} M_2 \\
\dot{M}_3 &= k_{M_3}^{eff}(G_1, G_2, G_3) G_3 - \delta_{M_3} M_3 \\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2, M_3) M_1 - \delta_{P_1} P_1 \\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2, M_3) M_2 - \delta_{P_2} P_2 \\
\dot{P}_3 &= k_{P_3}^{eff}(M_1, M_2, M_3) M_3 - \delta_{P_3} P_3 \\
\dot{C}_{M_1,P_3} &= \eta M_1 P_3 - (\nu + \delta_{C_{M_1,P_3}}) C_{M_1,P_3}
\end{aligned}$$

In this set of equations, M_1 denotes the mRNA species of the X-tra protein P_1. The capacity monitor mRNA is represented by M_2 and its protein by P_2. The RNA-binding protein L7Ae is denote by P_3 and its mRNA species is given by M_3.

$$\begin{aligned}
P_1^* &= \frac{\alpha_{P_1} \frac{\beta_{M_1}\gamma_{M_1}(1+\nu)}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\nu+\rho_{M_1} P_3^*)}}{1 + \frac{\beta_{M_1}\gamma_{M_1}(1+\nu)}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\nu+\rho_{M_1} P_3^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}} + \frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}} \\
P_2^* &= \frac{\alpha_{P_2} \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}}{1 + \frac{\beta_{M_1}\gamma_{M_1}(1+\nu)}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\nu+\rho_{M_1} P_3^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}} + \frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}} \\
P_3^* &= \frac{\alpha_{P_3} \frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}}{1 + \frac{\beta_{M_1}\gamma_{M_1}(1+\nu)}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\nu+\rho_{M_1} P_3^*)} + \frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}} + \frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}}
\end{aligned}$$

Here, we introduce a new lumped parameter $\rho_{M_1} := \frac{\eta}{\delta_{M_1}}$. Due to the complexity of the expressions, the system has not been fully solved for its steady state. To obtain the expressions used for fitting to the data, the steady state expression for P_3^* above was solved for P_3^* and plugged into the steady state expressions of P_1^* and P_2^*.

Ms2-cNOT7:

The system of ordinary differential equations used to obtain the steady state expressions is:

$$\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_2, G_3)G_1 - (\delta_{M_1} + \eta P_3)M_1 \\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_2, G_3)G_2 - \delta_{M_2}M_2 \\
\dot{M}_3 &= k_{M_3}^{eff}(G_1, G_2, G_3)G_3 - \delta_{M_3}M_3 \\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2, M_3)M_1 - \delta_{P_1}P_1 \\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2, M_3)M_2 - \delta_{P_2}P_2 \\
\dot{P}_3 &= k_{P_3}^{eff}(M_1, M_2, M_3)M_3 - \delta_{P_3}P_3
\end{aligned}$$

Again, M_1 represents the mRNA and P_1 the protein of X-tra. M_2 and P_2 denote the mRNA and the protein of the capacity monitor. The RNA-binding protein Ms2-cNOT7 is captured in P_3, with the mRNA species M_3. Solving the above system for steady state yields.

$$\begin{aligned}
P_1^* &= \frac{\alpha_{P_1}\frac{\beta_{M_1}\gamma_{M_1}}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\rho_{M_1}P_3^*)}}{1+\frac{\beta_{M_1}\gamma_{M_1}}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\rho_{M_1}P_3^*)}+\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}} \\
P_2^* &= \frac{\alpha_{P_2}\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}}{1+\frac{\beta_{M_1}\gamma_{M_1}}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\rho_{M_1}P_3^*)}+\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}} \\
P_3^* &= \frac{\alpha_{P_3}\frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}}{1+\frac{\beta_{M_1}\gamma_{M_1}}{(1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3})(1+\rho_{M_1}P_3^*)}+\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}+\frac{\beta_{M_3}\gamma_{M_3}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_{M_3}}}
\end{aligned}$$

As before, the expressions fit to the data are obtained by solving the equation for P_3^* and plugin the result into the other two expressions.

miRNA:

The system of ordinary differential equations used to obtain the steady state expressions is:

$$\begin{aligned}
\dot{M}_1 &= k_{M_1}^{eff}(G_1, G_1, G_m)G_1 - (\delta_{M_1} + \eta m)M_1 \\
\dot{M}_2 &= k_{M_2}^{eff}(G_1, G_1, G_m)G_1 - \delta_{M_2}M_2 \\
\dot{m} &= k_m^{eff}(G_1, G_1, G_m)G_m - \delta_m m \\
\dot{P}_1 &= k_{P_1}^{eff}(M_1, M_2)M_1 - \delta_{P_1}P_1 \\
\dot{P}_2 &= k_{P_2}^{eff}(M_1, M_2)M_2 - \delta_{P_2}P_2
\end{aligned}$$

In this system of equations, we denote the miRNA sensor mRNA by M_1 and the respective fluorescent protein output as P_1. The capacity monitor is expressed through the mRNA M_2 and the protein P_2. The miRNA itself is represented by m. Solving for steady state yields:

$$\begin{aligned}
P_1^* &= \frac{\alpha_{P_1}\frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_m(1+\theta_{M_1}\lambda\tau)}}{1+\frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_m}+\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_m(1+\theta_{M_1}\lambda\tau)}} \\
P_2^* &= \frac{\alpha_{P_2}\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_m}}{1+\frac{\beta_{M_1}\gamma_{M_1}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_m}+\frac{\beta_{M_2}\gamma_{M_2}}{1+\gamma_{M_1}+\gamma_{M_2}+\gamma_m(1+\theta_{M_1}\lambda\tau)}}
\end{aligned}$$

We use the additional lumping parameter:

$$\theta_{M_i} := \frac{\eta_{M_i}}{\delta_{M_i}} \frac{k_{M_2}^{cat} R_M^{total}}{\delta_m}$$

To account for the placement of the miRNA targets in different UTRs and different numbers of targets we introduce the two parameters λ (UTR) and τ (target number). Specifically, $\lambda \in \{\lambda_{3'}, \lambda_{5'}\}$ and $\tau \in \{\tau_{1x}, \tau_{3x}\}$.

A.7 Supplementary Note 7

Use of geometric means as a relative measure of cellular capacity.

To show how the ratio of geometric means measures gene expression capacity we consider a simple model of gene expression:

$$\begin{aligned}
\dot{C}_{M_i} &= k_{M_i}^{+} G_i R_{f_M} - (k_{M_i}^{-} + k_{M_i}^{cat}) C_{M_i} \\
\dot{M}_i &= k_{M_i}^{cat} C_{M_i} - \delta_{M_i} M_i \\
\dot{C}_{P_i} &= k_{P_i}^{+} M_i R_{f_P} - (k_{P_i}^{-} + k_{P_i}^{cat}) C_{P_i} \\
\dot{P}_i &= k_{P_i}^{cat} C_{P_i} - \delta_{P_i} P_i
\end{aligned}$$

In this model, gene G_i can bind free transcriptional resources R_{f_M}. This forms the complex C_{M_i} with rate $k_{M_i}^{+}$. This complex can disassociate with rate $k_{M_i}^{-}$ or give rise to the mRNA species M_i with rate k_i^{cat}. The mRNA species M_i may also degrade with rate δ_{M_i}. Analogously to this process, the mRNA species can bind free translational resources R_{f_P} to form the complex C_{P_i} with rate $k_{P_i}^{+}$. Again, this complex may dissociate with rate $k_{P_i}^{-}$ or produce the protein species P_i. This protein species is removed with rate δ_{P_i}. When the left-hand side of the equations above are set to zero, we can solve for the steady-state expression for the protein species:

$$P_i^* = \alpha_i G_i R_{f_M}^* R_{f_P}^* \text{ with } \alpha_i := \frac{k_{P_i}^{cat}}{\delta_{P_i}} \frac{k_{P_i}^{+}}{k_{P_i}^{-} + k_{P_i}^{cat}} \frac{k_{M_i}^{+}}{k_{M_i}^{-} + k_{M_i}^{cat}}$$

By treating G_i, $R_{f_M}^*$, $R_{f_P}^*$ and P_i^* as random variables we can write the geometric mean of P_i^* as:

$$\begin{aligned}
e^{\mathrm{E}[\log(P_i^*)]} &= e^{\mathrm{E}[\log(\alpha_i G_i R_{f_M}^* R_{f_P}^*)]} \\
&= e^{\mathrm{E}[\log(\alpha_i)+\log(G_i)+\log(R_{f_M}^*)+\log(R_{f_P}^*)]} \\
&= e^{\log(\alpha_i)+\mathrm{E}[\log(G_i)]+\mathrm{E}[\log(R_{f_M}^*)]+\mathrm{E}[\log(R_{f_P}^*)]} \\
&= \alpha_i e^{\mathrm{E}[\log(G_i)]} e^{\mathrm{E}[\log(R_{f_M}^*)]} e^{\mathrm{E}[\log(R_{f_P}^*)]}
\end{aligned}$$

If we consider P_i^* as our capacity monitor, which we use to measure the change in gene expression capacity relative to the capacity monitor measures for a baseline capacity$P_{i,0}^*$, then we write:

$$\frac{e^{\mathrm{E}[\log(P_i^*)]}}{e^{\mathrm{E}[\log(P_{i,0}^*)]}} = \frac{\alpha_i e^{\mathrm{E}[\log(G_i)]} e^{\mathrm{E}[\log(R_{f_M}^*)]} e^{\mathrm{E}[\log(R_{f_P}^*)]}}{\alpha_i e^{\mathrm{E}[\log(G_i)]} e^{\mathrm{E}[\log(R_{f_M,0}^*)]} e^{\mathrm{E}[\log(R_{f_P,0}^*)]}}$$

Assuming that the plasmid take-up distribution of the capacity monitor plasmid G_i does not change across samples we get:

$$\begin{aligned}
\frac{e^{\mathrm{E}[\log(P_i^*)]}}{e^{\mathrm{E}[\log(P_{i,0}^*)]}} &= \frac{e^{\mathrm{E}[\log(R_{f_M}^*)]} e^{\mathrm{E}[\log(R_{f_P}^*)]}}{e^{\mathrm{E}[\log(R_{f_M,0}^*)]} e^{\mathrm{E}[\log(R_{f_P,0}^*)]}} \\
&= \frac{e^{\mathrm{E}[\log(R_{f_M}^*)]+\mathrm{E}[\log(R_{f_P}^*)]}}{e^{\mathrm{E}[\log(R_{f_M,0}^*)]+\mathrm{E}[\log(R_{f_P,0}^*)]}} \\
&= \frac{e^{\mathrm{E}[\log(R_{f_M}^*)+\log(R_{f_P}^*)]}}{e^{\mathrm{E}[\log(R_{f_M,0}^*)+\log(R_{f_P,0}^*)]}} \\
&= \frac{e^{\mathrm{E}[\log(R_{f_M}^* R_{f_P}^*)]}}{e^{\mathrm{E}[\log(R_{f_M,0}^* R_{f_P,0}^*)]}}
\end{aligned}$$

$$
\begin{aligned}
e^{\mathrm{E}[\log(P_i^*)]-\mathrm{E}[\log(P_{i,0}^*)]} &= e^{\mathrm{E}[\log(R^*_{f_M}R^*_{f_P})]-\mathrm{E}[\log(R^*_{f_M,0}R^*_{f_P,0})]} \\
e^{\mathrm{E}[\log(P_i^*)-\log(P_{i,0}^*)]} &= e^{\mathrm{E}[\log(R^*_{f_M}R^*_{f_P})-\log(R^*_{f_M,0}R^*_{f_P,0})]} \\
e^{\mathrm{E}[\log(\frac{P_i^*}{P_{i,0}^*})]} &= e^{\mathrm{E}[\log(\frac{R^*_{f_M}R^*_{f_P}}{R^*_{f_M,0}R^*_{f_P,0}})]}
\end{aligned}
$$

This shows that our approach can report on relative changes in available free resources. In comparison, when the same analysis is performed with the same assumptions but the arithmetic mean is used instead of the geometric mean we get:

$$
\begin{aligned}
\frac{\mathrm{E}[P_i^*]}{\mathrm{E}[P_{i,0}^*]} &= \frac{\mathrm{E}[\alpha_i G_i R^*_{f_M}R^*_{f_P}]}{\mathrm{E}[\alpha_i G_i R^*_{f_M,0}R^*_{f_P,0}]} \\
&= \frac{\mathrm{E}[G_i R^*_{f_M}R^*_{f_P}]}{\mathrm{E}[G_i R^*_{f_M,0}R^*_{f_P,0}]}
\end{aligned}
$$

This expression is analogous to the ratio of the weighted arithmetic mean, where the weighting is given by G_i. If we consider the gene of the capacity monitor to be integrated, the expression simplifies to:

$$
\frac{\mathrm{E}[P_i^*]}{\mathrm{E}[P_{i,0}^*]} = \frac{\mathrm{E}[R^*_{f_M}R^*_{f_P}]}{\mathrm{E}[R^*_{f_M,0}R^*_{f_P,0}]}
$$

This holds because the gene copy number G_i is fixed after integration and can be pulled out of the expected value. The expression shows that integration of the capacity monitor into the genome of a cell would permit a more direct way of measuring the cellular capacity.

Table A.1: Transfection table for Figure 2.2a.

500 ng total	pGLM49	pTTF72	pGLM171
1:1	62.5 ng	62.5 ng	375 ng
1:2	62.5 ng	125 ng	312.5 ng
1:3	62.5 ng	187.5 ng	250 ng
1:4	62.5 ng	250 ng	187.5 ng
2:1	125 ng	62.5 ng	312.5 ng
2:2	125 ng	125 ng	250 ng
2:3	125 ng	187.5 ng	187.5 ng
2:4	125 ng	250 ng	125 ng
3:1	187.5 ng	62.5 ng	250 ng
3:2	187.5 ng	125 ng	187.5 ng
3:3	187.5 ng	187.5 ng	125 ng
3:4	187.5 ng	250 ng	62.5 ng
4:1	250 ng	62.5 ng	187.5 ng
4:2	250 ng	125 ng	125 ng
4:3	250 ng	187.5 ng	62.5 ng
4:4	250 ng	250 ng	0 ng
Reagent/cells			
Optimem	to 50 µL		
PEI	1.5 µL		
HEK293T	62500		

50 ng total	pGLM49	pTTF72	pGLM171
1:1	6.25 ng	6.25 ng	487.5 ng
1:2	6.25 ng	12.5 ng	481.25 ng
1:3	6.25 ng	18.75 ng	475 ng
1:4	62.5 ng	25 ng	468.75 ng
2:1	12.5 ng	6.25 ng	481.25 ng
2:2	12.5 ng	12.5 ng	475 ng
2:3	12.5 ng	18.75 ng	468.75 ng
2:4	12.5 ng	25 ng	462.5 ng
3:1	18.75 ng	6.25 ng	475 ng
3:2	18.75 ng	12.5 ng	468.75 ng
3:3	18.75 ng	18.75 ng	462.5 ng
3:4	18.75 ng	25 ng	456.25 ng
4:1	25 ng	6.25 ng	468.75 ng
4:2	25 ng	12.5 ng	462.5 ng
4:3	25 ng	18.75 ng	456.25 ng
4:4	25 ng	25 ng	450 ng
Reagent/cells			
Optimem	to 50 µL		
PEI	1.5 µL		
HEK293T	62500		

Table A.2: Transfection table for Supplementary Figure A.2.

	pTTF57	pTTF181	pGLM49	pTTF72	pGLM171
EFS/EFS	174.9 ng	174.8 ng	0 ng	0 ng	96.8 ng
EF1a/EFS	0 ng	174.8 ng	215.0 ng	0 ng	56.7 ng
EFS/EF1a	174.9 ng	0 ng	0 ng	215.5 ng	56.0 ng
EF1a/EF1a	0 ng	0 ng	215.0 ng	215.5 ng	15.9 ng
Reagent/cells					
Optimem	to 50 µL				
PEI	1.5 µL				
HEK293T	50000/75000				

Table A.3: Transfection table for Supplementary Figure A.3.

	pTTF57	pTTF181	pGLM49	pTTF72	pGLM171
EFS/EFS	174.9 ng	174.8 ng	0 ng	0 ng	96.8 ng
EF1a/EFS	0 ng	174.8 ng	215.0 ng	0 ng	56.7 ng
EFS/EF1a	174.9 ng	0 ng	0 ng	215.5 ng	56.0 ng
EF1a/EF1a	0 ng	0 ng	215.0 ng	215.5 ng	15.9 ng
Reagent/cells					
Optimem	to 50 µL				
PEI	1.5 µL				
HEK293T	50000/75000				

Table A.4: Transfection table for Figure 2.2b.

	pGLM49	pTTF72	pGLM171
1:1	62.5 ng	62.5 ng	375 ng
Reagent/cells			
Optimem	to 50 µL		
PEI	1.5 µL		
HEK293T	62500		

Table A.5: Transfection table for Supplementary Figure A.4.

	pL-A1	pBI-G	pEMPTY
1.0	120 ng	120 ng	360 ng
1.5	120 ng	180 ng	200 ng
2.0	120 ng	240 ng	140 ng
2.5	120 ng	300 ng	80 ng
Reagent/cells			
Optimem	50 µL		
Lipofectemine 3000	0.75 µL		
P3000	1 µL		
H1299	150000		
U2OS	200000		
HeLa	200000		
HEK293T	150000		
CHO-K1	150000		

Table A.6: Transfection table for Figure 2.2c and supplementary figure A.1b-c, A.8a, A.9a, A.10a.

	pL-A1	pBI-G	pEMPTY
1.0	120 ng	120 ng	360 ng
1.5	120 ng	180 ng	200 ng
2.0	120 ng	240 ng	140 ng
2.5	120 ng	300 ng	80 ng
Reagent/cells			
Optimem	50 µL		
Lipofectemine 3000	0.75 µL		
P3000	1 µL		
H1299	150000		
U2OS	200000		
HeLa	200000		
HEK293T	150000		
CHO-K1	150000		

Table A.7: Transfection table for Supplementary Figure A.1d-e.

	ai274	pBI-G	pEMPTY
1.0	120 ng	120 ng	360 ng
1.5	120 ng	180 ng	200 ng
2.0	120 ng	240 ng	140 ng
2.5	120 ng	300 ng	80 ng
Reagent/cells			
Optimem	50 µL		
Lipofectemine 3000	0.75 µL		
P3000	1 µL		
H1299	150000		
HEK	150000		

Table A.8: Transfection table for Figure 2.2d and Supplementary Figure A.5, A.23, A.24.

	pBI-F3G	pBI-H3G
noTS	500 ng	
miR31 iFFL		500 ng
Reagent/cells		
Optimem	250 µL	
Lipofectemine 3000	11 µL	
P3000	8.25 µL	
H1299	1650000	

Table A.9: Transfection table for Figure 2.2e and Supplementary Figure A.7.

	pTTF218	pTTF219
HDV (-)	0 ng	500 ng
HDV (+)	500 ng	0 ng
Reagent/cells		
Optimem	to 50 µL	
PEI	1.5 µL	
HEK293T	70000	

Table A.10: Transfection table for Figure 2.2f and Supplementary Figure A.8b, A.9b, A.10b, A.11b.

	pL-A1	Oipron	pBI-G	pEMPTY
(-) synthetic intron	50 ng		50 ng	200 ng
(+) synthtic intron		50 ng	50 ng	200 ng
Reagent/cells				
Optimem	50 µL			
Lipofectemine 3000	0.75 µL			
P3000	1 µL			
H1299	150000			
HEK293T	150000			
U2OS	200000			
HeLa	200000			
CHO-K1	150000			

Table A.11: Transfection table for Figure 2.2g and Supplementary Figure A.8c, A.9c, A.10c, A.11c.

	pL-S1	pL-C1	p125	pL-R1	pL-A1	pEMPTY
L7Ae control	50 ng				50 ng	200 ng
L7Ae	50 ng		50 ng		50 ng	150 ng
Ms2-cNOT7 control		50 ng			50 ng	200 ng
Ms2-cNOT7		50 ng		50 ng	50 ng	150 ng
Reagent/cells						
Optimem	50 µL					
Lipofectemine 3000	0.75 µL					
P3000	1 µL					
H1299	150000					
HEK	150000					
U2OS	200000					
HeLa	200000					
CHO-K1	150000					

Table A.12: Transfection table for Figure 2.2h and Supplementary Figure A.8d, A.9d, A.10d, A.11d.

	pL-A1	pH-7	pH-22	pH-14	pL-S1	pEMPTY
H1299 control	50 ng				50 ng	200 ng
H1299 reppressed		50 ng			50 ng	200 ng
U2OS control	50 ng				50 ng	200 ng
U2OS repressed			50 ng		50 ng	200 ng
HeLa control	50 ng				50 ng	200 ng
HeLa repressed				50 ng	50 ng	200 ng
Reagent/cells						
Optimem	50 µL					
Lipofectemine 2000	1.5 µL					
P3000	1 µL					
H1299	150000					
HeLa	200000					
U2-OS	200000					

Table A.13: Transfection table for Figure 2.3b,c and Supplementary Figure A.12.

	pBI-F3G	pBI-H1G/pBI-H3G/pBI-H5G/pBI-H7G	pEMPTY
Control	100 ng		200 ng
miR-31 TS		100 ng	200 ng
miR-31 TS + inhibitor (20pmol)		100 ng	200 ng
Reagent/cells			
Optimem	50 µL		
Lipofectemine 3000	0.75 µL		
P3000	1 µL		
H1299	150000		

Table A.14: Transfection table for Supplementary Figure A.13.

	pL-A1	pH-1/pH-2/pH-3/pH-5/pH-6/pH-7	pL-S1	pEMPTY
noTS	50 ng		50 ng	200 ng
miR-31 TS		50 ng	50 ng	200 ng
Reagent/cells				
Optimem	50 µL			
Lipofectemine 3000	0.75 µL			
P3000	1 µL			
H1299	150000			

Table A.15: Transfection table for Supplementary Figure A.14.

	pL-A1	pH-31	pL-S1	pEMPTY
noTS	50 ng		50 ng	200 ng
miR-31 TS		50 ng	50 ng	200 ng
Reagent/cells				
Optimem	50 µL			
Lipofectemine 3000	0.75 µL			
P3000	1 µL			
HEK293T	150000			
U2-OS	200000			

Table A.16: Transfection table for Supplementary Figure A.15, A.16.

