
Peer reviewed version

Link to published version (if available):
10.1021/acssynbio.5b00141

Link to publication record in Explore Bristol Research
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via American Chemical Society at http://dx.doi.org/10.1021/acssynbio.5b00141. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research
General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
Automatic control of gene expression in mammalian cells.

Chiara Fracassi,† Lorena Postiglione,†,‡ Gianfranco Fiore,† and Diego di Bernardo*,†,‡

TeleThon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Italy, and Department of Electrical Engineering and Information Technology, University of Naples Federico II, Naples, Italy

E-mail: dibernardo@tigem.it

Abstract

Automatic control of gene expression in living cells is paramount importance to characterise both endogenous gene regulatory networks and synthetic circuits. In addition, such a technology can be used to maintain the expression of synthetic circuit’s components in an optimal range, in order to ensure reliable performances. Here we present a microfluidics-based method to automatically control gene expression from the Tetracycline-inducible promoter in mammalian cells in real-time. Our approach is based on the negative-feedback control engineering paradigm. We validated our method in a monoclonal population of cells constitutively expressing a fluorescent reporter protein (d2EYFP), downstream of a minimal CMV promoter with seven tet-responsive operator motifs (CMV-TET). These cells also constitutively express the Tetracycline-transactivator protein (tTA). In cells grown in standard growth medium, tTA is able
to bind the $CMV\cdot TET$ promoter causing d2EYFP to be maximally expressed. Upon addition of Tetracycline to the culture medium, tTA detaches from the $CMV\cdot TET$ promoter thus preventing d2EYFP expression. We tested two different model-independent control algorithms (relay and PI) to force a monoclonal population of cells to express an intermediate level of d2EYFP equal to 50% of its maximum expression level, for up to 3,500 minutes. The control input is either Tetracycline-rich or standard growing medium. We demonstrated that both the relay and the PI controller can regulate gene expression at the desired level, despite oscillations (dampened in the case of the PI) around the chosen set-point.

**Keywords**

gene expression, synthetic biology, control engineering, microfluidics, mammalian cell

Biological systems control their functions in order to maintain homeostasis and respond to endogenous and exogenous cues. In 1954, Fred Grodins and colleagues were among the first to describe a biological function, i.e. the regulation of CO$_2$ by the respiratory system, in terms of a feedback-control system, by drawing an analogy to an electric circuit (1). The study of Molecular Biology from a Control Theory perspective has been, however, mostly theoretical for the last 60 years, mainly because of a lack of quantitative measurements of molecular processes. The advent of new technologies to quantitatively measure molecular species and manipulate cells has now enabled practical applications of control engineering to living systems.

Quantitative regulation of gene expression from inducible promoters in yeast cells, using feedback-control strategies, has been recently demonstrated (2–4). The aim is to force a population of cells, or even a single cell, to express a predetermined amount of a protein of
interest, which can be either constant or time-varying, by automatically administering an inducer molecule, light stimulus or environmental perturbation, to the cells harbouring an inducible promoter.

In 2011, a semi-automated control strategy was implemented to regulate gene expression in yeast from the \textit{GAL1} promoter, using a synthetic light-inducible transcription factor\cite{2}. A more advanced fully automated strategy was described in 2012, to control the expression of a reporter protein from the Hog1-responsive promoter using a microfluidics platform \cite{3} with osmotic pressure as control input. We reported the first control strategy to regulate gene expression from the \textit{GAL1} promoter using galactose and glucose as control inputs\cite{4} by means of an ad-hoc microfluidic platform.

The application of control engineering to mammalian systems is still at its infancy. A pioneering work was performed in 2011 by Toettcher and colleagues \cite{5}, where they described an optogenetic approach to control the post-translational activation of the PI3K cascade by automatically adjusting the light input to single cells in order to buffer for cell-to-cell variability. However, to our knowledge, control of gene expression from inducible promoters in mammalian cells has not been yet achieved.

Transcription in mammalian cells is up to one order of magnitude slower than in bacteria or simple eukaryotes, therefore the time-scale involved in the control of gene expression are much longer (days instead of hours) \cite{6}. Moreover, mammalian cells have much more stringent requirements for culture conditions, such as temperature, pH, \text{CO}_2 concentration, higher sensitivity to photo-toxicity, which have to be taken into account in order to successfully maintain a cell culture in a microfluidics chamber over several days under a time-lapse microscope. These needs are a distinctive feature of mammalian cells that have so far hindered the application of automatic control to mammalian systems.