SF 16	pL-A1	pH-16/pH-17/pH-18/pH-20/pH-21/pH-22	pL-S1	pEMPTY
noTS	50 ng		50 ng	200 ng
miR-221 TS		50 ng	50 ng	200 ng
SF 17	pL-A1	pH-8/pH-9/pH-10/pH-12/pH-13/pH-14	pL-S1	pEMPTY
noTS	50 ng		50 ng	200 ng
miR-21 TS		50 ng	50 ng	200 ng
Reagent/cells				
Optimem	50 µL			
Lipofectemine 3000	0.75 µL			
HeLa	200000			
U2-OS	200000			

Table A.17: Transfection table for Figure 2.4.

	pGLM49	pGLM91	pGLM92	pGLM102	pGLM103	pGLM171
OLP 0 ng	140 ng	0 ng	180 ng	0 ng	0 ng	180 ng
OLP 20 ng	140 ng	0 ng	180 ng	0 ng	20 ng	140 ng
OLP 60 ng	140 ng	0 ng	180 ng	0 ng	60 ng	100 ng
OLP 100 ng	140 ng	0 ng	180 ng	0 ng	100 ng	60 ng
OLP 140 ng	140 ng	0 ng	180 ng	0 ng	140 ng	20 ng
OLP 180 ng	140 ng	0 ng	180 ng	0 ng	180 ng	0 ng
IFF 0 ng	140 ng	180 ng	0 ng	0 ng	0 ng	180 ng
IFF 20 ng	140 ng	180 ng	0 ng	0 ng	20 ng	140 ng
IFF 60 ng	140 ng	180 ng	0 ng	0 ng	60 ng	100 ng
IFF 100 ng	140 ng	180 ng	0 ng	0 ng	100 ng	60 ng
IFF 140 ng	140 ng	180 ng	0 ng	0 ng	140 ng	20 ng
IFF 180 ng	140 ng	180 ng	0 ng	0 ng	180 ng	0 ng
FBK 0 ng	140 ng	0 ng	180 ng	0 ng	0 ng	180 ng
FBK 20 ng	140 ng	0 ng	180 ng	20 ng	0 ng	140 ng
FBK 60 ng	140 ng	0 ng	180 ng	60 ng	0 ng	100 ng
FBK 100 ng	140 ng	0 ng	180 ng	100 ng	0 ng	60 ng
FBK 140 ng	140 ng	0 ng	180 ng	140 ng	0 ng	20 ng
FBK 180 ng	140 ng	0 ng	180 ng	180 ng	0 ng	0 ng
HYB 0 ng	140 ng	180 ng	0 ng	0 ng	0 ng	180 ng
HYB 20 ng	140 ng	180 ng	0 ng	20 ng	0 ng	140 ng
HYB 60 ng	140 ng	180 ng	0 ng	60 ng	0 ng	100 ng
HYB 100 ng	140 ng	180 ng	0 ng	100 ng	0 ng	60 ng
HYB 140 ng	140 ng	180 ng	0 ng	140 ng	0 ng	20 ng
HYB 180 ng	140 ng	180 ng	0 ng	180 ng	0 ng	0 ng
Reagent/cells						
Optimem	to 50 µL					
PEI	1.5 µL					
HEK293T	62500					

Table A.18: Transfection table for Supplementary Figure A.21.

	pL-A1	p-H3/p-H18	pBI-G	pEMPTY
0.25 noTS	120 ng		30 ng	150 ng
0.25 miR-TS		120 ng	30 ng	150 ng
0.50 noTS	120 ng		60 ng	120 ng
0.50 miR-TS		120 ng	60 ng	120 ng
0.75 noTS	120 ng		90 ng	90 ng
0.75 miR-TS		120 ng	90 ng	90 ng
1.00 noTS	120 ng		120 ng	60 ng
1.00 miR-TS		120 ng	120 ng	60 ng
1.25 noTS	120 ng		150 ng	30 ng
1.25 miR-TS		120 ng	150 ng	30 ng
1.50 noTS	120 ng		180 ng	
1.50 miR-TS		120 ng	180 ng	
Reagent/cells				
Optimem	50 µL			
Lipofectemine 3000	0.75 µL			
P3000	1 µL			
H1299	150000			
U2OS	200000			
HEK293T	150000			

Table A.19: Transfection table for Figure 2.5c and Supplementary Figure A.18, A.19, A.20, A.22.

	pL-A1	p-H3/p-H18	pBI-G	pEMPTY
0.25 noTS	120 ng		30 ng	150 ng
0.25 miR-TS		120 ng	30 ng	150 ng
0.50 noTS	120 ng		60 ng	120 ng
0.50 miR-TS		120 ng	60 ng	120 ng
0.75 noTS	120 ng		90 ng	90 ng
0.75 miR-TS		120 ng	90 ng	90 ng
1.00 noTS	120 ng		120 ng	60 ng
1.00 miR-TS		120 ng	120 ng	60 ng
1.25 noTS	120 ng		150 ng	30 ng
1.25 miR-TS		120 ng	150 ng	30 ng
1.50 noTS	120 ng		180 ng	
1.50 miR-TS		120 ng	180 ng	
Reagent/cells				
Optimem	50 µL			
Lipofectemine 3000	0.75 µL			
P3000	1 µL			
H1299	150000			
U2OS	200000			
HEK293T	150000			

Table A.20: Transfection table for Supplementary Figure A.21.

	pL-A1	pH-18	pL-S1	pEMPTY
noTS	50 ng		50 ng	200 ng
miR-221 TS		50 ng	50 ng	200 ng
Reagent/cells				
Optimem	50 µL			
Lipofctemine 3000	0.75 µL			
P3000	1 µL			
HEK	150000			

Table A.21: Transfection table for Figure 2.5d.

	pTTF220	pTTF223	pTTF138	pGLM171
0.25 mCit.-3xTFF5/mRub.-3xTFF5	155.1 ng	0 ng	24.2 ng	120.7 ng
0.50 mCit.-3xTFF5/mRub.-3xTFF5	155.1 ng	0 ng	48.3 ng	96.6 ng
0.75 mCit.-3xTFF5/mRub.-3xTFF5	155.1 ng	0 ng	72.5 ng	72.4 ng
1.00 mCit.-3xTFF5/mRub.-3xTFF5	155.1 ng	0 ng	96.6 ng	48.3 ng
1.25 mCit.-3xTFF5/mRub.-3xTFF5	155.1 ng	0 ng	120.8 ng	24.1 ng
1.50 mCit.-3xTFF5/mRub.-3xTFF5	155.1 ng	0 ng	144.9 ng	0 ng
0.25 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	155.1 ng	24.2 ng	120.7 ng
0.50 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	155.1 ng	48.3 ng	96.6 ng
0.75 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	155.1 ng	72.5 ng	72.4 ng
1.00 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	155.1 ng	96.6 ng	48.3 ng
1.25 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	155.1 ng	120.8 ng	24.1 ng
1.50 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	155.1 ng	144.9 ng	0 ng
Reagent/cells				
Optimem	to 50 µL			
Lipofectamine 2000	0.6 µL			
mES E14	70000			

Table A.22: Transfection table for Supplementary Figure A.25.

	pTTF220	pTTF223	pTTF138	pGLM171
0.25 mCit.-3xTFF5/mRub.-3xTFF5	258.4 ng	0 ng	40.3 ng	201.3 ng
0.50 mCit.-3xTFF5/mRub.-3xTFF5	258.4 ng	0 ng	80.5 ng	161.1 ng
0.75 mCit.-3xTFF5/mRub.-3xTFF5	258.4 ng	0 ng	120.8 ng	120.8 ng
1.00 mCit.-3xTFF5/mRub.-3xTFF5	258.4 ng	0 ng	161.0 ng	80.6 ng
1.25 mCit.-3xTFF5/mRub.-3xTFF5	258.4 ng	0 ng	201.3 ng	40.3 ng
1.50 mCit.-3xTFF5/mRub.-3xTFF5	258.4 ng	0 ng	241.6 ng	0 ng
0.25 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	258.4 ng	40.3 ng	201.3 ng
0.50 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	258.4 ng	80.5 ng	161.1 ng
0.75 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	258.4 ng	120.8 ng	120.8 ng
1.00 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	258.4 ng	161.0 ng	80.6 ng
1.25 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	258.4 ng	201.3 ng	40.3 ng
1.50 mCit.-3xTFF4/mRub.-3xTFF4	0 ng	258.4 ng	241.6 ng	0 ng
Reagent/cells				
Optimem	to 50 µL			
PEI	1.5 µL			
HEK293T	70000			

Table A.23: List of the plasmids used in this study. Plasmids sequences are available on GenBank.

Fig.	Short plasmid name	Full plasmid name	Parts from	GenBank accession code
2c-e-f-g, 5b & SF1, 8-11, 13-16, 18-21	pL-A1	pT-GTW6-CMV-mKate	(Siciliano et al., 2018)	MT891367
SF1	ai274	pT-PGK-mCherry	(Matsumiya et al., 2018)	MT891341
2f & SF8-11	Oipron	pT-GTW6-CMV-SImKate	(Lu et al., 2013)	
2c-e, 5c & SF1, 8-11, 18-19	pBI-G	pBI-CMV1_EGFP	Clontech 631630	MT891343
2g-h & SF8-11, 13-16, 20	pL-S1	pBoxCDGC_2xKMet_EGFP	(Wroblewska et al., 2015)	MT891368
2g & SF8-11	pL-C1	pBoxCDGCmut_KMetEGFP-8xMS2-pA	(Wroblewska et al., 2015)	MH883358
2g & SF8-11	p125	pT-GTW6-CMV-L7AescFv35	(Wroblewska et al., 2015)	MH883336
2g & SF8-11	pL-R1	pT-GTW6-CMV-MS2- CNOT7	(Wroblewska et al., 2015)	MH883359
SF13	pH-1	pT-GTW6-CMV-mKate_1xmiR31TS5’	(Cella et al., 2018)	MT891348
SF13	pH-2	pT-GTW6-CMV-mKate_2xmiR31TS5’	(Cella et al., 2018)	MT891349
5c & SF13	pH-3	pT-GTW6-CMV-mKate_3xmiR31TS5’	(Cella et al., 2018)	MT891350
SF13	pH-5	pT-GTW6-CMV-mKate_1xmiR31TS3’	(Cella et al., 2018)	MT891351
SF13	pH-6	pT-GTW6-CMV-mKate_2xmiR31TS3’	(Cella et al., 2018)	MT891352
SF8, 13-14	pH-7	pT-GTW6-CMV-mKate_3xmiR31TS3’	(Cella et al., 2018)	MT891353
2d, 3 & SF5, 12, 23-24	pBI-F3G	pBI-CMV1_EGFP_mKate	Clontech 631630	MT891342
3 & SF12	pBI-H1G	pBI-CMV1_EGFP_mKate_1xmiR31TS5’	Clontech 631630	MT891344
3 & SF5, 12, 23-24	pBI-H3G	pBI-CMV1_EGFP_mKate_3xmiR31TS5’	Clontech 631630	MT891345
3 & SF12	pBI-H5G	pBI-CMV1_EGFP_mKate_1xmiR31TS3’	Clontech 631630	MT891346
3	pBI-H7G	pBI-CMV1_EGFP_mKate_3xmiR31TS3’	Clontech 631630	MT891347
SF15	pH-8	pT-GTW6-CMV-mKate_1xmiR21TS5’	(Cella et al., 2018)	MT891354
SF15	pH-9	pT-GTW6-CMV-mKate_2xmiR21TS5’	(Cella et al., 2018)	MT891355
SF15	pH-10	pT-GTW6-CMV-mKate_3xmiR21TS5’	(Cella et al., 2018)	MT891356
SF15	pH-12	pT-GTW6-CMV-mKate_1xmiR21TS3’	(Cella et al., 2018)	MT891357
SF15	pH-13	pT-GTW6-CMV-mKate_2xmiR21TS3’	(Cella et al., 2018)	MT891358
SF10-11, 15	pH-14	pT-GTW6-CMV-mKate_3xmiR21TS3’	(Cella et al., 2018)	MT891359
SF16	pH-16	pT-GTW6-CMV-mKate_1xmiR221TS5’	(Cella et al., 2018)	MT891361
SF16	pH-17	pT-GTW6-CMV-mKate_2xmiR221TS5’	(Cella et al., 2018)	MT891362
SF16, 20-22	pH-18	pT-GTW6-CMV-mKate_3xmiR221TS5’	(Cella et al., 2018)	MT891363
SF16	pH-20	pT-GTW6-CMV-mKate_1xmiR221TS3’	(Cella et al., 2018)	MT891364
SF16	pH-21	pT-GTW6-CMV-mKate_2xmiR221TS3’	(Cella et al., 2018)	MT891365
2h & SF16	pH-22	pT-GTW6-CMV-mKate_3xmiR221TS3’	(Cella et al., 2018)	MT891366
2a, 4, SF2 & SF3	pGLM49	INS-bGHpA-$P_{EF1\alpha}$-mCitrine-SV40pA-INS		MT891334
SF3	pGLM177	P_{SV40}-PuroR-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-sigW-SV40pA-INS	(Lillacci et al., 2018)	MT891340
2a, SF2 & SF4	pTT72	INS-bGHpA-$P_{EF1\alpha}$-mRuby3-SV40pA-INS	(Lillacci et al., 2018)	MT891324
2a, 4, 5e, SF3, SF25 & SF27	pGLM171	AmpR-INS-bGHpA-SV40pA-INS-pUCori	(Lillacci et al., 2018)	MT891339
2b	pTTF84	INS-bGHpA-synpA-$P_{EF1\alpha}$-tTA2::Cerulean-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-mCitrine-SV40pA	(Lillacci et al., 2018)	MT891325
2b	pTTF145	INS-bGHpA-mRuby3-PbiTRE-miRFP670-SV40pA-INS	(Lillacci et al., 2018)	MT891327
2e & SF7	pTTF218	P_{TRE}-(HDV)mCitrine-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-tTA2::P2A::mRuby3-SV40pA-P_{SV40}-puΔtk-SV40pA	(Lillacci et al., 2018)	MT891330
2e & SF7	pTTF219	P_{TRE}-(dHDV)mCitrine-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-tTA2::P2A::mRuby3-SV40pA-P_{SV40}-puΔtk-SV40pA	(Lillacci et al., 2018)	MT891331
SF4	pTTF57	INS-bGHpA-PEFS-mCitrine-SV40pA-INS	(Lillacci et al., 2018)	MT891323
SF4	pTTF181	INS-bGHpA-PEFS-mRuby3-SV40pA-INS	(Lillacci et al., 2018)	MT891328
4	pGLM91	INS-bGHpA-P_{TRE}-DsRed(FF4)-TFF4x3-SV40pA-INS		MT891335
4	pGLM92	INS-bGHpA-P_{TRE}-DsRed(FF4)-TFF5x3-SV40pA-INS		MT891336
4b	pGLM103	P_{SV40}-PuroR-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-tTA::Cerulean-TFF5x3-SV40pA-INS		MT891338
4c	pGLM102	P_{SV40}-PuroR-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-tTA::Cerulean-TFF4x3-SV40pA-INS		MT891337
5e & SF25	pTTF138	INS-bGHpA-$P_{EF1\alpha}$-miRFP670-SV40pA-INS	(Lillacci et al., 2018)	MT891326
5e & SF25	pTTF220	P_{SV40}-PuroR-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-mCitrine-TFF5x3-SV40pA-INS-$P_{EF1\alpha}$-mRuby3(FF4)-TFF5x3-SV40pA-INS	(Lillacci et al., 2018)	MT891332
5e & SF25	pTTF223	P_{SV40}-PuroR-SV40pA-INS-bGHpA-$P_{EF1\alpha}$-mCitrine-TFF4x3-SV40pA-INS-$P_{EF1\alpha}$-mRuby3(FF4)-TFF4x3-SV40pA-INS	(Lillacci et al., 2018)	MT891333
SF4	pTTF194	INS-bGHpA-$P_{EF1\alpha}$-mCitrine-SV40pA-INS-P_{SV40}-miRFP670-SV40pA-INS	(Lillacci et al., 2018)	MT891329

Table A.24: List of the primers and oligos used to generate miRNA target sites.

Construct	Primer name	Primer Sequence
mKate_1xmiR31TS5’	Fw1	GATCCAGCTATGCCAGCATCTTGCCTG
mKate_1xmiR31TS5’	Rv1	CTAGCAGGCAAGATGCTGGCATAGCTG
mKate_2xmiR31TS5’	Fw2	GATCCAGCTATGCCAGCATCTTGCCTAGCTATGCCAGCATCTTGCCTG
mKate_2xmiR31TS5’	Rv2	CTAGCAGGCAAGATGCTGGCATAGCTAGGCAAGATGCTGGCATAGCTG
mKate_3xmiR31TS5’	Fw3	GATCCAGCTATGCCAGCATCTTGCCTAGCTATGCCAGCATCT
mKate_3xmiR31TS5’	Rv3	TAGCTAGGCAAGATGCTGGCATAGCTG
mKate_3xmiR31TS5’	Fw4	TGCCTAGCTATGCCAGCATCTTGCCTG
mKate_3xmiR31TS5’	Rv4	CTAGCAGGCAAGATGCTGGCATAGCTAGGCAAGATGCTGGCA
mKate_1xmiR31TS3’	Fw5	AGCTTAGCTATGCCAGCATCTTGCCTTTAAT
mKate_1xmiR31TS3’	Rv5	TAAAGGCAAGATGCTGGCATAGCTA
mKate_2xmiR31TS3’	Fw6	AGCTTAGCTATGCCAGCATCTTGCCTAGCTATGCCAGCATCTTGCCTTTAAT
mKate_2xmiR31TS3’	Rv6	TAAAGGCAAGATGCTGGCATAGCTAGGCAAGATGCTGGCATAGCTA
mKate_3xmiR31TS3’	Fw7	AGCTTAGCTATGCCAGCATCTTGCCTAGCTATGCCAGCATCT
mKate_3xmiR31TS3’	Rv7	TAGCTAGGCAAGATGCTGGCATAGCTA
mKate_3xmiR31TS3’	Fw8	TGCCTAGCTATGCCAGCATCTTGCCTTTAAT
mKate_3xmiR31TS3’	Rw8	TAAAGGCAAGATGCTGGCATAGCTAGGCAAGATGCTGGCA
mKate_1xmiR21TS5’	Fw9	GATCCTAGCTTATCAGACTGATGTTGAG
mKate_1xmiR21TS5’	Rv9	CTAGCTCAACATCAGTCTGATAAGCTAG
mKate_2xmiR21TS5’	Fw10	GATCCTAGCTTATCAGACTGATGTTGATAGCTTATCAGACTGATGTTGAG
mKate_2xmiR21TS5’	Rv10	CTAGCTCAACATCAGTCTGATAAGCTATCAACATCAGTCTGATAAGCTAG
mKate_3xmiR21TS5’	Fw11	GATCCTAGCTTATCAGACTGATGTTGATAGCTTATCAGACTGAT
mKate_3xmiR21TS5’	Rv11	AGCTATCAACATCAGTCTGATAAGCTAG
mKate_3xmiR21TS5’	Fw12	GTTGATAGCTTATCAGACTGATGTTGAG
mKate_3xmiR21TS5’	Rv12	CTAGCTCAACATCAGTCTGATAAGCTATCAACATCAGTCTGATA
mKate_1xmiR21TS3’	Fw15	AGCTTTAGCTTATCAGACTGATGTTGATTAAT
mKate_1xmiR21TS3’	Rv15	TAATCAACATCAGTCTGATAAGCTAA
mKate_2xmiR21TS3’	Fw16	AGCTTTAGCTTATCAGACTGATGTTGATAGCTTATCAGACTGATGTTGATTAAT
mKate_2xmiR21TS3’	Rv16	TAATCAACATCAGTCTGATAAGCTATCAACATCAGTCTGATAAGCTAA
mKate_3xmiR21TS3’	Fw17	AGCTTTAGCTTATCAGACTGATGTTGATAGCTTATCAGACTGAT
mKate_3xmiR21TS3’	Rv17	AGCTATCAACATCAGTCTGATAAGCTAA
mKate_3xmiR21TS3’	Fw18	GTTGATAGCTTATCAGACTGATGTTGATTAAT
mKate_3xmiR21TS3’	Rv18	TAATCAACATCAGTCTGATAAGCTATCAACATCAGTCTGATA
mKate_1xmiR221TS5’	Fw21	GATCCACCTGGCATACAATGTAGATTTG
mKate_1xmiR221TS5’	Rv21	CTAGCAAATCTACATTGTATGCCAGGTG
mKate_2xmiR221TS5’	Fw22	GATCCACCTGGCATACAATGTAGATTTACCTGGCATACAATGTAGATTTG
mKate_2xmiR221TS5’	Rv22	CTAGCAAATCTACATTGTATGCCAGGTAAATCTACATTGTATGCCAGGTG
mKate_3xmiR221TS5’	Fw23	GATCCACCTGGCATACAATGTAGATTTACCTGGCATACAATGT
mKate_3xmiR221TS5’	Rv23	CAGGTAAATCTACATTGTATGCCAGGTG
mKate_3xmiR221TS5’	Fw24	AGATTTACCTGGCATACAATGTAGATTTG
mKate_3xmiR221TS5’	Rv24	CTAGCAAATCTACATTGTATGCCAGGTAAATCTACATTGTATGC
mKate_1xmiR221TS3’	Fw25	AGCTTACCTGGCATACAATGTAGATTTTTAAT
mKate_1xmiR221TS3’	Rv25	TAAAAATCTACATTGTATGCCAGGTA
mKate_2xmiR221TS3’	Fw26	AGCTTACCTGGCATACAATGTAGATTTACCTGGCATACAATGTAGATTTTTAAT
mKate_2xmiR221TS3’	Rv26	TAAAAATCTACATTGTATGCCAGGTAAATCTACATTGTATGCCAGGTA
mKate_3xmiR221TS3’	Fw27	AGCTTACCTGGCATACAATGTAGATTTACCTGGCATACAATGTA
mKate_3xmiR221TS3’	Rv27	CAGGTAAATCTACATTGTATGCCAGGTA
mKate_3xmiR221TS3’	Fw28	GATTTACCTGGCATACAATGTAGATTTTTAAT
mKate_3xmiR221TS3’	Rv28	TAAAAATCTACATTGTATGCCAGGTAAATCTACATTGTATGC

Table A.25: List of the target sites for miRNAs used in this study (Griffiths-Jones et al., 2006).

miRNA	Sequence
hsa-miR-31-5p	AGCTATGCCAGCATCTTGCCT
hsa-miR-21	TCAACATCAGTCTGATAAGCTA
hsa-miR-221	AAATCTACATTGTATGCCAGGT
miR-FF4	CCGCTTGAAGTCTTTAATTAAA
miR-FF5	AAGCACTCTGATTTGACAATTA

Table A.26: List of the primers used for qPCR analyses.

Primer	**Function**	**Sequence (5'-3')**
F7	Forward primer for mKate	GGTGTCTAAGGGCGAAGAGC
F8	Reverse primer for mKate	GCTGGTAGCCAGGATGTCGA
qPCR-EGFP-F	Forward primer for EGFP	AAGGGCATCGACTTCAAG
qPCR-EGFP-R	Reverse primer for EGFP	TGCTTGTCGGCCATGATATG
qPCR-18S-F	Forward primer for 18S	GCTTAATTTGACTCAACACGGGA
qPCR-18S-R	Reverse primer for 18S	AGCTATCAATCTGTCAATCCTGTC

Table A.27: 5x isothermal reaction buffer recipe.

Component	**Concentration**
PEG-800	25 %
Tris-HCl, pH 7.5	500 mM
MgCl2	50 mM
DTT	50 mM
dATP	1 mM
dTTP	1 mM
dCTP	1 mM
dGTP	1 mM
NAD	5 mM
To 3 mL with ddH2O	

Table A.28: Parameter fits related to Figure 2.4b,c and Supplementary Note A.3.

Parameter	**Unit**	**Value**
α_{P_1}	$Arbitrary fluorescence$	3108.52
α_{P_2}	$Arbitrary fluorescence$	28007.3
α_{P_3}	$Arbitrary fluorescence$	2198.66
β_{M_1}	$Unitless$	9275.79
β_{M_2}	$Unitless$	1017.34
β_{M_3}	$Unitless$	709.432
$k_{m_{M_1}}$	ng	0.00400709
γ_{M_2}	$Unitless$	1.35779
γ_{M_3}	$Unitless$	14.5078
θ_{M_1}	$Unitless$	286.573
θ_{M_2}	$Unitless$	126.876
κ	$Arbitrary fluorescence$	17.2643
h	$Unitless$	31

Table A.29: Parameter fits related to Figure 2.5c and Supplementary Note A.4 endogenous microRNA-based iFFL.

Parameter	Unit	Value
β_{M_1}	*Unitless*	0.746056
β_{M_2}	*Unitless*	32.429
$k_{m_{M_1}}$	*Unitless*	59.104
γ_{M_2}	*Unitless*	0.480552
γ_m	*Unitless*	33.2046
θ_{M_2}	*Unitless*	227.144

Table A.30: Parameter fits related to Figure 2.5d and Supplementary Note A.4 synthetic microRNA-based iFFL.

Parameter	Unit	Value
β_{M_1}	*Unitless*	1.89289
β_{M_2}	*Unitless*	0.0152265
β_{M_3}	*Unitless*	1.43347e-18
$k_{m_{M_1}}$	*Unitless*	3.21455
γ_{M_2}	*Unitless*	2.04539e-16
γ_{M_3}	*Unitless*	1.91369
θ_{M_2}	*Unitless*	9.80712
θ_{M_3}	*Unitless*	30.272

Table A.31: Parameter fits related to Supplementary Figure A.19 and Supplementary Note A.4 endogenous microRNA-based iFFL.

Parameter	Unit	Value
β_{M_1}	*Unitless*	1.01232e-9
β_{M_2}	*Unitless*	0.943493
$k_{m_{M_1}}$	*Unitless*	3.57119
γ_{M_2}	*Unitless*	2.85075e-11
γ_m	*Unitless*	0.173874
θ_{M_2}	*Unitless*	10.9904

Table A.32: Parameter fits related to Supplementary Figure A.21 and Supplementary Note A.4 endogenous microRNA-based iFFL.

Parameter	**Unit**	**Value**
β_{M_1}	*Unitless*	2.81166
β_{M_2}	*Unitless*	1.95439e-5
$k_{m_{M_1}}$	*Unitless*	2.32439
γ_{M_2}	*Unitless*	1.1604e-13
γ_m	*Unitless*	6.65805
θ_{M_2}	*Unitless*	6.38791

Table A.33: Parameter fits related to Supplementary Figure A.22 and Supplementary Fig. A.4 synthetic microRNA-based iFFL.

Parameter	**Unit**	**Value**
β_{M_1}	*Unitless*	9.61413e-16
β_{M_2}	*Unitless*	0.00222615
β_{M_3}	*Unitless*	3.91565e-17
$k_{m_{M_1}}$	*Unitless*	33.6238
γ_{M_2}	*Unitless*	3.19321e-14
γ_{M_3}	*Unitless*	13.4858
θ_{M_2}	*Unitless*	6.45491
θ_{M_3}	*Unitless*	30.2392

Table A.34: Parameter fits related to Supplementary Figure A.17 and Supplementary Note A.6 Doxycycline titration.

Parameter	Unit	Value
α_{P_1}	$Arbitrary fluorescence$	8467.86
α_{P_2}	$Arbitrary fluorescence$	11843.7
α_{P_3}	$Arbitrary fluorescence$	34208.2
α_{P_4}	$Arbitrary fluorescence$	7872.41
β_{M_1}	$Unitless$	36149.8
β_{M_2}	$Unitless$	1763.75
β_{M_3}	$Unitless$	2970.3
β_{M_4}	$Unitless$	9847.66
γ_{M_1}	$Unitless$	276.803
γ_{M_2}	$Unitless$	11620.
γ_{M_3}	$Unitless$	1914.02
γ_{M_4}	$Unitless$	13025.9
κ_a	$Arbitrary fluorescence$	0.000376413
κ_i	$Arbitrary fluorescence$	0.402255
h_a	$Unitless$	1
h_i	$Unitless$	1

Table A.35: Parameter fits related to Supplementary Figure A.17 and Supplementary Note A.6 HDV.

Parameter	Unit	Value
α_{P_1}	$Arbitrary fluorescence$	2349.49
α_{P_2}	$Arbitrary fluorescence$	5194.24
β_{M_1}	$Unitless$	2.17088
β_{M_2}	$Unitless$	0.389635
γ_{M_1}	$Unitless$	269.809
γ_{M_2}	$Unitless$	308.635
ω_{M_1}	$Unitless$	19.1555
κ_a	$Arbitrary fluorescence$	0.00309458
κ_i	$Arbitrary fluorescence$	1.36083
h_a	$Unitless$	2
h_i	$Unitless$	1

Table A.36: Parameter fits related to Supplementary Figure A.17 and Supplementary Note A.6 RNA-binding proteins (L7Ae).

Parameter	**Unit**	**Value**
α_{P_1}	$Arbitrary fluorescence$	41764.7
α_{P_2}	$Arbitrary fluorescence$	2604.06
α_{M_3}	$Arbitrary fluorescence$	16208.
β_{M_1}	$Unitless$	657.816
β_{M_2}	$Unitless$	3508.74
β_{M_3}	$Unitless$	386.641
γ_{M_1}	$Unitless$	142.346
γ_{M_2}	$Unitless$	23.082
γ_{M_3}	$Unitless$	10.0254
ν	$Time^{-1}$	36932.
ρ_{M_1}	$Arbitrary fluorescence^{-1}$	1093.94

Table A.37: Parameter fits related to Supplementary Figure A.17 and Supplementary Note A.6 RNA-binding proteins (Ms2-cNOT7).

Parameter	**Unit**	**Value**
α_{P_1}	$Arbitrary fluorescence$	19444.9
α_{P_2}	$Arbitrary fluorescence$	14480.
α_{M_3}	$Arbitrary fluorescence$	1.77696
β_{M_1}	$Unitless$	2115.61
β_{M_2}	$Unitless$	858.958
β_{M_3}	$Unitless$	74.2967
γ_{M_1}	$Unitless$	41.1893
γ_{M_2}	$Unitless$	104.793
γ_{M_3}	$Unitless$	31.3533
ρ_{M_1}	$Arbitrary fluorescence^{-1}$	213.586

Table A.38: Parameter fits related to Supplementary Figure A.17 and Supplementary Note A.6 miRNA.

Parameter	**Unit**	**Value**
α_{P_1}	$Arbitrary fluorescence$	433.658
α_{P_2}	$Arbitrary fluorescence$	4855.31
β_{M_1}	$Unitless$	992.56
β_{M_2}	$Unitless$	895.094
γ_{M_1}	$Unitless$	2232.66
γ_{M_2}	$Unitless$	518.349
γ_{M_3}	$Unitless$	2.56322
$\lambda_{3'}$	$Unitless$	155.969
$\lambda_{5'}$	$Unitless$	926.904
τ_{1x}	$Unitless$	3.0671
τ_{3x}	$Unitless$	70.1826

A.8 Bibliography

Lillacci, G., Benenson, Y., & Khammash, M. (2018). Synthetic control systems for high performance gene expression in mammalian cells. *Nucleic acids research*, *46*(18), 9855–9863.

Siciliano, V., DiAndreth, B., Monel, B., Beal, J., Huh, J., Clayton, K. L., Wroblewska, L., McKeon, A., Walker, B. D., & Weiss, R. (2018). Engineering modular intracellular protein sensor-actuator devices. *Nature communications*, *9*(1), 1–7.

Matsumiya, M., Tomita, T., Yoshioka-Kobayashi, K., Isomura, A., & Kageyama, R. (2018). Es cell-derived presomitic mesoderm-like tissues for analysis of synchronized oscillations in the segmentation clock. *Development*, *145*(4), dev156836.

Lu, J., Zhang, F., & Kay, M. A. (2013). A mini-intronic plasmid (mip): A novel robust transgene expression vector in vivo and in vitro. *Molecular Therapy*, *21*(5), 954–963.

Wroblewska, L., Kitada, T., Endo, K., Siciliano, V., Stillo, B., Saito, H., & Weiss, R. (2015). Mammalian synthetic circuits with rna binding proteins for rna-only delivery. *Nature biotechnology*, *33*(8), 839–841.

Cella, F., Wroblewska, L., Weiss, R., & Siciliano, V. (2018). Engineering protein-protein devices for multilayered regulation of mrna translation using orthogonal proteases in mammalian cells. *Nature communications*, *9*(1), 1–9.

Griffiths-Jones, S., Grocock, R. J., Van Dongen, S., Bateman, A., & Enright, A. J. (2006). Mirbase: Microrna sequences, targets and gene nomenclature. *Nucleic acids research*, *34*(suppl_1), D140–D144.

Appendix B

A genetic mammalian proportional-integral feedback control circuit for robust and precisegene regulation

B.1 Mathematical Modeling of the Circuit in Figure 3.2 without Network Perturbation

Consider the circuit depicted in Figure 3.2(a) without network perturbation, where an RNA-based molecular realization of the antithetic integral controller regulates the production of a particular protein of interest, namely the transcription factor tTA-mCitrine. The circuit can operate in either open loop or closed loop. We first present a detailed (mechanistic) mathematical model and then carry out a model reduction technique that allows us to analyze the steady-state behavior of the output protein. Finally, we provide the technical details of properly calibrating the model to the experimental data.

Full Model Description

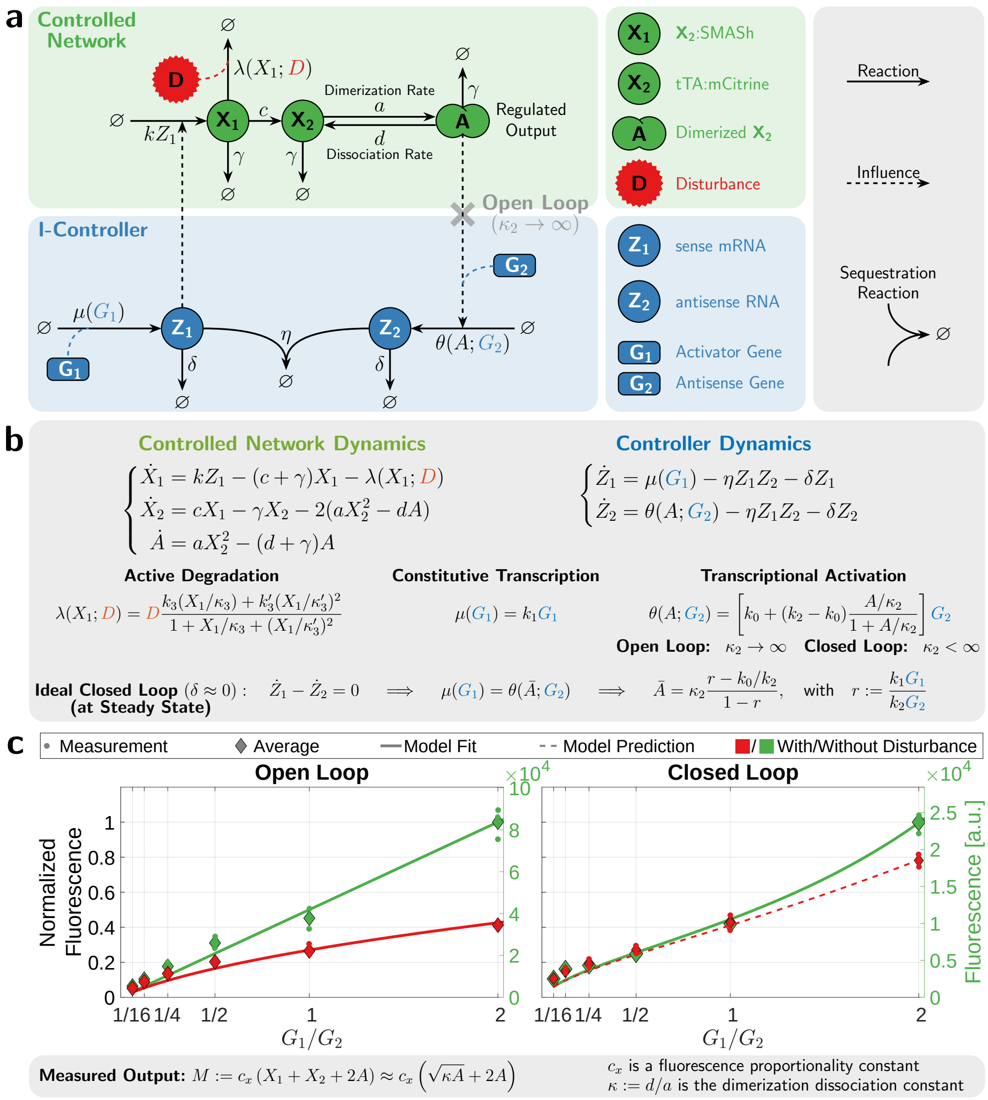
A detailed biochemical reaction network that describes the interactions between the various biochemical species, depicted in Table B.1, is given in Table B.2. These tables are sufficient to provide a mathematical model for the circuit in Figure 3.2(a). The open-loop circuit can be obtained by setting $a_2 = 0$ (in Table B.2) thus preventing the transcriptional activator **A** from binding to the promoter of the anti-sense gene $\mathbf{G_2^F}$. Note that if **A** and **B** are two species, then **A**:**B** is understood to be the complex formed when **A** and **B** are bound together.

Model Reduction

In this section, the full model given in Table B.2 is mathematically reduced to the model described schematically in Figure B.1(a) and mathematically in Figure B.1(b).

Note that Figure B.1 is a special case of Figure 3.4(a) in the main text, where there is only integral control and no network perturbation. The model reduction procedure is based on the following assumption:

Assumption 1. *The binding/unbinding reactions are fast.*



$$\begin{cases}\dot X_1 = kZ_1-(c+\gamma)X_1-\lambda(X_1;D)\\ \dot X_2 = cX_1-\gamma X_2-2(aX_2^2-dA)\\ \dot A = aX_2^2-(d+\gamma)A\end{cases}\qquad \begin{cases}\dot Z_1=\mu(G_1)-\eta Z_1Z_2-\delta Z_1\\ \dot Z_2=\theta(A;G_2)-\eta Z_1Z_2-\delta Z_2\end{cases}$$

$$\lambda(X_1;D)=D\frac{k_3(X_1/\kappa_3)+k_3'(X_1/\kappa_3')^2}{1+X_1/\kappa_3+(X_1/\kappa_3')^2}\qquad \mu(G_1)=k_1G_1\qquad \theta(A;G_2)=\left[k_0+(k_2-k_0)\frac{A/\kappa_2}{1+A/\kappa_2}\right]G_2$$

$$\dot Z_1-\dot Z_2=0\implies \mu(G_1)=\theta(\bar A;G_2)\implies \bar A=\kappa_2\frac{r-k_0/k_2}{1-r},\quad r:=\frac{k_1G_1}{k_2G_2}$$

$$M:=c_x(X_1+X_2+2A)\approx c_x\left(\sqrt{\kappa A}+2A\right)$$

Figure B.1: Mathematical Modeling of the I-Circuit without Network Perturbation in Figure 3.2. (a)/(b) Schematic/Mathematical Description of the Reduced Model. This is a special case of the compact model presented in Figure 3.4(a) where the proportional controller and network perturbation are removed to model only the integral control action. The model for the open-loop circuit is obtained by setting $\kappa_2 \to \infty$, where κ_2 is the dissociation constant of **A** from $\mathbf{G_2}$. This removes the the feedback from the regulated output **A** since $\theta(A; G_2)$ becomes k_0G_2. In the ideal operation of the antithetic integral controller, where the dilution rate δ is negligible with respect to the other rates of the controller, the regulated output **A** has a steady-state concentration, denoted by $\bar{A}$, that is independent of the controlled network parameters. This ensures robust perfect adaptation of the regulated output to external disturbances such as **D**. (c)Model Calibration to Experimental Data. This panel is the same as Figure 3.4(b), but the x-axis is plotted on a linear scale here to examine the concavity of the curves.

Species	**Symbol**	**Species**	**Symbol**
Activator Gene	$\mathbf{G_1}$	Antisense Gene (Free)	$\mathbf{G_2^F}$
Sense mRNA	$\mathbf{Z_1}$	Anti-sense RNA	$\mathbf{Z_2}$
Asunaprevir Disturbance (Free)	$\mathbf{D^F}$	tTA:mCitrine:SMASh	$\mathbf{X_1}$
$\mathbf{X_1}$:$\mathbf{D^F}$	$\mathbf{X_1^*}$	tTA:mCitrine	$\mathbf{X_2}$
Dimerized tTA:mCitrine	$\mathbf{A}$	$\mathbf{G_2^F}$:$\mathbf{A}$	$\mathbf{G_2^*}$

Table B.1: List of Biochemical Species

Reaction	**Mechanism**	**Constants**
Transcription 1	$\mathbf{G_1} \xrightarrow{k_1} \mathbf{G_1} + \mathbf{Z_1}$	
Transcription 2	$\mathbf{G_2^F} + \mathbf{A} \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathbf{G_2^*} \xrightarrow{k_2} \mathbf{G_2^*} + \mathbf{Z_2}$	$\kappa_2 := \frac{d_2}{a_2}$
Leaky Transcription 2	$\mathbf{G_2^F} \xrightarrow{k_0} \mathbf{G_2^F} + \mathbf{Z_2}$	
Translation	$\mathbf{Z_1} \xrightarrow{\boldsymbol{k}} \mathbf{Z_1} + \mathbf{X_1}$	
Active Degradation	$\mathbf{X_1} + \mathbf{D^F} \underset{d_3}{\overset{a_3}{\rightleftharpoons}} \mathbf{X_1^*} \xrightarrow{k_3} \mathbf{D^F}$	$\kappa_3 := \frac{d_3+k_3}{a_3}$
Conversion	$\mathbf{X_1} \xrightarrow{\boldsymbol{c}} \mathbf{X_2}$	
Dimerization	$\mathbf{X_2} + \mathbf{X_2} \underset{\boldsymbol{d}}{\overset{\boldsymbol{a}}{\rightleftharpoons}} \mathbf{A}$	$\kappa := \frac{d}{a}$
Dilution/Degradation	$\mathbf{Z_{\boldsymbol{i}}} \xrightarrow{\delta} \varnothing \quad \mathbf{X_{\boldsymbol{i}}} \xrightarrow{\gamma} \varnothing \quad (\mathrm{i} = 1{,}2) \quad \mathbf{A} \xrightarrow{\gamma} \varnothing$	
Sequestration	$\mathbf{Z_1} + \mathbf{Z_2} \xrightarrow{\eta} \varnothing$	

Table B.2: List of Biochemical Reactions

Assumption 1 allows us to exploit a time-scale separation principle based on the fact that the (un)binding reactions are much faster than the other reactions in the system. As a result, a Quasi-Steady-State Approximation (QSSA) is applied.

Now, we show the mathematical derivation of the reduced model. The conservation laws are given in terms of the total concentrations of bound and free antisense gene and ASV denoted by G_2 and D, respectively. That is, we have

$$\begin{aligned} G_2^F + G_2^* &=: G_2 \\ D^F + X_1^* &=: D. \end{aligned} \tag{B.1}$$

Note that G_1, G_2 and D are constants and are considered to act as external inputs and disturbance to the circuit. Since the binding reactions are much faster than the other reactions in the network (Assumption 1), one can invoke the Quasi-Steady-State Approximation (QSSA) as follows

$$\begin{aligned} \dot{G}_2^* \approx 0 \quad &\implies \quad a_2 G_2^F A - d_2 G_2^* \approx 0 \quad \implies G_2^* \approx \frac{A G_2^F}{\kappa_2} \\ \dot{X}_1^* \approx 0 \quad &\implies \quad a_3 X_1 D^F - (d_3 + k_3) X_1^* \approx 0 \quad \implies X_1^* \approx \frac{X_1 D^F}{\kappa_3}, \end{aligned} \tag{B.2}$$

where the dissociation and Michaelis-Menten constants, κ_2 and κ_3, are given in Table B.2. By substituting the quasi-steady-state approximation of G_2^* in the conservation law $G_2^F + G_2^* = G_2$, we obtain the following expressions

$$G_2^F \approx G_2 \frac{1}{1 + A/\kappa_2}, \quad G_2^* \approx G_2 \frac{A/\kappa_2}{1 + A/\kappa_2}.$$

Similarly, by substituting the quasi-steady-state approximation of X_1^* in the conservation law $D^F + X_1^* = D$, we obtain

$$D^F \approx D \frac{1}{1 + X_1/\kappa_3}, \quad X_1^* \approx D \frac{X_1/\kappa_3}{1 + X_1/\kappa_3}.$$

Equipped with the quasi-steady-state approximations, we can now write down a set of *Ordinary Differential Equations (ODEs)* that describe the evolution of X_1, X_2, A, Z_1 and Z_2.

$$\begin{aligned}
\dot{X}_1 &= kZ_1 - a_3 X_1 D^F + d_3 X_1^* - (c+\gamma)X_1 \approx kZ_1 - k_3 X_1^* - (c+\gamma)X_1 \\
&\approx kZ_1 - \left((c+\gamma)X_1 + k_3 D \frac{X_1/\kappa_3}{1 + X_1/\kappa_3}\right) \\
\dot{X}_2 &= cX_1 - \gamma X_2 - 2\left(aX_2^2 - dA\right) \\
\dot{A} &= aX_2^2 - dA - \gamma A \\
\dot{Z}_1 &= k_1 G_1 - \eta Z_1 Z_2 - \delta Z_1 \\
\dot{Z}_2 &= k_2 G_2^* + k_0 G_2^F - \eta Z_1 Z_2 - \delta Z_2 \approx G_2 \left(\frac{k_0 + k_2(A/\kappa_2)}{1 + A/\kappa_2}\right) - \eta Z_1 Z_2 - \delta Z_2.
\end{aligned}$$

Finally, the dynamics of the reduced model can be more compactly written as

$$\begin{cases}
\dot{X}_1 = kZ_1 - (c+\gamma)X_1 - \lambda(X_1; D) \\
\dot{X}_2 = cX_1 - \gamma X_2 - 2\left(aX_2^2 - dA\right) \\
\dot{A} = aX_2^2 - (d+\gamma)A \\
\dot{Z}_1 = \mu(G_1) - \eta Z_1 Z_2 - \delta Z_1 \\
\dot{Z}_2 = \theta(A; G_2) - \eta Z_1 Z_2 - \delta Z_2,
\end{cases} \quad \text{with} \quad
\begin{cases}
\lambda(X_1; D) := k_3 D \dfrac{X_1/\kappa_3}{1 + X_1/\kappa_3} \\
\mu(G_1) := k_1 G_1 \\
\theta(A; G_2) := \left[k_0 + (k_2 - k_0)\dfrac{A/\kappa_2}{1 + A/\kappa_2}\right] G_2,
\end{cases} \tag{B.3}$$

where $k_0 << k_2$ since leaky transcription is usually much slower than activated transcription. Note that in the open-loop circuit, only the leaky transcription reaction can occur, that is $a_2 = 0$. As a result, $\kappa_2 \to \infty$, and therefore the function θ becomes independent of A, i.e. $\theta(A; G_2) = k_0 G_2$.

The only difference between the dynamics given in (B.3) and Figure B.1(b) lies in the active degradation function λ. In fact, by setting $k_3' = 0$ and $\kappa_3' \to \infty$ in Figure B.1(b), we obtain (B.3). The active degradation function of Figure B.1(b) is, in general, of higher order and involves squared terms X_1^2. It turns out that this higher order function is necessary to fit the data properly (refer to Supplementary Information B.1 for a detailed explanation). The mechanism underlying this higher-order active degradation function is explained in the subsequent section.

Higher Order Active Degradation

The active degradation function $\lambda(X_1; D)$ in (B.3) takes the form of a hill function multiplied by the (disturbance) Asunaprevir concentration D. We now consider the more general hill function of Figure B.1(b) which is given by

$$\lambda(X_1; D) = D \frac{k_3 X_1/\kappa_3 + k_3' (X_1/\kappa_3')^2}{1 + X_1/\kappa_3 + (X_1/\kappa_3')^2}. \tag{B.4}$$

This function has a higher order (hill coefficient) since it involves squared terms X_1^2. We show next how this hill function can mechanistically arise from the interactions of the various species in the circuit.

In addition to the Active Degradation reaction given in Table B.2, we allow $\mathbf{X_1}$ to dimerize to form the complex $\mathbf{A_1}$ (Dimerized tTA:mCitrine:SMASh) which can also be actively degraded by $\mathbf{D}$. These additional mechanisms are modeled by appending the previous model with the additional active degradation and dimerization reactions listed in Table B.3. Note that, theoretically, $\mathbf{A_1}$ can still release the SMASh tag at some rate c' and/or may still be able to initiate transcription of $\mathbf{G_2}$ at some rate k_2'. However, we assume that the bulk dimer $\mathbf{A_1}$ is very unstable and tends to dissociate at a rate $d_3' >> c'$ and its transcription rate is much slower than that of $\mathbf{A}$. As a result, c' and k_2' can be neglected and thus the corresponding reactions are not listed in Table B.3.

Now, we show the mathematical derivation of the higher order active degradation function $\lambda(X_1; D)$. The mathematical procedure is, once again, based on Assumptions 1; however, an additional assumption is added here as well.

Reaction	Mechanism	Constants
Dimerization 2	$\mathbf{X_1} + \mathbf{X_1} \underset{d_1'}{\overset{a_1'}{\rightleftharpoons}} \mathbf{A_1}$	$\kappa_1' := \frac{d_1'}{a_1'}$
Active Degradation 2	$\mathbf{A_1} + \mathbf{D^F} \underset{d_3'}{\overset{a_3'}{\rightleftharpoons}} \mathbf{A_1^*} \xrightarrow{k_3'} \mathbf{D^F}$	$\kappa_3' := \frac{d_3'+k_3'}{a_3'}$

Table B.3: Additional Biochemical Reactions for Higher Order Active Degradation

Assumption 2. *The concentration of the free ASV molecules D^F is low.*

This assumption means that the majority of the ASV molecules are in their bound state. The conservation law that can be seen from Tables B.2 and B.3 is given by

$$D^F + X_1^* + A_1^* =: D. \tag{B.5}$$

This replaces the conservation law for D given in (B.1) since the species $\mathbf{A_1^*}$ is a complex formed of $\mathbf{A_1}$ and $\mathbf{D^F}$. With Assumption 1 in mind, one can invoke the Quasi-Steady-State Approximation (QSSA) as follows

$$\begin{aligned} \dot{A}_1^* \approx 0 \quad &\implies \quad a_3' A_1 D^F - (d_3' + k_3') A_1^* \approx 0 \quad \implies A_1^* \approx \frac{A_1 D^F}{\kappa_3'} \\ \dot{A}_1 \approx 0 \quad &\implies \quad a_1' X_1^2 - d_1' A_1 - (a_3' A_1 D^F - d_3' A_1^*) \approx 0 \quad \implies A_1 \approx \frac{X_1^2}{\kappa_1' + \frac{k_3'}{a_1' \kappa_3'} D^F}, \end{aligned} \tag{B.6}$$

where the constants κ_1' and κ_3' are given in Table B.3. By substituting the quasi-steady-state approximations of X_1^* from (B.2) and (A_1, A_1^*) from (B.6) in the conservation law $D^F + X_1^* + A_1^* = D$, while invoking Assumption 2 (more precisely $\kappa_1' >> \frac{k_3'}{a_1'\kappa_3'} D^F$), we obtain

$$D^F \approx D \frac{1}{1 + X_1/\kappa_3 + X_1^2/\kappa_1'\kappa_3'}, \quad X_1^* \approx D \frac{X_1/\kappa_3}{1 + X_1/\kappa_3 + X_1^2/\kappa_1'\kappa_3'}, \quad A_1^* \approx D \frac{X_1^2/\kappa_1'\kappa_3'}{1 + X_1/\kappa_3 + X_1^2/\kappa_1'\kappa_3'}.$$

Equipped with the quasi-steady-state approximations, we can now update the *Ordinary Differential Equation (ODE)* that describes the evolution of X_1.

$$\begin{aligned} \dot{X}_1 &= kZ_1 - (a_3 X_1 D^F - d_3 X_1^*) - (c + \gamma) X_1 - 2(a_1' X_1^2 - d_1' A_1) \\ &\approx kZ_1 - k_3 X_1^* - (c + \gamma) X_1 - 2(a_3' A_1 D^F - d_3' A_1^*) \\ &\approx kZ_1 - k_3 X_1^* - (c + \gamma) X_1 - 2k_3' A_1^* \\ &\approx kZ_1 - (c + \gamma) X_1 - D \frac{k_3 X_1/\kappa_3 + 2k_3' X_1^2/\kappa_1'\kappa_3'}{1 + X_1/\kappa_3 + X_1^2/\kappa_1'\kappa_3'} =: kZ_1 - (c + \gamma) X_1 - \lambda(X_1; D), \end{aligned}$$

and thus the active degradation function λ takes the intended form given in (B.4) (with slight abuse of notation) and shown in Figure B.1(b).

Mathematical Model of the Measured Output: Fluorescence

Let $M^i(G_1, D)$ denote the measured fluorescence at a given concentration of the activator gene G_1 and drug disturbance D, where $i = o$ corresponds to open-loop measurements while $i = c$ corresponds to closed-loop measurements. Note that G_2 is held constant throughout the paper, and thus the explicit dependence of the measured fluorescence M on G_2 is not emphasized here. Fluorescence is emitted from all molecules containing mCitrine, that is species $\mathbf{X_1}$, $\mathbf{X_2}$ and $\mathbf{A}$ (see Figure B.1(a)). Hence we have

$$M^i(G_1, D) = c_x \left(X_1^i + X_2^i + 2A^i \right) \qquad i \in \{o, c\}, \tag{B.7}$$

where c_x is a proportionality constant that maps concentrations (in nM) to fluorescence (in a.u.). Recall that the dynamics of X_1, X_2 and A (in both open and closed loop) are given by

$$\begin{cases}\dot{X}_1 = kZ_1 - (c+\gamma)X_1 - \lambda(X_1; D)\\ \dot{X}_2 = cX_1 - \gamma X_2 - 2(aX_2^2 - dA),\\ \dot{A} = aX_2^2 - (d+\gamma)A.\end{cases}$$

At steady state, we have $\bar{A} = \bar{X}_2^2/\kappa$ with $\kappa := (d+\gamma)/a$. which implies that $\bar{X}_1 = \frac{\gamma}{c}\left(\bar{X}_2 + 2\bar{A}\right)$. This implies that the measurement at steady state is given by

$$\bar{M}^i(G_1, D) = c_x\frac{\gamma+c}{c}\left(\bar{X}_2^i + 2\bar{A}^i\right) = c_x\frac{\gamma+c}{c}\left(\sqrt{\kappa\bar{A}^i} + 2\bar{A}^i\right).$$

This equation links the measured fluorescence at steady state, $\bar{M}^i$, to the steady-state concentration of the regulated output $\bar{A}^i$ for both open- and closed-loop circuits. In fact, this equation is also approximately valid transiently (not only at steady state) under the following assumption.

Assumption 3. *Transcription is much slower than releasing the SMASh tag.*

This assumption is reasonable since $\mathbf{X_1}$ is very unstable and tends to release the SMASh tag very quickly (i.e. c is large). Under this assumption, a quasi-steady-state argument can be used to obtain the following approximations $A \approx X_2^2/\kappa$ and $X_1 \approx \frac{\gamma}{c}(X_2 + 2A)$ which yield the same measurement equation at any time t (not just at steady state). Furthermore, we can drop the factor $(\gamma + c)/c$ since $c >> \gamma$ to obtain the expression shown in the bottom of Figure B.1(c) and Figure 3.4(a).

Model Calibration to the Experimental Data

In this section, we calibrate the model depicted in Figures B.1(a) and (b) to fit the experimentally collected data at steady state.

Mathematical Representation of the Data

The experiments that are carried out allows us to obtain the data set visualized in Figure B.2. The mathematical notation for the available data is described in Table B.4.

	No Disturbance (D = 0 nM)	**With Disturbance** (D = 33 nM)
Open Loop	$\left\{G_1^j, \hat{M}_j^{o,ND}\right\}_{j=1}^{N^o}$	$\left\{G_1^j, \hat{M}_j^{o,D}\right\}_{j=1}^{N^o}$
Closed Loop	$\left\{G_1^j, \hat{M}_j^{c,ND}\right\}_{j=1}^{N^c}$	$\left\{G_1^j, \hat{M}_j^{c,D}\right\}_{j=1}^{N^c}$

Table B.4: Data Representation: Set of inputs G_1 and measured outputs $\hat{M}$ for the open- and closed-loop circuits with and without disturbance. N^o (resp. N^c) denotes the number of different concentrations of $\mathbf{G_1}$ that are applied to the open-loop (resp. closed-loop) circuit. For each G_1^j ($j = 1, ..., N^o$ or N^c), four measurements $\hat{M}_j^{(.)}$ are obtained. $\hat{M}_j^{o,ND}$ (resp. $\hat{M}_j^{o,D}$) denotes the measurement for the open-loop circuit without (resp. with) disturbance; whereas, $\hat{M}_j^{c,ND}$ (resp. $\hat{M}_j^{c,D}$) denotes the measurement for the closed-loop circuit without (resp. with) disturbance. Constant disturbances $D = 30$ nM are applied and the concentration of the antisense plasmid $G_2 = 0.004$ pmol is kept constant throughout all the experiments. Note that the measurements $\hat{M}$ represent the average of the experimentally obtained triplicates.

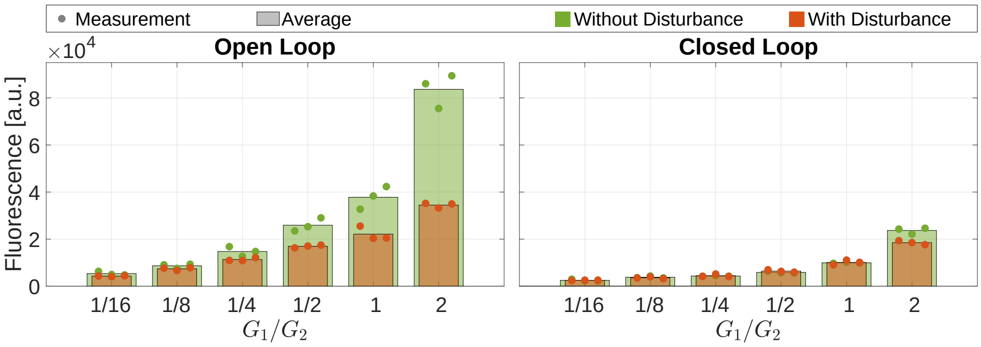


Figure B.2: Data used for fitting the circuit realizing an antithetic integral controller. The fluorescence data are obtained for a wide range of plasmid ratios G_1/G_2, for open/closed loop settings, and with/without disturbance by fixing $G_2 = 0.004$ pmol and sweeping G_1 accordingly.

Necessity of Higher Order Active Degradation

We select the higher order active degradation function introduced in Supplementary Information B.1 because it will be shown that the negative concavity of the fitted curve in the open-loop setting with disturbance (see Figure B.1(c)) cannot be captured by a simple first order hill function λ.

The fixed point of the dynamics in the open-loop setting, denoted by $\left(\bar{X}_1^o, \bar{X}_2^o, \bar{A}^o, \bar{Z}_1^o, \bar{Z}_2^o\right)$ satisfies the following set of nonlinear algebraic equations.

$$\begin{cases} 0 = k\bar{Z}_1^o - (c+\gamma)\bar{X}_1^o - \lambda(\bar{X}_1^o; D) \\ 0 = c\bar{X}_1^o - \gamma\bar{X}_2^o - 2\left(a\left(\bar{X}_2^o\right)^2 - d\bar{A}^o\right) \\ 0 = a\left(\bar{X}_2^o\right)^2 - (d+\gamma)\bar{A}^o \\ 0 = k_1G_1 - \eta\bar{Z}_1^o\bar{Z}_2^o - \delta\bar{Z}_1^o \\ 0 = k_0G_2 - \eta\bar{Z}_1^o\bar{Z}_2^o - \delta\bar{Z}_2^o, \end{cases} \tag{B.8}$$

and the measured output denoted by $\bar{M}^o(G_1, D)$ is given by

$$\bar{M}^o(G_1, D) = c_x\frac{\bar{c}}{c}\left(\bar{X}_2^o + 2\bar{A}^o\right), \tag{B.9}$$

where $\bar{c} := c + \gamma$. To analyze the concavity of $\bar{M}^o$ as a function of G_1, we study the sign of the second derivative $\frac{\partial^2\bar{M}^o}{\partial G_1^2}$. First, observe that using the second and third equations in (B.8), the measured output can be rewritten as

$$\begin{aligned} \bar{M}^o(G_1, D) &= c_x\frac{\bar{c}}{c}\left(\bar{X}_2^o + \frac{2}{\kappa}\left(\bar{X}_2^o\right)^2\right) \\ &= c_x\frac{\bar{c}}{c}\frac{c}{\gamma}\bar{X}_1^o \\ &= c_x\frac{\bar{c}}{\gamma}\bar{X}_1^o. \end{aligned}$$

Hence the sign of $\frac{\partial^2\bar{M}^o}{\partial G_1^2}$ is the same as that of $\frac{\partial^2\bar{X}_1^o}{\partial G_1^2}$ which we derive next. Taking the second derivative of the first equation in (B.8) with respect to G_1 yields

$$k\frac{\partial^2\bar{Z}_1^o}{\partial G_1^2} - \bar{c}\frac{\partial^2\bar{X}_1^o}{\partial G_1^2} - \frac{\partial^2\lambda(\bar{X}_1^o; D)}{\partial X_1^2}\left(\frac{\partial\bar{X}_1^o}{\partial G_1}\right)^2 - \frac{\partial\lambda(\bar{X}_1^o; D)}{\partial X_1}\frac{\partial^2\bar{X}_1^o}{\partial G_1^2} = 0 \implies \frac{\partial^2\bar{X}_1^o}{\partial G_1^2} = \frac{k\frac{\partial^2\bar{Z}_1^o}{\partial G_1^2} - \frac{\partial^2\lambda(\bar{X}_1^o;D)}{\partial X_1^2}\left(\frac{\partial\bar{X}_1^o}{\partial G_1}\right)^2}{\bar{c} + \frac{\partial\lambda(\bar{X}_1^o;D)}{\partial X_1}}. \tag{B.10}$$

Next we derive an expression for $\frac{\partial^2 \bar{Z}_1^o}{\partial G_1^2}$. Subtracting the last two equations in (B.8) from each other yields

$$\frac{k_1 G_1 - k_0 G_2}{\delta} = \bar{Z}_1^o - \bar{Z}_2^o \implies \begin{cases} \dfrac{\partial \bar{Z}_1^o}{\partial G_1} - \dfrac{\partial \bar{Z}_2^o}{\partial G_1} = \dfrac{k_1}{\delta} \\ \dfrac{\partial^2 \bar{Z}_1^o}{\partial G_1^2} = \dfrac{\partial^2 \bar{Z}_2^o}{\partial G_1^2}. \end{cases} \tag{B.11}$$

By taking the first derivative of the fourth equation in (B.8) with respect to G_1 and exploiting (B.11), we obtain an expression for $\frac{\partial \bar{Z}_1^o}{\partial G_1}$ given by

$$\frac{\partial \bar{Z}_1^o}{\partial G_1} = \frac{k_1}{\delta} \frac{\eta \bar{Z}_1^o + \delta}{\eta(\bar{Z}_1^o + \bar{Z}_2^o) + \delta}.$$

Furthermore, by taking the second derivative of the fourth equation in (B.8) with respect to G_1 and exploiting (B.11), we obtain an expression for $\frac{\partial^2 \bar{Z}_1^o}{\partial G_1^2}$ given by

$$\frac{\partial^2 \bar{Z}_1^o}{\partial G_1^2} = 2\eta \frac{k_1^2}{\delta^2} \frac{\left(\eta \bar{Z}_1^o + \delta\right)^2}{\left[\eta(\bar{Z}_1^o + \bar{Z}_2^o) + \delta\right]^3} > 0.$$

Therefore, observe using (B.10), that as long as λ is an increasing function of X_1 $\left(\frac{\partial \lambda(X_1;D)}{\partial X_1} \geq 0\right)$ with negative concavity $\left(\frac{\partial^2 \lambda(X_1;D)}{\partial X_1^2} \leq 0\right)$, we have $\frac{\partial^2 \bar{X}_1^o}{\partial G_1^2} > 0$. Therefore for a first order active degradation function $\lambda(X_1; D) = k_3 D \frac{X_1/\kappa_3}{1+X_1/\kappa_3}$, we have that $\frac{\partial^2 \bar{M}^o}{\partial G_1^2} > 0$ and, as a result, it cannot capture the negative concavity of the measurements in the open-loop setting with disturbance shown in Figure B.1(c). On the other hand a second order active degradation function λ given in (B.4) is capable of changing the concavity and hence is adopted in the paper.

Choice of Parameter Groups

Consider the circuit depicted in Figures B.1(a) and (b). The measured output for the open- ($i = o$) and closed-loop ($i = c$) settings is given by

$$\begin{aligned} &\bar{M}^i(G_1, D) = c_x\left(\sqrt{\kappa \bar{A}} + 2\bar{A}\right), \quad \text{for } i \in \{o, c\} \\ &\text{with } \begin{cases} 0 = k\bar{Z}_1 - (c+\gamma)\bar{X}_1 - D\dfrac{k_3 \bar{X}_1/\kappa_3 + k_3' \left(\bar{X}_1/\kappa_3'\right)^2}{1 + \bar{X}_1/\kappa_3 + \left(\bar{X}_1/\kappa_3'\right)^2} \\ 0 = c\bar{X}_1 - \gamma \bar{X}_2 - 2(a\bar{X}_2^2 - d\bar{A}) \\ 0 = a\bar{X}_2^2 - (d+\gamma)\bar{A} \\ 0 = k_1 G_1 - \eta \bar{Z}_1 \bar{Z}_2 - \delta \bar{Z}_1 \\ 0 = \theta^i(\bar{A}; G_2) - \eta \bar{Z}_1 \bar{Z}_2 - \delta \bar{Z}_2, \qquad \theta^i(A; G_2) = \begin{cases} k_0 G_2, & \text{for } i = o \\ \left[k_0 + (k_2 - k_0)\dfrac{A/\kappa_2}{1 + A/\kappa_2}\right] G_2, & \text{for } i = c. \end{cases} \end{cases} \end{aligned} \tag{B.12}$$

The system model has 15 parameters: $\{k, k_0, k_1, k_2, k_3, k_3', \kappa_2, \kappa_3, \kappa_3', c, a, d, \delta, \gamma, \eta\}$ to be calibrated to the data. Furthermore, the measurement equation has an additional parameter c_x to be calibrated as well and thus summing up to 16 total parameters. However, steady-state measurements cannot uniquely identify all of those parameters. For this reason, we carry out a suitable choice of re-parameterization to obtain a minimal number of (aggregated) parameter groups that can be uniquely identified from the steady-state measurements. Particularly, define the following parameter groups

$$\begin{aligned} \Delta_1 &:= \frac{k_3}{\kappa_3} \frac{1}{c+\gamma}, & \Delta_2 &:= \frac{k_3'}{\kappa_3'^2} \frac{\gamma}{c+\gamma} \frac{1}{cc_x}, & \Delta_1' &:= \frac{1}{\kappa_3} \frac{\gamma}{cc_x}, & \Delta_2' &:= \frac{1}{\kappa_3'^2} \frac{\gamma^2}{c^2 c_x^2}, \\ \bar{\delta} &:= \frac{\delta}{k_1} & \bar{\gamma} &:= \frac{\gamma}{k} \frac{c+\gamma}{c} \frac{1}{c_x}, & \bar{\kappa} &:= c_x \kappa, & \bar{\kappa}_2 &:= c_x \sqrt{\kappa \kappa_2} \\ \bar{\eta} &:= \frac{\eta}{k_1}, & \bar{k}_0 &:= \frac{k_0}{k_1}, & \bar{k}_2 &:= \frac{k_2}{k_1}. \end{aligned} \tag{B.13}$$

and the following transformed variables

$$\tilde{X}_1 := c_x \frac{c}{\gamma} \bar{X}_1 \qquad \text{and} \qquad \tilde{X}_2 := c_x \bar{X}_2. \tag{B.14}$$

Then the steady-state measurements can be rewritten in terms of the parameter groups and transformed variables as

$$\begin{aligned} \bar{M}^i(G_1, D) = \tilde{X}_1, \quad \text{for } i \in \{o, c\}, \qquad \text{such that} \quad & \begin{cases} \tilde{X}_1 = \tilde{X}_2 + 2\dfrac{\tilde{X}_2^2}{\bar{\kappa}} \\ \bar{Z}_1 = \bar{\gamma}\left[\tilde{X}_1 + D\dfrac{\Delta_1\tilde{X}_1 + \Delta_2\tilde{X}_1^2}{1 + \Delta_1'\tilde{X}_1 + \Delta_2'\tilde{X}_1^2}\right] \\ \bar{Z}_2 = \dfrac{G_1}{\bar{\eta}\bar{Z}_1} - \dfrac{\bar{\delta}}{\bar{\eta}} \\ 0 = \bar{\theta}^i(\tilde{X}_2; G_2) - \bar{\eta}\bar{Z}_1\bar{Z}_2 - \bar{\delta}\bar{Z}_2 \end{cases} \\ \text{where} \qquad \bar{\theta}^i(\tilde{X}_2; G_2) = & \begin{cases} \bar{k}_0 G_2, & \text{for } i = o \\ \left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\dfrac{\left(\tilde{X}_2/\bar{\kappa}_2\right)^2}{1 + \left(\tilde{X}_2/\bar{\kappa}_2\right)^2}\right] G_2, & \text{for } i = c. \end{cases} \end{aligned} \tag{B.15}$$

Note that $i = o$ and $i = c$ correspond to the open- and closed-loop settings, respectively. Hence given the input G_1 and the disturbance D, one can use (B.15) to compute the mCitrine measurement in the open- and closed-loop settings. To do so, one has to solve the set of nonlinear algebraic equations. This is done in Matlab by recasting the set of nonlinear algebraic equations as a single but high order polynomial in $\tilde{X}_2$ using the symbolic toolbox and then solving the polynomial using the command "roots". This allows us to solve the system of equations more efficiently (by computing eigenvalues of a companion matrix associated with the obtained polynomial) without requiring an initial guess as in Newton-Raphson-like methods. This is particularly important since model calibration may require solving this system of equations thousands of times. We close this section by observing that we have now reduced the parameters to be calibrated down to 11 (as compared to 16).

Model Calibration Steps

The model calibration is carried out in four steps to avoid over-fitting. In the first step, the parameters of the model for the open-loop circuit in the absence of disturbance ($D = 0$) are fit to the data $\left\{G_1^j, \hat{M}_j^{o,ND}\right\}_{j=1}^{N^o}$ by solving the following optimization problem for $\theta^{o,ND} := \begin{bmatrix}\bar{\gamma} & \bar{\delta} & \bar{\eta} & \bar{k}_0\end{bmatrix}$.

$$\begin{aligned} \min_{\Theta^{o,ND}} J(\Theta^{o,ND}) \; = \; & \sum_{j=1}^{j=N^o} \left[\hat{M}_j^{o,ND} - \bar{M}^o(G_1^j, 0)\right]^2 \\ \text{s.t.} \qquad \bar{M}^o(G_1^j, 0) = \tilde{X}_1 \quad \text{and} \quad & \begin{cases} \bar{Z}_1 = \bar{\gamma}\tilde{X}_1 \\ \bar{Z}_2 = \dfrac{G_1^j}{\bar{\eta}\bar{Z}_1} - \dfrac{\bar{\delta}}{\bar{\eta}} \\ 0 = \bar{k}_0 G_2 - \bar{\eta}\bar{Z}_1\bar{Z}_2 - \bar{\delta}\bar{Z}_2, \end{cases} \end{aligned}$$

where $G_2 = 0.004$ pmol. Note that this system of equations can be rewritten in terms of $\tilde{X}_1$ explicitly as a second degree polynomial.

In the second step, we first fix the parameters that are obtained from the previous fit. Then, the parameters of the model for the closed-loop circuit in the absence of disturbance ($D = 0$) are fit to the data $\left\{G_1^j, \hat{M}_j^{c,ND}\right\}_{j=1}^{N^c}$ by solving the following optimization problem for $\theta^{c,ND} := \begin{bmatrix}\bar{\kappa} & \bar{\kappa}_2 & \bar{k}_2\end{bmatrix}$.

$$\min_{\Theta^{c,ND}} J(\Theta^{c,ND}) = \sum_{j=1}^{j=N^c} \left[\hat{M}_j^{c,ND} - \bar{M}^c(G_1^j, 0)\right]^2$$

$$\text{s.t.} \qquad \bar{M}^c(G_1^j, 0) = \tilde{X}_1 \quad \text{and} \quad \begin{cases} \tilde{X}_1 = \tilde{X}_2 + 2\dfrac{\tilde{X}_2^2}{\bar{\kappa}} \\ \bar{Z}_1 = \bar{\gamma}\tilde{X}_1 \\ \bar{Z}_2 = \dfrac{G_1^j}{\bar{\eta}\bar{Z}_1} - \dfrac{\bar{\delta}}{\bar{\eta}} \\ 0 = \left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\dfrac{\tilde{X}_2^2/\bar{\kappa}_2}{1 + \tilde{X}_2^2/\bar{\kappa}_2}\right] G_2 - \bar{\eta}\bar{Z}_1\bar{Z}_2 - \bar{\delta}\bar{Z}_2, \end{cases}$$

where $G_2 = 0.004$ pmol. Note that the last equation can be rewritten in terms of $\tilde{X}_2$ explicitly as a sixth degree polynomial.

In the third step, we first fix the parameters that are obtained from the previous fits. Then, the parameters of the model for the open-loop circuit in the presence of disturbance ($D = 33$ nM) are fit to the data $\left\{G_1^j, \hat{M}_j^{o,D}\right\}_{j=1}^{N^o}$ by solving the following optimization problem for $\theta^{o,D} := \begin{bmatrix}\Delta_1 & \Delta_2 & \Delta_1' & \Delta_2'\end{bmatrix}$.

$$\min_{\Theta^{o,D}} J(\Theta^{o,D}) = \sum_{j=1}^{j=N^o} \left[\hat{M}_j^{o,D} - \bar{M}^o(G_1^j, D)\right]^2$$

$$\text{s.t.} \qquad \bar{M}^o(G_1^j, 0) = \tilde{X}_1 \quad \text{and} \quad \begin{cases} \bar{Z}_1 = \bar{\gamma}\left[\tilde{X}_1 + D\dfrac{\Delta_1\tilde{X}_1 + \Delta_2\tilde{X}_1^2}{1 + \Delta_1'\tilde{X}_1 + \Delta_2'\tilde{X}_1^2}\right] \\ \bar{Z}_2 = \dfrac{G_1^j}{\bar{\eta}\bar{Z}_1} - \dfrac{\bar{\delta}}{\bar{\eta}} \\ 0 = \bar{k}_0 G_2 - \bar{\eta}\bar{Z}_1\bar{Z}_2 - \bar{\delta}\bar{Z}_2, \end{cases}$$

where $G_2 = 0.004$ pmol, $D = 33$ nM. Note that the last equation can be rewritten in terms of $\tilde{X}_1$ explicitly as a sixth degree polynomial.

In the last step, we already have all the parameter groups, and thus we use them to mathematically predict the measurements for the closed-loop circuit with disturbance ($D = 33$ nM). This is done by using (B.15) where the system of equations can be rewritten as a polynomial in $\tilde{X}_2$ of degree 13.

Estimated Parameter Groups: The model fit and prediction are shown in Figures 3.4(b) and B.1(c), where the optimally estimated parameter groups are given by

$$\begin{aligned} &\Delta_1 \approx 0\ \mathrm{nM}^{-1}, && \Delta_2 = 1.26 \times 10^{-6}\ \mathrm{a\,u}^{-1}\mathrm{nM}^{-1}, && \Delta_1' \approx 0\ \mathrm{a\,u}^{-1}, && \Delta_2' \approx 0\ \mathrm{a\,u}^{-2}, \\ &\bar{\delta} = 9.5 \times 10^{-3} && \bar{\gamma} = 1.01 \times 10^{-5}\ \mathrm{a\,u}^{-1}, && \bar{\kappa} = 10.1 \times 10^{4}\ \mathrm{a.u.}, && \bar{\kappa}_2 = 1.22 \times 10^{4}\ \mathrm{a\,u}^{-1} \\ &\bar{\eta} = 2.59 \times 10^{3}\ \mathrm{nM}^{-1}, && \bar{k}_0 \approx 0, && \bar{k}_2 = 2.13. \end{aligned} \tag{B.16}$$

The estimated parameter groups suggest that leaky transcription of the antisense gene is negligible. They also suggest that the active degradation function is approximately purely quadratic in X_1 that is $\lambda(X_1; D) \approx k_3 G(X_1/\kappa_3')^2$ which means that the degradation of the dimerized tTA:mCitrine:SMASh dominates the degradation of the monomer.

B.2 Mathematical Modeling of the Circuit in Figure 3.2 with Network Perturbation

Consider the circuit depicted in Figure 3.2(a) with network perturbation. The difference here compared to the circuit analyzed in Supplementary Information B.1, lies in the additional plasmid encoding a gene that expresses an RNA-binding protein capable of inhibiting the translation of the sense mRNA. We first describe the additional mechanistic interactions that are introduced by this gene, and then carry out a model reduction technique that allows us to analyze the steady-state behavior of the output protein. Finally we provide the technical details of fitting the model to the experimental data.

Full Model Description

A detailed biochemical reaction network that models the dynamics of the circuit in Figure 3.2(a) can be obtained by appending Tables B.1 and B.2 (that model the circuit of Figure 3.2(a) without network perturbation) by the list of biochemical species and reactions given in Tables B.5 and B.6, respectively. These species and reactions describe the biochemical reaction sub-network that is introduced by the additional gene encoding for NES-L7Ae.

Species	Symbol	Species	Symbol
NES-L7Ae Gene (Free)	$\mathbf{G'^F_2}$	Activated Gene $\mathbf{G'^*_2}$	$\mathbf{G'^F_2{:}A}$
NES-L7Ae mRNA	$\mathbf{Z'_2}$	NES-L7Ae Protein	$\mathbf{X'_1}$
$\mathbf{Z'_1}$	$\mathbf{Z_1{:}X'_1}$		

Table B.5: List of Additional Biochemical Species

Reaction	Mechanism	Constants
Transcription 2′	$\mathbf{G'^F_2} + \mathbf{A} \underset{d_2}{\overset{a_2}{\rightleftharpoons}} \mathbf{G'^*_2} \xrightarrow{k'_2} \mathbf{G'^*_2} + \mathbf{Z'_2}$	$\kappa_2 := \frac{d_2}{a_2}$
Leaky Transcription 2′	$\mathbf{G'^F_2} \xrightarrow{k'_0} \mathbf{G'^F_2} + \mathbf{Z'_2}$	
Translation 2′	$\mathbf{Z'_2} \xrightarrow{k'} \mathbf{Z'_2} + \mathbf{X'_1}$	
Inhibition	$\mathbf{X'_1} + \mathbf{Z_1} \underset{d'}{\overset{a'}{\rightleftharpoons}} \mathbf{Z'_1}$	
Dilution/Degradation	$\mathbf{X'_1} \xrightarrow{\gamma'} \varnothing \qquad \mathbf{Z'_1} \xrightarrow{\delta} \varnothing \qquad \mathbf{Z'_2} \xrightarrow{\delta'} \varnothing$	
Sequestration 2	$\mathbf{Z'_1} + \mathbf{Z_2} \xrightarrow{\eta} \varnothing$	

Table B.6: List of Additional Biochemical Reactions

Model Reduction

In this section, the full model given in Tables B.2 and B.6 is mathematically reduced to the model described schematically in Figure B.3(a) and mathematically in Figure B.3(b). Note that Figure B.3 is a special case of Figure 3.4(a) in the main text, where there is only integral control but with network perturbation. The model reduction procedure is based on the reduced model that is obtained previously in Supplementary Information B.1.

One additional conservation law is appended here to the previous conservation laws in (B.1). It is given in terms of the total concentration of the bound and free NES-L7Ae gene denoted by ${G'}_2$. That is, we have

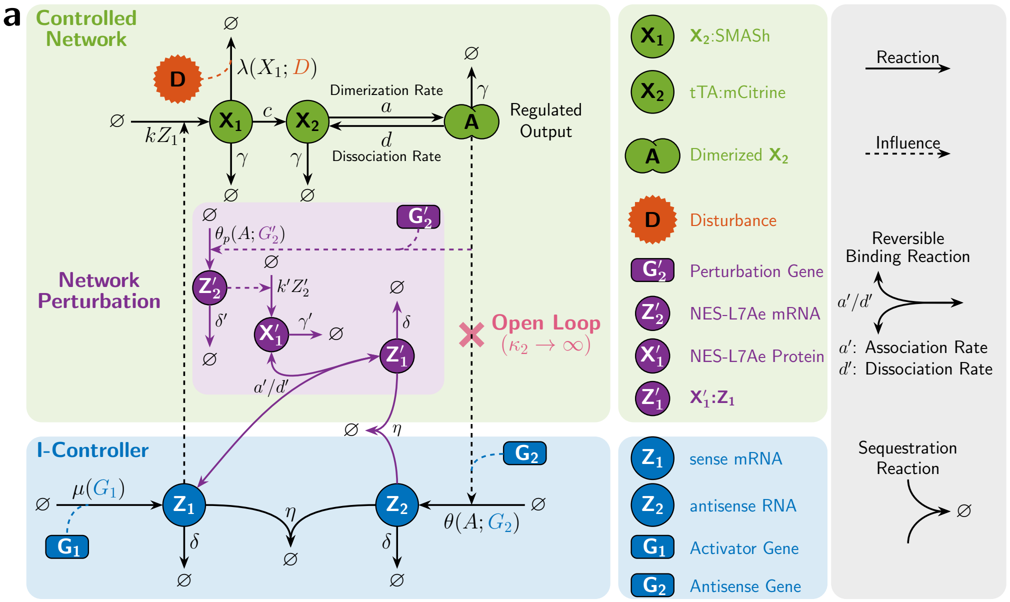
$$G'^F_2 + G'^*_2 =: G'_2. \tag{B.17}$$

Note that ${G'}_2$ is a constant that is considered to act as an external perturbation to the circuit. Since the binding reactions are much faster than the other reactions in the network (Assumption 1), one can invoke the Quasi-Steady-State Approximation (QSSA) as follows

$$\dot{G'}^*_2 \approx 0 \quad \implies \quad a_2 {G'}^F_2 A - d_2 {G'}^*_2 \approx 0 \quad \implies {G'}^*_2 \approx \frac{A G'^F_2}{\kappa_2}, \tag{B.18}$$

where the dissociation constant κ_2 is given in Table B.6 and is assumed to be equal to that in Table B.2 since $\mathbf{G_2}$ and $\mathbf{G'_2}$ have the same promoters. By substituting the quasi-steady-state approximation of ${G'}^*_2$ in the conservation law ${G'}^F_2 + {G'}^*_2 =: {G'}_2$, we obtain the following expressions

$${G'}^F_2 \approx {G'}_2 \frac{1}{1 + A/\kappa_2}, \qquad {G'}^*_2 \approx {G'}_2 \frac{A/\kappa_2}{1 + A/\kappa_2}.$$



b

Controlled Network Dynamics

$$\begin{cases} \dot{X}_1 = kZ_1 - (c+\gamma)X_1 - \lambda(X_1; D) \\ \dot{X}_2 = cX_1 - \gamma X_2 - 2(aX_2^2 - dA) \\ \dot{A} = aX_2^2 - (d+\gamma)A \end{cases}$$

Controller Dynamics

$$\begin{cases} \dot{Z}_1 = \mu(G_1) - \eta Z_1 Z_2 - \delta Z_1 - (a'X'_1 Z_1 - d'Z'_1) \\ \dot{Z}_2 = \theta(A; G_2) - \eta(Z_1 + Z'_1)Z_2 - \delta Z_2 \end{cases}$$

Network Perturbation Dynamics

$$\begin{cases} \dot{Z}'_1 = (a'X'_1 Z_1 - d'Z'_1) - \eta Z_2 Z'_1 - \delta Z'_1 \\ \dot{Z}'_2 = \theta_p(A; G'_2) - \delta' Z'_2 \\ \dot{X}'_1 = k'Z'_2 - \gamma' X'_1 - (a'X'_1 Z_1 - d'Z'_1) \end{cases}$$

Transcriptional Activation: (Network Perturbation) $\theta_p(A; G'_2) = \left[k'_0 + (k'_2 - k'_0)\dfrac{A/\kappa_2}{1 + A/\kappa_2}\right] G'_2$

Ideal Closed Loop $(\delta \approx 0)$: **(at Steady State)** $\dot{Z}_1 + \dot{Z}'_1 - \dot{Z}_2 = 0 \implies \mu(G_1) = \theta(\bar{A}; G_2) \implies \bar{A} = \kappa_2 \dfrac{r - k_0/k_2}{1 - r}$, with $r := \dfrac{k_1 G_1}{k_2 G_2}$

Figure B.3: Mathematical Modeling of the I-Circuit with Network Perturbation in Figure 3.2. (a)/(b) Schematic/Mathematical Description of the Reduced Model. This is a special case of the compact model presented in Figure 3.4(a) where the proportional controller is removed, that is **X'$_1$** cannot be produced from **Z$_1$**. As a result, this network models the integral control action and the network perturbation introduced by the NES-L7Ae gene, denoted by **G'$_2$** which gives rise to the sub-network in purple. Here, the dimer **A** acts as a transcription factor for both genes **G$_2$** and **G'$_2$**. When **G'$_2$** is activated, it is transcribed into **Z'$_2$** at a rate $\theta_p(X_2; G'_2)$ which in turn is translated into **X'$_1$** at a rate $k'Z'_2$. Then **X'$_1$** is capable of inhibiting the translation of **Z$_1$** by binding to it. Note that the complex **Z'$_1$** formed from the binding reaction can still sequester the antisense RNA **Z$_2$**. In the ideal operation of the antithetic integral controller, where the dilution rate δ is negligible with respect to the other rates of the controller, the regulated output **A** has a steady-state concentration, denoted by $\bar{A}$, that is unaffected by the network perturbation. This ensures robust perfect adaptation of the output not only to external disturbances in the controlled network as illustrated in Figure 3.2(c), but also to network perturbations as well.

Equipped with these quasi-steady-state approximations, we can update the ODEs of Z_1 and Z_2 from (B.3) and write down the additional ODEs for Z_1', Z_2' and X_1'.

$$\begin{aligned}
\dot{Z}_1 &= \mu(G_1) - \eta Z_1 Z_2 - \delta Z_1 - (a' X'_1 Z_1 - d' Z_1') \\
\dot{Z}_2 &= \theta(A; G_2) - \eta Z_1 Z_2 - \delta Z_2 - \eta Z_2 Z_1' \\
\dot{X}_1' &= -a' X_1' Z_1 + d' Z_1' + k' Z_2' - \gamma' X_1' \\
\dot{Z}_1' &= a' X_1' Z_1 - d' Z_1' - \eta Z_1' Z_2 - \delta Z_1' \\
\dot{Z}_2' &= k_2' G'^{*}_2 + k_0' G'^{F}_2 - \delta' Z_2' \approx k_2' G_2' \left(\frac{k_0'/k_2' + A/\kappa_2}{1 + A/\kappa_2} \right) - \delta' Z_2'.
\end{aligned}$$

Finally, the dynamics of the reduced model can be written as

$$\begin{cases}
\dot{X}_1 = k Z_1 - (c+\gamma) X_1 - \lambda(X_1; D) \\
\dot{X}_2 = c X_1 - \gamma X_2 - 2(a X_2^2 - d A) \\
\dot{A} = a X_2^2 - (d+\gamma) A \\
\dot{Z}_1 = \mu(G_1) - \eta Z_1 Z_2 - \delta Z_1 - (a' X'_1 Z_1 - d' Z_1') \\
\dot{Z}_2 = \theta(A; G_2) - \eta(Z_1 + Z_1') Z_2 - \delta Z_2 \\
\dot{Z}_1' = a' X_1' Z_1 - d' Z_1' - \eta Z_1' Z_2 - \delta Z_1' \\
\dot{Z}_2' = \theta_p(A; G'_2) - \delta' Z_2'; \quad \text{with} \quad \theta_p(A; G_2) := \left[k_0' + (k_2' - k_0') \dfrac{(A/\kappa_2)}{1 + (A/\kappa_2)} \right] G_2' \\
\dot{X}_1' = k' Z_2' - \gamma' X_1' - a' X_1' Z_1 + d' Z_1',
\end{cases} \tag{B.19}$$

where $k_0' << k_2'$ (since leaky transcription is usually much slower than activated transcription) and λ, μ, θ are all functions given in (B.3). Note that by setting $G'_2 = 0$, we obtain the circuit in Figure B.1(a) where there is no network perturbation.

Model Calibration to the Experimental Data

In this section, we calibrate the model depicted in Figures B.3(a) and (b) to fit the experimentally collected data at steady state. The mathematical model of the measurement described in Supplementary Information B.1 is applicable here as well. However we add here the argument G'_2 to the measurement function $M(G_1, D, G'_2)$ to explicitly show the dependence of the measurement on the concentration of the gene **G'₂**. The experimental data are collected for both open- and closed-loops with/without disturbance and with/without network perturbation over two plasmid ratios G_1/G_2 as shown in Figure 3.4(c). The disturbance is introduced via $D = 30$ nM, while network perturbation is introduced via $G'_2 = 0.001$ pmol.

Choice of Parameter Groups

Consider the circuit depicted in Figures B.3(a) and (b). The measured output for the open- ($i = o$) and closed-loop ($i = c$) settings are given by

$$\bar{M}^i(G_1, D, G'_2) = c_x \left(\sqrt{\kappa \bar{A}} + 2\bar{A} \right), \quad \text{for } \; i \in \{o, c\}$$

$$\text{with} \; \begin{cases}
0 = k\bar{Z}_1 - (c+\gamma)\bar{X}_1 - \lambda(\bar{X}_1; D) \\
0 = c\bar{X}_1 - \gamma\bar{X}_2 - 2(a\bar{X}_2^2 - d\bar{A}) \\
0 = a\bar{X}_2^2 - (d+\gamma)\bar{A} \\
0 = \mu(G_1) - \eta\bar{Z}_1\bar{Z}_2 - \delta\bar{Z}_1 - (a'\bar{X}'_1\bar{Z}_1 - d'\bar{Z}_1') \\
0 = \theta^i(\bar{A}; G_2) - \eta(\bar{Z}_1 + \bar{Z}_1')\bar{Z}_2 - \delta\bar{Z}_2; \; \theta^i(A; G_2) = \begin{cases} k_0 G_2, & \text{for } i = o \\ \left[k_0 + (k_2 - k_0) \dfrac{A/\kappa_2}{1 + A/\kappa_2} \right] G_2, & \text{for } i = c. \end{cases} \\
0 = a'\bar{X}_1'\bar{Z}_1 - d'\bar{Z}_1' - \eta\bar{Z}_1'\bar{Z}_2 - \delta\bar{Z}_1' \\
0 = \theta_p(\bar{A}; G'_2) - \delta'\bar{Z}_2'; \qquad \theta_p(A; G_2) = \left[k_0' + (k_2' - k_0') \dfrac{A/\kappa_2}{1 + A/\kappa_2} \right] G_2' \\
0 = k'\bar{Z}_2' - \gamma'\bar{X}_1' - a'\bar{X}_1'\bar{Z}_1 + d'\bar{Z}_1',
\end{cases} \tag{B.20}$$

The system model has 7 additional parameters (compared to the circuit in Figure B.1) to be calibrated to the data: $a', d', k_0', k_1', k', \delta'$ and γ'. Once again, we carry out a suitable choice of re-parameterization to obtain a minimal number of (lumped) parameter groups that can be uniquely identified from the steady-state measurements. To specify the choice of the parameter groups, we first express all the variables as rational functions of $\bar{X}_2$. The first three and seventh equations in (B.20) can be rewritten as

$$\begin{cases} \bar{X}_1 = \frac{\gamma}{c}\left(\bar{X}_2 + 2\frac{\bar{X}_2^2}{\kappa}\right) \\ \bar{A} = \frac{\bar{X}_2^2}{\kappa}; \quad \kappa := \frac{d+\gamma}{a} \\ \bar{Z}_1 = \frac{1}{k}\left[(c+\gamma)\bar{X}_1 - \lambda(\bar{X}_1; D)\right] \\ \bar{Z}'_2 = \frac{\theta_p(\bar{A}; G'_2)}{\delta'}. \end{cases} \tag{B.21}$$

Hence, we expressed now $\bar{X}_1, \bar{A}, \bar{Z}_1$ and $\bar{Z}'_2$ as rational functions of $\bar{X}_2$. Next, we express $\bar{Z}_2$ as a rational function of $\bar{Z}_1$ and $\bar{A}$ (and thus $\bar{X}_2$). To do so we obtain the following equations

$$\bar{Z}'_1 = \bar{Z}_2 + \frac{\mu(G_1) - \theta(A; G_2)}{\delta} - \bar{Z}_1 \tag{B.22}$$

$$\bar{Z}'_1(\eta\bar{Z}_2 + \delta) = \mu - \bar{Z}_1(\eta\bar{Z}_2 + \delta) \tag{B.23}$$

$$\bar{X}'_1 = \frac{1}{\gamma'}\left[k'\bar{Z}'_2 - \bar{Z}'_1(\eta\bar{Z}_2 + \delta)\right], \tag{B.24}$$

where the first equation is obtained by subtracting the fifth equation in (B.20) from the sum of the fourth and sixth equations, the second equation is obtained by summing up the fourth and sixth equations in (B.20), and the third equation is obtained by summing up the sixth and eighth equations in (B.20). By substituting for $\bar{X}'_1$ in the sixth equation of (B.20), we obtain

$$\frac{a'k'}{\gamma'}\bar{Z}_1\bar{Z}'_2 - \left(\frac{a'}{\gamma'}\bar{Z}_1 + 1\right)\bar{Z}'_1(\eta\bar{Z}_2 + \delta) - d'\bar{Z}'_1 = 0. \tag{B.25}$$

By substituting the expressions for $\bar{Z}'_1(\eta\bar{Z}_2 + \delta), \bar{Z}'_1$ and $\bar{Z}'_2$ from (B.23), (B.22) and (B.21), respectively we obtain

$$\bar{Z}_2 = \frac{\left(\bar{Z}_1 + \frac{\gamma'}{a'} + \frac{d'}{a'}\frac{\gamma'}{\delta}\right)\left(G_1 - \frac{\delta}{k_1}\bar{Z}_1\right) - \frac{d'}{a'}\frac{\gamma'}{\delta}\frac{\theta}{k_1} - \frac{k'}{\delta'}\bar{Z}_1\frac{\theta_p}{k_1}}{\frac{\eta}{k_1}\bar{Z}_1\left(\bar{Z}_1 + \frac{\gamma'}{a'}\right) - \frac{d'}{a'}\frac{\gamma'}{k_1}}. \tag{B.26}$$

Finally, one can substitute the expressions for $\bar{Z}_1, \bar{Z}'_1$ and $\bar{Z}_2$ in the fifth equation of (B.20) to obtain a single (high order) polynomial equation in $\bar{X}_2$ solely which can be solved efficiently in Matlab using the command “root”, instead of solving the set of nonlinear algebraic equations (B.20).

Next, we rewrite the obtained equations in terms of the parameter groups and transformed variables. By recalling the transformed variable $\tilde{X}_2$ given in (B.14) and the parameter groups given in (B.13) and introducing the following five additional parameter groups

$$\bar{\gamma}' := \frac{\gamma'}{a'}, \qquad \bar{\kappa}' := \frac{d'}{a'}\frac{\gamma'}{\delta}, \qquad \bar{k}' := \frac{k'}{\delta'}, \qquad \bar{k}_0' := \frac{k_0'}{k_1}, \qquad \bar{k}_2' := \frac{k_2'}{k_1}, \tag{B.27}$$

we can rewrite $\bar{Z}_2$ as

$$\bar{Z}_2 = \frac{\left(\bar{Z}_1 + \bar{\gamma}' + \bar{\kappa}'\right)\left(G_1 - \bar{\delta}\bar{Z}_1\right) - \bar{\kappa}'\left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\frac{\tilde{X}_2^2/\bar{\kappa}_2^2}{1+\tilde{X}_2^2/\bar{\kappa}_2^2}\right]G_2 - \bar{k}'\bar{Z}_1\left[\bar{k}_0' + (\bar{k}_2' - \bar{k}_0')\frac{\tilde{X}_2^2/\bar{\kappa}_2^2}{1+\tilde{X}_2^2/\bar{\kappa}_2^2}\right]G_2'}{\bar{\eta}\bar{Z}_1\left(\bar{Z}_1 + \bar{\gamma}'\right) - \bar{\delta}\bar{\kappa}'}. \tag{B.28}$$

Therefore, the steady-state measurements can be rewritten in terms of the parameter groups and transformed variables as

$$
\bar{M}^i\,(G_1, D, G'_2) = \tilde{X}_1, \quad \text{for } i \in \{o, c\}
$$

$$
\begin{cases}
\tilde{X}_1 = \tilde{X}_2 + 2\dfrac{\tilde{X}_2^2}{\bar{\kappa}} \\
\bar{Z}_1 = \bar{\gamma}\left[\tilde{X}_1 + D\dfrac{\Delta_1\tilde{X}_1 + \Delta_2\tilde{X}_1^2}{1 + \Delta'_1\tilde{X}_1 + \Delta'_2\tilde{X}_1^2}\right] \\
\bar{Z}_2 = \dfrac{\left(\bar{Z}_1 + \bar{\gamma}' + \bar{\kappa}'\right)\left(G_1 - \bar{\delta}\bar{Z}_1\right) - \bar{\kappa}'\bar{\theta}^i(\tilde{X}_2; G_2)\ - \bar{k}'\bar{Z}_1\bar{\theta}_p(\tilde{X}_2; G'_2)}{\bar{\eta}\bar{Z}_1\left(\bar{Z}_1 + \bar{\gamma}'\right) - \bar{\delta}\bar{\kappa}'} \\
\bar{Z}'_1 = \bar{Z}_2 + \dfrac{G_1 - \bar{\theta}^i(\tilde{X}_2; G_2)}{\bar{\delta}} - \bar{Z}_1 \\
0 = \bar{\theta}^i(\tilde{X}_2; G_2) - \bar{\eta}(\bar{Z}_1 + \bar{Z}'_1)\bar{Z}_2 - \bar{\delta}\bar{Z}_2
\end{cases}
\tag{B.29}
$$

$$
\text{where} \qquad \bar{\theta}^i(\tilde{X}_2; G_2) = \begin{cases}
\bar{k}_0 G_2, & \text{for } i = o \\
\left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\dfrac{\left(\tilde{X}_2/\bar{\kappa}_2\right)^2}{1 + \left(\tilde{X}_2/\bar{\kappa}_2\right)^2}\right] G_2, & \text{for } i = c.
\end{cases}
$$

$$
\text{and} \qquad \bar{\theta}_p(\tilde{X}_2; G'_2) = \left[\bar{k}'_0 + (\bar{k}'_2 - \bar{k}'_0)\dfrac{\left(\tilde{X}_2/\bar{\kappa}_2\right)^2}{1 + \left(\tilde{X}_2/\bar{\kappa}_2\right)^2}\right] G'_2.
$$

Note that $i = o$ and $i = c$ correspond to the open- and closed-loop settings, respectively. Hence given the input G_1, the perturbation G'_2, and the disturbance D, one can use (B.29) to compute the mCitrine measurement in the open- and closed-loop settings. We close this section by observing that we have now reduced the parameters to be calibrated down to 16 (as compared to 23).

Model Calibration Steps

The model fitting is carried out in three steps to avoid over-fitting. In the first step, the model without network perturbation ($G'_2 = 0$) that is obtained in Figure B.1 via the estimated group parameters given in (B.16) is re-calibrated to the new experimental conditions such as the change in the fluorescence proportionality constant c_x. In the second step, the parameters $\bar{k}'_0, \bar{k}'_2, \bar{k}', \bar{\gamma}'$ and $\bar{\kappa}'$ are estimated using the data of the scenario with network perturbation ($G'_2 = 0.002$ pmol) but without disturbance ($D = 0$ nM). Finally, a prediction step is carried out to further assess the model fitting procedure. In this step, the estimated parameters are used to predict the scenario where both network perturbation and disturbance are applied simultaneously.

Estimated Parameter Groups: The model fit and prediction are shown in Figure 3.4(c), where the optimally estimated parameter groups are given by

$$
\begin{array}{llll}
\Delta_1 \approx 0\ \text{nM}^{-1}, & \Delta_2 = 2.76 \times 10^{-7}\ \text{a u}^{-1}\text{nM}^{-1}, & \Delta'_1 \approx 0\ \text{a u}^{-1}, & \Delta'_2 \approx 0\ \text{a u}^{-2}, \\
\bar{\delta} = 9.5 \times 10^{-3} & \bar{\gamma} = 1.35 \times 10^{-6}\ \text{a u}^{-1}, & \bar{\kappa} = 82 \times 10^{4}\ \text{a.u.}, & \bar{\kappa}_2 = 10.5 \times 10^{4}\ \text{a u}^{-1} \\
\bar{\eta} = 2.59 \times 10^{3}\ \text{nM}^{-1}, & \bar{k}_0 \approx 0, & \bar{k}_2 = 2.44, & \bar{k}'_0 \approx 0 \\
\bar{k}'_2 = 0.42, & \bar{k}' = 2.12 \times 10^{3}, & \bar{\gamma}' = 0.11\ \text{nM}, & \bar{\kappa}' = 99.3\ \text{nM}
\end{array}
\tag{B.30}
$$

It is straight forward to see that the parameter groups that are common with (B.16) are numerically very close if one takes into consideration that the fluorescence proportionality constant c_x is increased by seven to eight times.

B.3 Mathematical Modeling of the Circuit in Figure 3.3

Consider the circuit depicted in Figure 3.3(a), with a Proportional-Integral controller, that can operate in either open or closed loop. We first present a detailed (mechanistic) mathematical model and then carry out a model reduction technique that allows us to analyze the steady-state behavior. Finally we provide the technical details of fitting the model to the experimentally obtained data.

Full Model Description

A detailed biochemical reaction network that models the dynamics of the circuit in Figure 3.3(a) can be obtained by replacing the translation reaction in Table B.2 with a slightly modified version, and appending additional inhibition, degradation/dilution and sequestration reactions depicted in Tables B.7 and B.8. These species and reactions describe the biochemical reaction sub-network that provides a proportional control action via the additional protein $\mathbf{X'_1}$ which is translated from the same mRNA $\mathbf{Z_1}$ as $\mathbf{X_1}$. This protein can bind to $\mathbf{Z_1}$ to form the complex $\mathbf{Z'_1}$ to inhibit translation. However, the complex $\mathbf{Z'_1}$ can still be sequestered by the anti-sense RNA $\mathbf{Z_2}$.

Species	Symbol	Species	Symbol
NES-L7Ae	$\mathbf{X'_1}$	$\mathbf{Z_1{:}X'_1}$	$\mathbf{Z'_1}$

Table B.7: List of Additional Biochemical Species (PI-Circuit)

Reaction	Mechanism	Constants
Translation	$\mathbf{Z_1} \xrightarrow{k} \mathbf{Z_1} + \mathbf{X_1} + \mathbf{X'_1}$	
Inhibition	$\mathbf{X'_1} + \mathbf{Z_1} \underset{d'}{\overset{a'}{\rightleftharpoons}} \mathbf{Z'_1}$	
Dilution/Degradation	$\mathbf{X'_1} \xrightarrow{\gamma'} \varnothing \qquad \mathbf{Z'_1} \xrightarrow{\delta} \varnothing$	
Sequestration	$\mathbf{Z'_1} + \mathbf{Z_2} \xrightarrow{\eta} \varnothing$	

Table B.8: List of Additional Biochemical Reactions (PI-Circuit)

Model Reduction

In this section, the full model given in Tables B.2 and B.8 is mathematically reduced to the model described schematically in Figure B.4(a) and mathematically in B.4(b). The model reduction procedure is based on the reduced model that is obtained previously in Supplementary Information B.1. In fact, one can update the ODEs of Z_1 and Z_2 from (B.3) and write down the additional ODEs for X'_1 and Z'_1.

$$\begin{aligned}
\dot{X}'_1 &= kZ_1 - a'X'_1Z_1 + d'Z'_1 - \gamma'X'_1 \\
\dot{Z}_1 &= \mu(G_1) - \eta Z_1Z_2 - \delta Z_1 - a'X'_1Z_1 + d'Z'_1 \\
\dot{Z}_2 &= \theta(A; G_2) - \eta(Z_1 + Z'_1)Z_2 - \delta Z_2 \\
\dot{Z}'_1 &= a'X'_1Z_1 - d'Z'_1 - \eta Z_2Z'_1 - \delta Z'_1.
\end{aligned}$$

Finally, the dynamics of the reduced model is thus given by

$$\begin{cases}
\dot{X}_1 = kZ_1 - (c+\gamma)X_1 - \lambda(X_1; D) \\
\dot{X}_2 = cX_1 - \gamma X_2 - 2\left(aX_2^2 - dA\right) \\
\dot{A} = aX_2^2 - (d+\gamma)A \\
\dot{X}'_1 = kZ_1 - \gamma'X'_1 - (a'X'_1Z_1 - d'Z'_1) \\
\dot{Z}_1 = \mu(G_1) - \eta Z_1Z_2 - \delta Z_1 - (a'X'_1Z_1 - d'Z'_1) \\
\dot{Z}_2 = \theta(A; G_2) - \eta(Z_1 + Z'_1)Z_2 - \delta Z_2 \\
\dot{Z}'_1 = a'X'_1Z_1 - d'Z'_1 - \eta Z'_1Z_2 - \delta Z'_1,
\end{cases} \tag{B.31}$$

where λ, μ and θ are all functions given in (B.3).

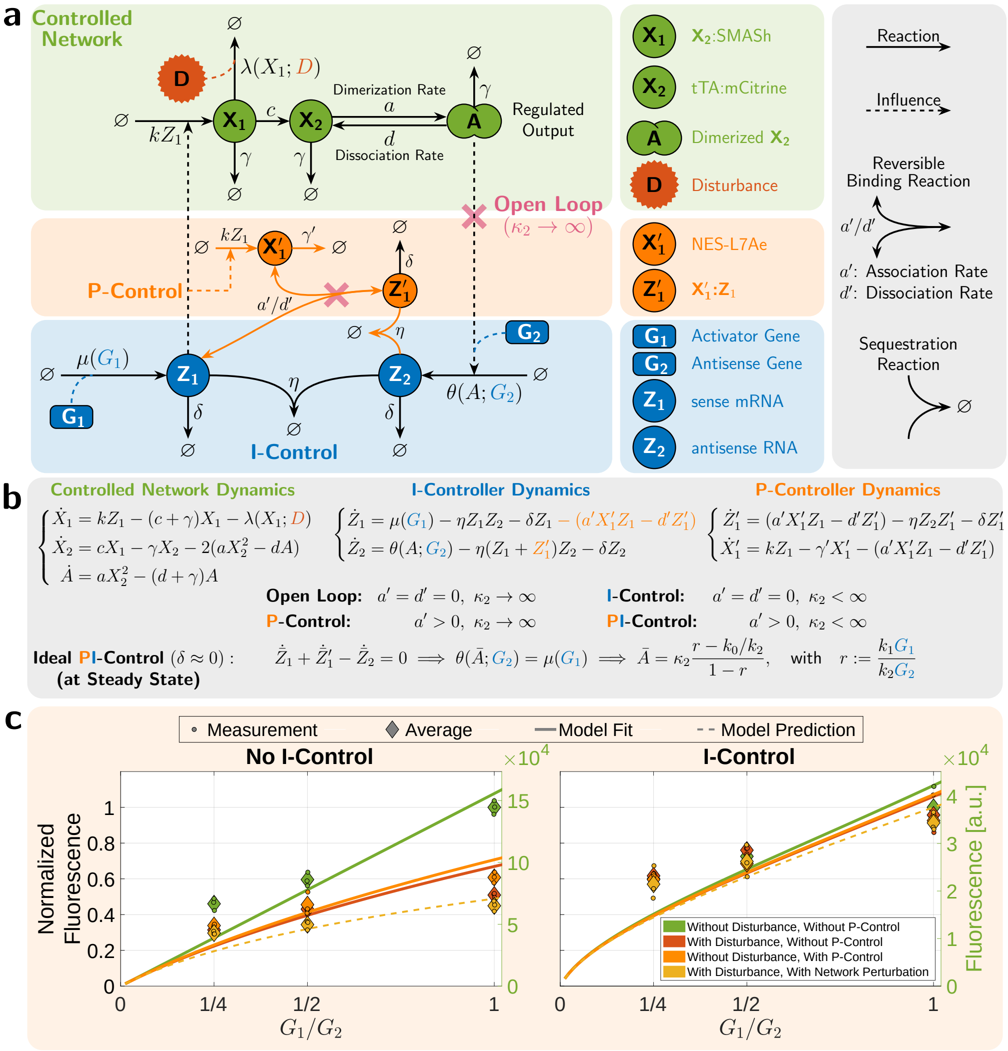


Figure B.4: Mathematical Modeling of the PI-Circuit in Figure 3.3. (a)/(b) Schematic/Mathematical Description of the Reduced Model. This is a special case of the compact model presented in Figure 3.4(a) where the network perturbation is removed, that is the additional gene **G'$_2$** is removed. As a result, this network models the Proportional-Integral (PI) control actions. The P-controller, depicted in the orange box, is realized via the production of **X'$_1$** by **Z$_1$** at the same rate k as that of **X$_1$**. This allows **X'$_1$** to act as a proxy for **X$_1$**. A negative feedback action is then achieved via the (un)binding reaction between **X'$_1$** and **Z$_1$** which inhibits the production of **X$_1$**. In the ideal operation of the controller, where the dilution rate δ is negligible with respect to the other rates of the controller, the regulated output **A** has a steady-state concentration, denoted by $\bar{A}$, that is unaffected by the P-controller. This ensures that the robust perfect adaptation of the regulated output is not influenced by the proportional controller. In fact, the P-controller has the effect of shaping the transient dynamics and reducing the steady state variance of **A** while leaving $\bar{A}$ unchanged.

Model Calibration to the Experimental Data

In this section, we calibrate the model depicted in Figures B.4(a) and (b) to fit the experimentally collected data at steady state. The mathematical model of the measurement described in Supplementary Information B.1 is applicable here as well. The experimental data are collected for both open- and closed-loops with/without disturbance and with/without Proportional P-control over three different plasmid ratios G_1/G_2 as shown in Figure 3.4(d). The disturbance is introduced via $D = 30$ nM.

Choice of Parameter Groups

Consider the circuit depicted in Figures B.4(a) and (b). The measured output for the open- ($i = o$) and closed-loop ($i = c$) settings are given by

$$\bar{M}^i(G_1, D) = c_x\left(\sqrt{\kappa\bar{A}} + 2\bar{A}\right), \quad \text{for } \ i \in \{o, c\}$$
$$\text{with} \begin{cases} 0 = k\bar{Z}_1 - (c+\gamma)\bar{X}_1 - \lambda(\bar{X}_1; D) \\ 0 = c\bar{X}_1 - \gamma\bar{X}_2 - 2(a\bar{X}_2^2 - d\bar{A}) \\ 0 = a\bar{X}_2^2 - (d+\gamma)\bar{A} \\ 0 = \mu(G_1) - \eta\bar{Z}_1\bar{Z}_2 - \delta\bar{Z}_1 - (a'\bar{X}'_1\bar{Z}_1 - d'\bar{Z}_1') \\ 0 = \theta^i(\bar{A}; G_2) - \eta(\bar{Z}_1 + \bar{Z}_1')\bar{Z}_2 - \delta\bar{Z}_2; \ \theta^i(A; G_2) = \begin{cases} k_0 G_2, & \text{for } i = o \\ \left[k_0 + (k_2 - k_0)\dfrac{A/\kappa_2}{1 + A/\kappa_2}\right] G_2, & \text{for } i = c. \end{cases} \\ 0 = a'\bar{X}_1'\bar{Z}_1 - d'\bar{Z}_1' - \eta\bar{Z}_1'\bar{Z}_2 - \delta\bar{Z}_1' \\ 0 = k\bar{Z}_1 - \gamma'\bar{X}_1' - a'\bar{X}_1'\bar{Z}_1 + d'\bar{Z}_1', \end{cases} \tag{B.32}$$

Note that by open loop (resp. closed loop), we mean the circuit without (resp. with) the integral controller. The system model has 3 additional parameters (compared to the circuit in Figure B.1) to be calibrated to the data: a', d' and γ'. Once again, we carry out a suitable choice of re-parameterization to obtain a minimal number of (lumped) parameter groups that can be uniquely identified from the steady-state measurements. To specify the choice of the parameter groups, we first express all the variables as rational functions of $\bar{X}_2$. The first three equations in (B.32) can be rewritten as

$$\begin{cases} \bar{X}_1 = \dfrac{\gamma}{c}\left(\bar{X}_2 + 2\dfrac{\bar{X}_2^2}{\kappa}\right) \\ \bar{A} = \dfrac{\bar{X}_2^2}{\kappa}; \quad \kappa := \dfrac{d+\gamma}{a} \\ \bar{Z}_1 = \dfrac{1}{k}\left[(c+\gamma)\bar{X}_1 - \lambda(\bar{X}_1; D)\right] \end{cases} \tag{B.33}$$

Hence, we expressed now $\bar{X}_1$, $\bar{A}$ and $\bar{Z}_1$ as rational functions of $\bar{X}_2$. Next, we express $\bar{Z}_2$ as a rational function of $\bar{Z}_1$ and $\bar{A}$ (and thus $\bar{X}_2$). To do so we obtain the following equations

$$\bar{Z}'_1 = \bar{Z}_2 + \frac{\mu(G_1) - \theta(A; G_2)}{\delta} - \bar{Z}_1 \tag{B.34}$$
$$\bar{Z}'_1(\eta\bar{Z}_2 + \delta) = \mu - \bar{Z}_1(\eta\bar{Z}_2 + \delta) \tag{B.35}$$
$$\bar{X}'_1 = \frac{1}{\gamma'}\left[k\bar{Z}_1 - \bar{Z}'_1(\eta\bar{Z}_2 + \delta)\right], \tag{B.36}$$

where the first equation is obtained by subtracting the fifth equation in (B.32) from the sum of the fourth and sixth equations, the second equation is obtained by summing up the fourth and sixth equations in (B.32), and the third equation is obtained by summing up the last two equations in (B.32). By substituting for $\bar{X}'_1$ in the sixth equation of (B.32), we obtain

$$\frac{a'k}{\gamma'}\bar{Z}_1^2 - \left(\frac{a'}{\gamma'}\bar{Z}_1 + 1\right)\bar{Z}'_1(\eta\bar{Z}_2 + \delta) - d'\bar{Z}'_1 = 0. \tag{B.37}$$

By substituting the expressions for $\bar{Z}'_1(\eta\bar{Z}_2 + \delta)$ and $\bar{Z}'_1$ from (B.35) and (B.34), respectively we obtain

$$\bar{Z}_2 = \frac{\left(\bar{Z}_1 + \frac{\gamma'}{a'} + \frac{d'}{a'}\frac{\gamma'}{\delta}\right)\left(G_1 - \frac{\delta}{k_1}\bar{Z}_1\right) - \frac{d'}{a'}\frac{\gamma'}{\delta}\frac{\theta}{k_1} - \frac{k}{k_1}\bar{Z}_1^2}{\frac{\eta}{k_1}\bar{Z}_1\left(\bar{Z}_1 + \frac{\gamma'}{a'}\right) - \frac{d'}{a'}\frac{\gamma'}{k_1}}. \tag{B.38}$$

Finally, one can substitute the expressions for $\bar{Z}_1, \bar{Z}'_1$ and $\bar{Z}_2$ in the fifth equation of (B.32) to obtain a single (high order) polynomial equation in $\tilde{X}_2$ solely which can be solved efficiently in Matlab using the command "root", instead of solving the set of nonlinear algebraic equations (B.32).

Next, we rewrite the obtained equations in terms of the parameter groups and transformed variables. By recalling the transformed variable $\tilde{X}_2$ given in (B.14) and the parameter groups given in (B.13) and introducing the following three additional parameter groups

$$\bar{a}' := \frac{a'}{\gamma'}, \qquad \bar{d}' := \frac{d'}{\delta}, \qquad \bar{k} := \frac{k}{k_1}, \tag{B.39}$$

we can rewrite $\bar{Z}_2$ as

$$\bar{Z}_2 = \frac{\left(\bar{Z}_1 + \frac{\bar{d}'+1}{\bar{a}'}\right)(G_1 - \bar{\delta}\bar{Z}_1) - \frac{\bar{d}'}{\bar{a}'}\left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\frac{\tilde{X}_2^2/\bar{\kappa}_2^2}{1+\tilde{X}_2^2/\bar{\kappa}_2^2}\right]G_2 \; - \bar{k}\bar{Z}_1^2}{\bar{\eta}\bar{Z}_1\left(\bar{Z}_1 + \frac{1}{\bar{a}'}\right) - \bar{\delta}\frac{\bar{d}'}{\bar{a}'}}. \tag{B.40}$$

Therefore, the steady-state measurements can be rewritten in terms of the parameter groups and transformed variables as

$$\begin{gathered}
\bar{M}^i(G_1, D) = \tilde{X}_1, \quad \text{for } i \in \{o, c\} \\
\begin{cases}
\tilde{X}_1 = \tilde{X}_2 + 2\dfrac{\tilde{X}_2^2}{\bar{\kappa}} \\
\bar{Z}_1 = \bar{\gamma}\left[\tilde{X}_1 + D\dfrac{\Delta_1\tilde{X}_1 + \Delta_2\tilde{X}_1^2}{1 + \Delta_1'\tilde{X}_1 + \Delta_2'\tilde{X}_1^2}\right] \\
\bar{Z}_2 = \dfrac{\left(\bar{Z}_1 + \frac{\bar{d}'+1}{\bar{a}'}\right)(G_1 - \bar{\delta}\bar{Z}_1) - \frac{\bar{d}'}{\bar{a}'}\left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\frac{\tilde{X}_2^2/\bar{\kappa}_2^2}{1+\tilde{X}_2^2/\bar{\kappa}_2^2}\right]G_2 \; - \bar{k}\bar{Z}_1^2}{\bar{\eta}\bar{Z}_1\left(\bar{Z}_1 + \frac{1}{\bar{a}'}\right) - \bar{\delta}\frac{\bar{d}'}{\bar{a}'}} \\
\bar{Z}'_1 = \bar{Z}_2 + \dfrac{G_1 - \bar{\theta}^i(\tilde{X}_2; G_2)}{\bar{\delta}} - \bar{Z}_1 \\
0 = \bar{\theta}^i(\tilde{X}_2; G_2) - \bar{\eta}(\bar{Z}_1 + \bar{Z}'_1)\bar{Z}_2 - \bar{\delta}\bar{Z}_2
\end{cases} \\
\text{where} \qquad \bar{\theta}^i(\tilde{X}_2; G_2) = \begin{cases}
\bar{k}_0 G_2, & \text{for } i = o \\
\left[\bar{k}_0 + (\bar{k}_2 - \bar{k}_0)\dfrac{\left(\tilde{X}_2/\bar{\kappa}_2\right)^2}{1 + \left(\tilde{X}_2/\bar{\kappa}_2\right)^2}\right]G_2, & \text{for } i = c.
\end{cases}
\end{gathered} \tag{B.41}$$

Recall that $i = o$ and $i = c$ correspond to the open- and closed-loop settings, respectively. Hence given the input G_1 and the disturbance D, one can use (B.41) to compute the mCitrine measurement in the open- and closed-loop settings. We close this section by observing that we have now reduced the parameters to be calibrated down to 14 (as compared to 19).

Model Calibration Steps

The model fitting is carried out in three steps to avoid over-fitting. In the first step, the model without a proportional controller (the production of $\mathbf{X'_1}$ by $\mathbf{Z_1}$ is removed) that is obtained in Figure B.1 via the estimated group parameters given in (B.16) is re-calibrated to the new experimental conditions such as the change in the fluorescence proportionality constant c_x. In the second step, the parameters $\bar{a}', \bar{d}', \bar{k}$ are estimated using the data of the scenario with a proportional controller but without disturbance ($D = 0$ nM). Finally, a prediction step is carried out to further assess the model fitting procedure. In this step, the estimated parameters are used to predict the scenario where both a proportional controller and disturbance are applied simultaneously.

Estimated Parameter Groups: The model fit and prediction are shown in Figure 3.4(d), where the optimally estimated parameter groups are given by

$$\begin{array}{llll}
\Delta_1 \approx 0 \text{ nM}^{-1}, & \Delta_2 = 2.1 \times 10^{-7} \text{ a u}^{-1}\text{nM}^{-1}, & \Delta_1' \approx 0 \text{ a u}^{-1}, & \Delta_2' \approx 0 \text{ a u}^{-2}, \\
\bar{\delta} = 9.5 \times 10^{-3} & \bar{\gamma} = 2.7 \times 10^{-6} \text{ a u}^{-1}, & \bar{\kappa} = 35.1 \times 10^{4} \text{ a.u.}, & \bar{\kappa}_2 = 55.3 \times 10^{4} \text{ a u}^{-1} \\
\bar{\eta} = 2.59 \times 10^{3} \text{ nM}^{-1}, & \bar{k}_0 \approx 0, & \bar{k}_2 = 2.53, & \bar{a}' = 9.1\text{nM}^{-1} \\
\bar{d}' = 903, & \bar{k} = 1.8. & &
\end{array} \tag{B.42}$$

B.4 Antithetic Proportional-Integral Control of Plasma Glucose

In this section, we provide the mathematical details underlying the simulation results given in Figure 3.6 which demonstrate the tight regulation of glucose concentration in the plasma via an antithetic proportional-integral controller.

Brief Description of the Glucose-Insulin Network to be Controlled

The mathematical whole-body model describing the dynamics of the glucose-insulin system is adopted from Dalla Man et al., 2007. The FDA-approved mathematical model is comprised of 15 species, 19 reactions and 64 parameters and is conveniently implemented as part of the SimBiology toolbox in MATLAB The MathWorks, Inc. (2021). MATLAB (Version 2021a), n.d. It captures the dynamics of glucose and insulin across the various relevant organs/tissues in the body. For a healthy subject, the model parameters take particular values; whereas for a type-II diabetic patient, certain relevant parameters are modified to reflect the lower sensitivity to insulin (see Dalla Man et al., 2007, Table I and The MathWorks, Inc. (2021). MATLAB (Version 2021a), n.d. for details). In contrast, for a type-I diabetic patient, the endogenous insulin production reactions Dalla Man et al., 2007, Equations 23-26 are removed from the model to reflect the death or inactivity of the β-cells while the parameters are kept the same as those associated with a healthy subject. The output to be controlled here is the plasma glucose concentration in mg/dL; whereas the actuated input is the total secreted quantity of insulin in the plasma in pmol.

P-Type Proportional-Integral Control Motif

In the previous circuits of Figures 3.2, 3.3 and 3.5, the controlled networks have positive gains. That is, producing more input species leads to an increase in the output species. As a result, the designed controllers for these networks implement negative feedback to ensure closed-loop stability and are hence called N-type controllers (for Negative feedback). In contrast, the glucose-insulin network to be controlled here has a negative gain since producing more insulin (input species) leads to a decrease in plasma glucose levels (output species). Subsequently, the controller for this network should implement positive feedback (P-type) to achieve overall negative feedback for the closed loop. A P-type antithetic integral controller can be achieved by switching the roles of the sense and antisense RNAs, and a P-type proportional controller can be achieved by using a promoter that is activated in the presence of glucose to drive the expression of insulin.

Next, we provide the controller differential equations that we append to the glucose-insulin model. Let G and I denote the plasma glucose concentration (output) and total plasma insulin molecules (input) in mg/dL and pmol, respectively. The controller dynamics are thus given by

$$\begin{cases} \dot{Z}_1 = \mu - \eta Z_1 Z_2 \\ \dot{Z}_2 = \theta G - \eta Z_1 Z_2 \\ \dot{Z}_3 = h(G) - \gamma Z_3; \qquad h(G) := h_{\max} \dfrac{G^n}{G^n + \kappa^n} \\ \dot{I}_c = N(kZ_2 + \alpha Z_3) - dI_c \\ \dot{I} = dI_c + \cdots, \end{cases} \tag{B.43}$$

where Z_1 and Z_2 represent the average quantities (per cell) of the anti-sense and sense RNAs, respectively, while Z_3 represents the average quantity of another mRNA that is transcribed by the gene associated with the proportional controller (see Figure 3.6(a)). Note that the average quantities are taken across N cells in pmol. Furthermore, I_c denotes the total quantity of produced insulin in all the cells before they diffuse at a rate d to the plasma. Hence the rate of total insulin secretion into the plasma given by dI_c (in pmol/h) serves as the proportional-integral actuation to the glucose-insulin system. Note that the controller reactions and parameters are appended to the SimBiology glucose-insulin model in Matlab to close the loop.

Choice of Controller Parameter Values

To enhance the performance of any controller, the control parameters has to be properly tuned. However, in practice, the various biological controller parameters cannot be freely tuned since the time scales are governed by gene expression processes. The controller parameter values appearing in the system of differential equations of the controller (B.43) are listed in Table B.9 for Type I and II diabetic subjects.

Controller Parameters	**Numerical Values**		**Units**	**Brief Description**
	Type I	**Type II**		
μ	1.66×10^{-10}	1.66×10^{-10}	pmol/h	Constitutive transcription rate
θ	$\mu/100$	$\mu/100$	pmol dL/mg/h	Transcription rate constant
η	10^{10}	10^{10}	1/pmol/h	Transcription rate
N	10^9	10^9	Dimensionless	Number of implanted controller cells
k	$10N$	$10N$	1/h	Translation rate constant
α	$700N$	$700N$	1/h	Translation rate constant
δ	$\log(2)/0.5$	$\log(2)/0.5$	1/h	Fast mRNA degradation rate constant
d	$1/0.5$	$1/0.5$	1/h	Insulin diffusion rate constant
$h_{\max}$	μ	2.5μ	pmol/h	Transcription rate
κ	200	150	mg/dL	Dissociation constant
n	4	4	Dimensionless	Hill coefficient

Table B.9: List of plausible numerical values of the various biological controller parameters for controlling both Type I and Type II diabetic patients by the Proportional-Integral controller described in (B.43). Standalone proportional control is achieved by setting $k = 0$, while standalone integral control is achieved by setting $\alpha = 0$.

Next, we provide the rationale behind picking realistic numerical values of the various biological control parameters in the case of Type I diabetic subjects. This ensures that the modeling and simulation study summarized in Figure 3.6 is numerically realistic and plausible. In the case of Type II diabetic subjects, the control parameters are similar to the Type I case, with some additional fine tuning to enhance the performance. Two important numbers are of particular interest: transcription and translation rates. It is shown in Schwanhäusser et al., 2011 that the transcription rate ranges between 0.1 and 100 mRNAs per hour with extreme cases going up to more than 500 mRNAs per hour. Hence picking a transcription rate, for the antisense RNA $\mathbf{Z_1}$, of 100 mRNAs per hour allows us to set μ to

$$\mu = 100 \times \frac{10^{12}}{N_A} \approx 1.66 \times 10^{-10}\text{pmol/h},$$

where N_A is Avogadro's number. Since $h(G)$ is also a transcription rate whose maximum value (as $G \to \infty$) is $h_{\max}$, then we also set

$$h_{\max} = \mu = 1.66 \times 10^{-10}\text{pmol/h}.$$

The desired set-point of the controlled glucose levels in the plasma is chosen here to be 100mg/dL.

Hence for the proportional-integral controller to achieve this set-point given by μ/θ, we set θ to be

$$\theta = \mu/100 \approx 1.66 \times 10^{-12}\text{pmol dL/mg/h}.$$

Observe that for this value of θ, the transcription rate of the sense-mRNA (around the set-point) is given by μ which is already selected to be plausible. Next, it is also shown in Schwanhäusser et al., 2011 that the translation rate ranges between 1 and 1000 proteins per mRNA per hour. This allows us to set k and α to plausible values between 1 and 1000 h^{-1} multiplied by the number of implanted cells N which is taken to be 1 billion — equal to the number of β-cells in the human pancreas Scharfmann et al., 2019. The sequestration rate η is chosen to be large enough to reflect a fast sense/antisense RNA hybridization. Furthermore, κ is selected to tune the threshold of the Hill function with high sensitivity (Hill coefficient $n = 4$).

Ideally, a proportional controller exhibits an instantaneous feedback from the output into the input. Here, in a more practical setting, the proportional controller can be realized via gene expression (see Figure 3.6(a)) where the degradation rate of the associated mRNA $\mathbf{Z_3}$ is desired to be fast to mimic the ideal instantaneous proportional control action. Consequently, the degradation rate of the mRNA $\mathbf{Z_3}$ is picked to be fast enough, that is $\log(2)/0.5\text{h}^{-1}$ to reflect a relatively short half-life of 0.5 h. Finally, a conversion reaction $\mathbf{I_c} \xrightarrow{d} \mathbf{I}$ is introduced to reflect a delay caused by the diffusion of the insulin from the cells to the plasma at a rate of $d = 1/0.5\text{h}^{-1}$. This reflects a 30 min average time to secrete 1000 newly produced insulin molecules into the plasma.

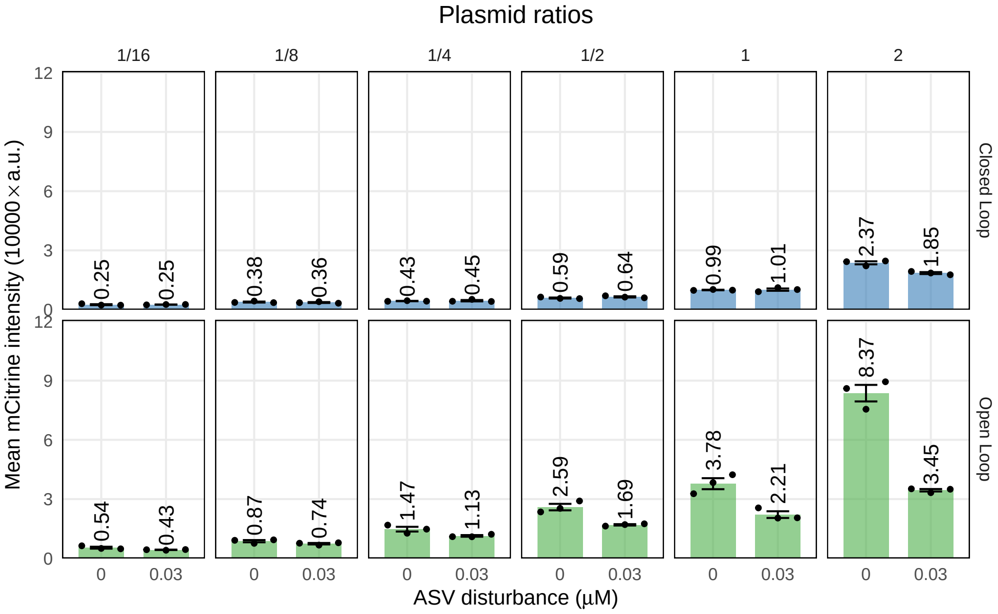


Figure B.5: Fluorescence values shown in Figure 3.2(c) in arbitrary fluorescence units. The genetic implementations of the open- and closed-loop circuits as shown in Figure 3.2(a) were transiently transfected in six different molar ratios (setpoint := activator/antisense) and perturbed with 30 nM ASV. The data was collected 48 hours after transfection and is plotted as mean fluorescence intensity ± standard error for N = 3 replicates. The data is provided in a separate file.

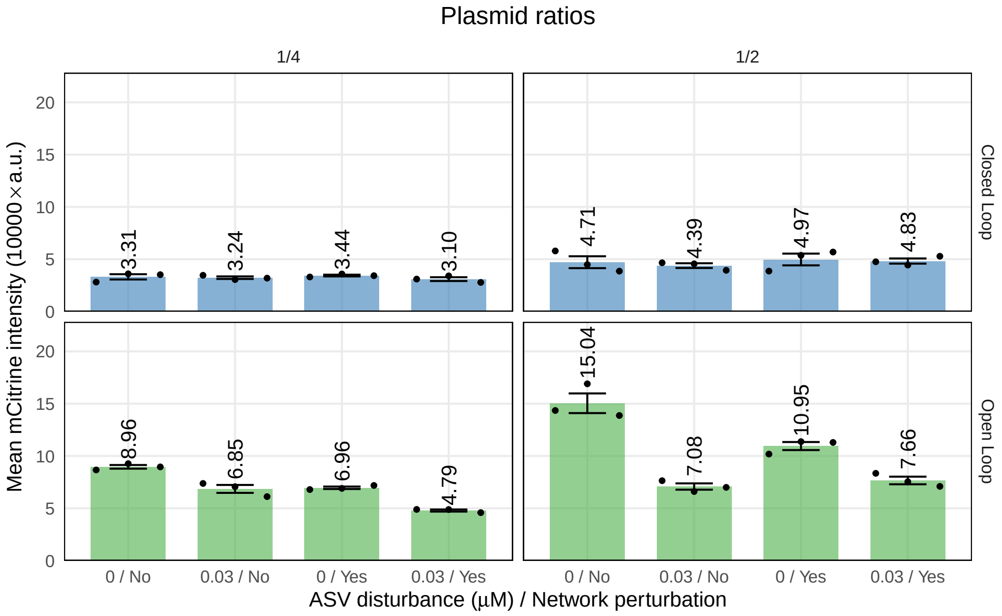


Figure B.6: Fluorescence values shown in Figure 3.2(d) in arbitrary fluorescence units. The closed- and open-loop circuits were perturbed by an additional negative feedback loop from L7Ae and by adding 30 nM ASV, as shown in Figure 3.2(a). This was done for two setpoints 1/4 and 1/2 (setpoint :=activator/antisense). The data was collected 48 hours after transfection and is plotted as mean fluorescence intensity ± standard error for N = 3 replicates. The data is provided in a separate file.

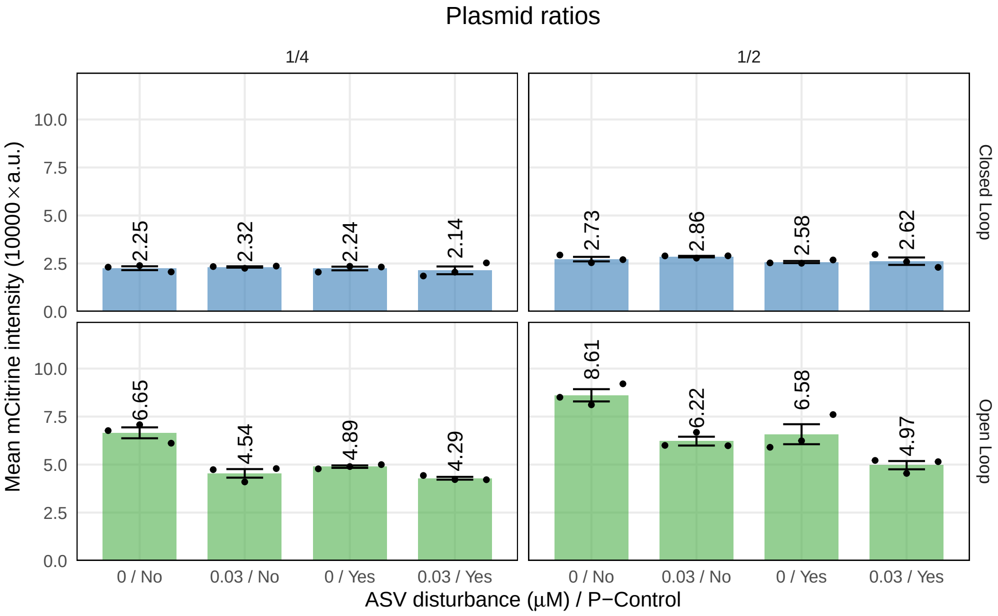


Figure B.7: Fluorescence values shown in Figure 3.3(b) in arbitrary fluorescence units. The P and PI circuits were implemented by adding a negative feedback loop from L7Ae to the open- and closed-loop circuits, as shown in Figure 3.3(a). All circuits were perturbed by adding 30 nM of ASV. This was done for two setpoints 1/4 and 1/2 (setpoint :=activator/antisense). The data was collected 48 hours after transfection and is plotted as mean fluorescence intensity ± standard error for N = 3 replicates. The data is provided in a separate file.

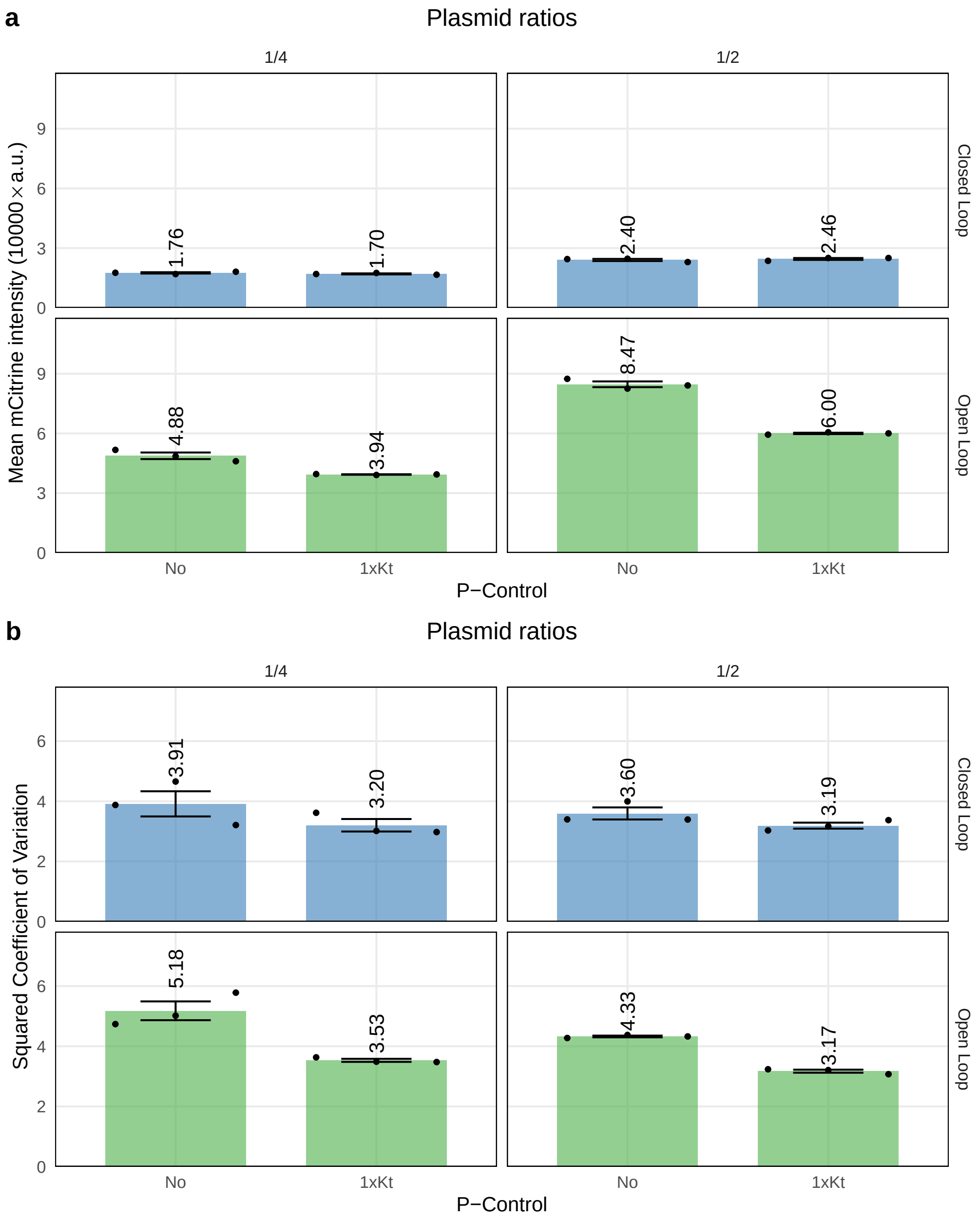


Figure B.8: Fluorescence values shown in Figure 3.3(c) in arbitrary fluorescence units. This experiment was performed in a 24-well plate rather than a 96-well plate because a large sample size is required to estimate the steady-state variance accurately. The P and PI circuits were implemented by adding a negative feedback loop from L7Ae to the open- and closed-loop circuits, as shown in Figure 3.3(a). This was done for two setpoints 1/4 and 1/2 (setpoint :=activator/antisense). (a) Expression levels of tTA-mCitrine-SMASh are plotted as mean fluorescence intensity ± standard error for N = 3 replicates. (b) The coefficient of variation squared is shown as the mean ± standard error for N = 3 replicates per condition. The data is provided in a separate file.

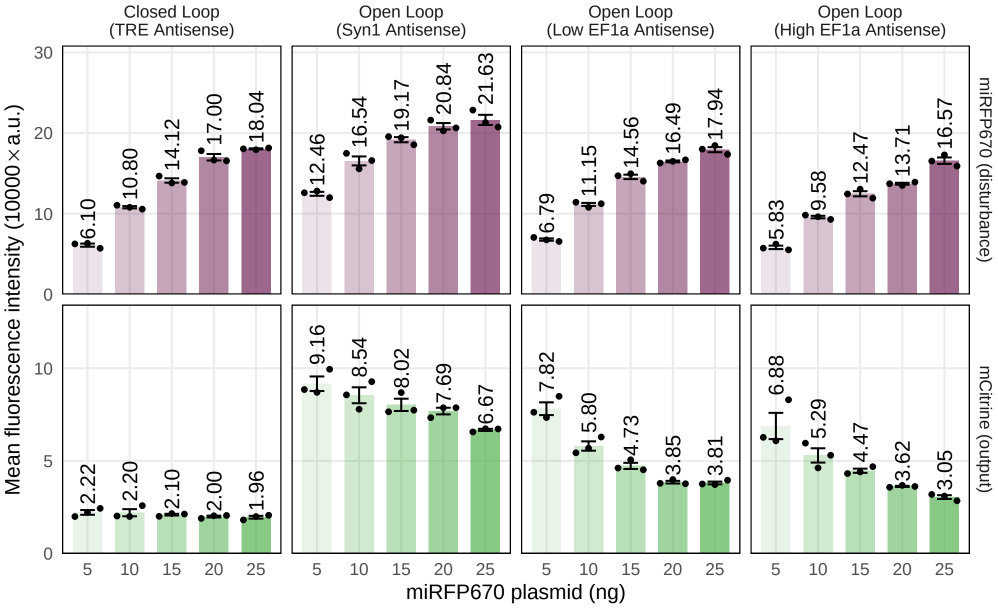


Figure B.9: Fluorescence values of different implementations of the open-loop circuit. Besides the open- and closed-loop circuit shown in Figure 3.2a, another open-loop implementation, in which the antisense RNA is expressed by a strong constitutive EF1*α* promoter, is shown. The amount of activator plasmid was fixed among different circuits, and therefore, different levels of dsRNA formation were achieved by different amounts of antisense plasmid. All of the conditions were co-transfected with different amounts of an additional disturbance plasmid that constitutively expresses the fluorescent protein miRFP670. Expression of miRFP670 not only reflects the potential effect of dsRNA formation on gene expression but also introduces a disturbance to the amount of available resources which indirectly affects the expression levels of tTA-mCitrine-SMASh. The data was collected 48 hours after transfection and is plotted as mean fluorescence intensity ± standard error for N = 3 replicates. The data is provided in a separate file.

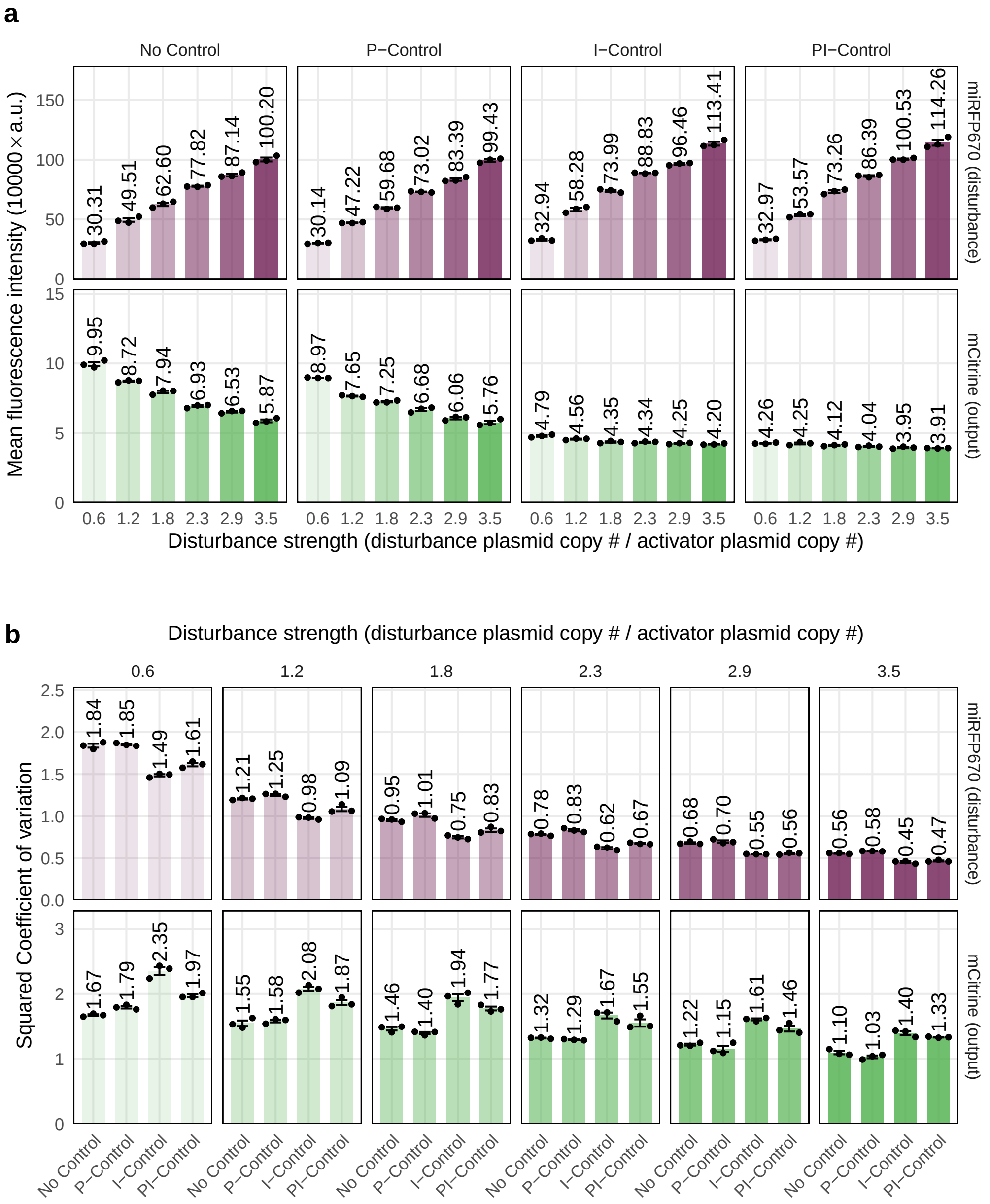


Figure B.10: See caption on the next page.

Figure B.10: Fluorescence values shown in Figure 3.5 in arbitrary fluorescence units. This experiment was performed in the 24-well plate rather than 96-well plate because a large sample size is required to estimate the steady-state variance accurately. The open-loop (No Control), proportional feedback (P-Control), antithetic integral feedback (I-Control) and proportional-integral feedback (PI-Control) circuits were perturbed by co-transfecting different amounts of an additional disturbance plasmid that constitutively expresses the fluorescent protein miRFP670, as shown in Figure 3.5(a). The increase of miRFP670 expression introduces a disturbance to the amount of available resources which indirectly affects the expression levels of tTA-mCitrine-SMASh. The activator plasmid and antisense plasmid for all controllers were transiently transfected at a setpoint ratio of 1/2 together with disturbance strengths varying from 0.6 to 3.5. The data was collected 48 hours after transfection. (a) Expression levels of tTA-mCitrine-SMASh and miRFP670 are plotted as mean fluorescence intensity ± standard error for N = 3 replicates. (b) The coefficient of variation squared is shown as the mean ± standard error for N= 3 replicates per condition. The data is provided in a separate file.

Table B.10: Transfection table regarding data shown in Figure 3.2.

Open Loop (OL)

100 ng total	pCH127	pCH124	pGLM171
Plasmid ratio = 2	31.08 ng	14.26 ng	54.66 ng
Plasmid ratio = 1	15.54 ng	14.26 ng	70.20 ng
Plasmid ratio = 1/2	7.77 ng	14.26 ng	77.97 ng
Plasmid ratio = 1/4	3.89 ng	14.26 ng	81.86 ng
Plasmid ratio= 1/8	1.94 ng	14.26 ng	83.80 ng
Plasmid ratio = 1/16	0.97 ng	14.26 ng	84.77 ng
Optimen	to 10 uL		
PEI	0.3 uL		
HEK293T	12000		

Closed Loop (CL)

100 ng total	pCH127	pCH122	pGLM171
Plasmid ratio = 2	31.08 ng	14.84 ng	54.08 ng
Plasmid ratio = 1	15.54 ng	14.84 ng	69.62 ng
Plasmid ratio = 1/2	7.77 ng	14.84 ng	77.39 ng
Plasmid ratio = 1/4	3.89 ng	14.84 ng	81.28 ng
Plasmid ratio = 1/8	1.94 ng	14.84 ng	83.22 ng
Plasmid ratio = 1/16	0.97 ng	14.84 ng	84.19 ng
Optimen	to 10 uL		
PEI	0.3 uL		
HEK293T	12000		

Table B.11: Transfection table regarding data shown in Figure 3.2(d). Open Loop (OL), Closed Loop (CL), With/Without Network Perturbation (NP).

Plasmid ratio = 1/2

100 ng total	pCH310	pCH124	pCH122	pCH135	pCH323	pGLM171
OL w/o NP	7.80 ng	14.26 ng	0 ng	0 ng	2.26 ng	75.68 ng
CL w/o NP	7.80 ng	0 ng	14.84 ng	0 ng	2.26 ng	75.1 ng
OL w/ NP	7.80 ng	14.26 ng	0 ng	2.41 ng	0 ng	75.53 ng
CL w/ NP	7.80 ng	0 ng	14.84 ng	2.41 ng	0 ng	74.95 ng
Optimen	to 10 uL					
PEI	0.3 uL					
HEK293T	12000					

Plasmid ratio = 1/4

100 ng total	pCH310	pCH124	pCH122	pCH135	pCH323	pGLM171
OL w/o NP	3.90 ng	14.26 ng	0 ng	0 ng	2.26 ng	79.58 ng
CL w/o NP	3.90 ng	0 ng	14.84 ng	0 ng	2.26 ng	79.43 ng
OL w/ NP	3.90 ng	14.26 ng	0 ng	2.41 ng	0 ng	79.00 ng
CL w/ NP	3.90 ng	0 ng	14.84 ng	2.41 ng	0 ng	78.85 ng
Optimen	to 10 uL					
PEI	0.3 uL					
HEK293T	12000					

Table B.12: Transfection table regarding data shown in Figure 3.3(b). Open Loop (OL), Closed Loop (CL), With/Without Proportional Control (PC).

Plasmid ratio = 1/2

100 ng total	pCH184	pCH183	pCH192	pCH191	pGLM171
OL w/o NP	8.24 ng	0 ng	15.26 ng	0 ng	76.49 ng
CL w/o NP	8.24 ng	0 ng	0 ng	15.84 ng	75.91 ng
OL w/ NP	0 ng	8.24 ng	15.26 ng	0 ng	76.49 ng
CL w/ NP	0 ng	8.24 ng	0 ng	15.84 ng	75.91 ng
Optimen	to 10 uL				
PEI	0.3 uL				
HEK293T	12000				

Plasmid ratio = 1/4

100 ng total	pCH184	pCH183	pCH192	pCH191	pGLM171
OL w/o NP	4.12 ng	0 ng	15.26 ng	0 ng	80.62 ng
CL w/o NP	4.12 ng	0 ng	0 ng	15.84 ng	80.04 ng
OL w/ NP	0 ng	4.12 ng	15.26 ng	0 ng	80.62 ng
CL w/ NP	0 ng	4.12 ng	0 ng	15.84 ng	80.04 ng
Optimen	to 10 uL				
PEI	0.3 uL				
HEK293T	12000				

Table B.13: Transfection table regarding data shown in Figure 3.3(c). Open Loop (OL), Closed Loop (CL), With/Without Proportional Control (PC).

Plasmid ratio = 1/2

500 ng total	pCH184	pCH183	pCH192	pCH191	pGLM171
OL w/o NP	41.21 ng	0 ng	76.32 ng	0 ng	382.47 ng
CL w/o NP	41.21 ng	0 ng	0 ng	79.22 ng	379.57 ng
OL w/ NP	0 ng	41.21 ng	76.32 ng	0 ng	382.47 ng
CL w/ NP	0 ng	41.21 ng	0 ng	79.22 ng	379.57 ng
Optimen	to 50 uL				
PEI	1.5 uL				
HEK293T	75000				

Plasmid ratio = 1/4

100 ng total	pCH184	pCH183	pCH192	pCH191	pGLM171
OL w/o NP	20.61 ng	0 ng	76.32 ng	0 ng	403.08 ng
CL w/o NP	20.61 ng	0 ng	0 ng	79.22 ng	400.18 ng
OL w/ NP	0 ng	20.61 ng	76.32 ng	0 ng	403.08 ng
CL w/ NP	0 ng	20.61 ng	0 ng	79.22 ng	400.18 ng
Optimen	to 50 uL				
PEI	1.5 uL				
HEK293T	75000				

Table B.14: Transfection table regarding data shown in Figure B.9.

Closed Loop (TRE Antisense)

100 ng total	pCH127	pCH122	pTTF138	pGLM171
DS = 1	7.77 ng	14.84 ng	10 ng	67.39 ng
DS = 2	7.77 ng	14.84 ng	20 ng	57.39 ng
DS = 3	7.77 ng	14.84 ng	30 ng	47.39 ng
DS = 4	7.77 ng	14.84 ng	40 ng	37.39 ng
DS = 5	7.77 ng	14.84 ng	50 ng	27.39 ng
Optimen	to 10 uL			
PEI	0.3 uL			
HEK293T	12000			

Open Loop (Syn1 Antisense)

100 ng total	pCH127	pCH124	pTTF138	pGLM171
DS = 1	7.77 ng	14.26 ng	10 ng	67.97 ng
DS = 2	7.77 ng	14.26 ng	20 ng	57.97 ng
DS = 3	7.77 ng	14.26 ng	30 ng	47.97 ng
DS = 4	7.77 ng	14.26 ng	40 ng	37.97 ng
DS = 5	7.77 ng	14.26 ng	50 ng	27.97 ng
Optimen	to 10 uL			
PEI	0.3 uL			
HEK293T	12000			

Open Loop (Low EF1α Antisense)

100 ng total	pCH127	pCH218	pTTF138	pGLM171
DS = 1	7.77 ng	1.94 ng	10 ng	80.29 ng
DS = 2	7.77 ng	1.94 ng	20 ng	70.29 ng
DS = 3	7.77 ng	1.94 ng	30 ng	60.29 ng
DS = 4	7.77 ng	1.94 ng	40 ng	50.29 ng
DS = 5	7.77 ng	1.94 ng	50 ng	40.29 ng
Optimen	to 10 uL			
PEI	0.3 uL			
HEK293T	12000			

Open Loop (High EF1α Antisense)

100 ng total	pCH127	pCH218	pTTF138	pGLM171
DS = 1	7.77 ng	3.88 ng	10 ng	78.35 ng
DS = 2	7.77 ng	3.88 ng	20 ng	68.35 ng
DS = 3	7.77 ng	3.88 ng	30 ng	58.35 ng
DS = 4	7.77 ng	3.88 ng	40 ng	48.35 ng
DS = 5	7.77 ng	3.88 ng	50 ng	38.35 ng
Optimen	to 10 uL			
PEI	0.3 uL			
HEK293T	12000			

Table B.15: Transfection table regarding data shown in Figure 3.5. No Control (NC), P-Control (P), I-Control (I), PI-Control (PI), Disturbance Strength (DS).

Plasmid ratio = 1/2

500 ng total	pCH184	pCH183	pCH218	pCH191	pTTF138	pGLM171
NC w/ DS = 0.6	43 ng	0 ng	10.13 ng	0 ng	20.02 ng	426.85 ng
NC w/ DS = 1.2	43 ng	0 ng	10.13 ng	0 ng	40.04 ng	406.83 ng
NC w/ DS = 1.8	43 ng	0 ng	10.13 ng	0 ng	60.06 ng	386.81 ng
NC w/ DS = 2.3	43 ng	0 ng	10.13 ng	0 ng	80.09 ng	366.78 ng
NC w/ DS = 2.9	43 ng	0 ng	10.13 ng	0 ng	100.11 ng	346.76 ng
NC w/ DS = 3.5	43 ng	0 ng	10.13 ng	0 ng	120.13 ng	326.74 ng
P w/ DS = 0.6	0 ng	43 ng	10.13 ng	0 ng	20.02 ng	426.85 ng
P w/ DS = 1.2	0 ng	43 ng	10.13 ng	0 ng	40.04 ng	406.83 ng
P w/ DS = 1.8	0 ng	43 ng	10.13 ng	0 ng	60.06 ng	386.81 ng
P w/ DS = 2.3	0 ng	43 ng	10.13 ng	0 ng	80.09 ng	366.78 ng
P w/ DS = 2.9	0 ng	43 ng	10.13 ng	0 ng	100.11 ng	346.76 ng
P w/ DS = 3.5	0 ng	43 ng	10.13 ng	0 ng	120.13 ng	326.74 ng
I w/ DS = 0.6	43 ng	0 ng	0 ng	82.52 ng	20.02 ng	354.46 ng
I w/ DS = 1.2	43 ng	0 ng	0 ng	82.52 ng	40.04 ng	334.44 ng
I w/ DS = 1.8	43 ng	0 ng	0 ng	82.52 ng	60.06 ng	314.42 ng
I w/ DS = 2.3	43 ng	0 ng	0 ng	82.52 ng	80.09 ng	294.39 ng
I w/ DS = 2.9	43 ng	0 ng	0 ng	82.52 ng	100.11 ng	274.37 ng
I w/ DS = 3.5	43 ng	0 ng	0 ng	82.52 ng	120.13 ng	254.35 ng
PI w/ DS = 0.6	0 ng	43 ng	0 ng	82.52 ng	20.02 ng	354.46 ng
PI w/ DS = 1.2	0 ng	43 ng	0 ng	82.52 ng	40.04 ng	334.44 ng
PI w/ DS = 1.8	0 ng	43 ng	0 ng	82.52 ng	60.06 ng	314.42 ng
PI w/ DS = 2.3	0 ng	43 ng	0 ng	82.52 ng	80.09 ng	294.39 ng
PI w/ DS = 2.9	0 ng	43 ng	0 ng	82.52 ng	100.11 ng	274.37 ng
PI w/ DS = 3.5	0 ng	43 ng	0 ng	82.52 ng	120.13 ng	254.35 ng
Optimen	to 50 uL					
PEI	1.5 uL					
HEK293T	150000					

Table B.16: List of the plasmids used in this study. Plasmid sequences enclosed in separate file.

Fig.	Short name	Description
2, 3, 4, 6, s5	pGLM171	AmpR-INS-bGHpA-SV40pA-INS-pUCori
6, s5	pTTF138	INS-bGHpA-PEF1a-miRFP670-SV40pA-INS
2, 3, s5	pCH122	ConLS-pTRE-antisense-tTA-mCitrine-SMASh(Ai)-FF5-SV40p(A)-INS-ConR1
2, 3, s5	pCH124	ConLS-pSyn1-antisense-tTA-mCitrine-SMASh(Ai)-FF5-SV40p(A)-INS-ConR1
2,s5	pCH127	ConL1-pEF1a-tTA-mCitrine-SMASh(Ai)-FF5-SV40p(A)-ConRE
3	pCH135	ConL2-pTRE-NES-L7AE-SV40p(A)-INS-ConR1
4b, 4c, 6	pCH183	ConL1-pEF1a-KT hairpin-tTA-mCitrine-SMASh(Ai)-P2A-L7Ae-SV40p(A)-ConRE
4b, 4c, 6	pCH184	ConL1-pEF1a-KT hairpin(mut)-tTA-mCitrine-SMASh(Ai)-P2A-L7Ae-SV40p(A)-ConRE
4b, 4c, 6	pCH191	ConLS-pTRE-antisense-tTA-mCitrine-SMASh(Ai)-P2A-L7Ae-SV40p(A)-INS-ConR1
4b, 4c	pCH192	ConLS-pSyn1-antisense-tTA-mCitrine-SMASh(Ai)-P2A-L7Ae-SV40p(A)-INS-ConR1
6, s5	pCH218	ConL2-pEF1a-anti-sense-tTA-mCitrine-SMASh(Ai)-FF5-SV40p(A)-ConRE
3	pCH310	ConL1-pEF1a-KT hairpin-tTA-mCitrine-SMASh(Ai)-FF5-SV40p(A)-ConRE
3	pCH323	ConL2-pSyn1-NES-L7AE-SV40p(A)-INS-ConR1

B.5 Bibliography

Dalla Man, C., Rizza, R. A., & Cobelli, C. (2007). Meal simulation model of the glucose-insulin system. *IEEE Transactions on biomedical engineering*, *54*(10), 1740–1749.

The MathWorks, Inc. (2021). MATLAB (Version 2021a). (n.d.). Simulating the glucose-insulin response.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, *473*(7347), 337–342.

Scharfmann, R., Staels, W., Albagli, O., et al. (2019). The supply chain of human pancreatic β cell lines. *The Journal of clinical investigation*, *129*(9), 3511–3520.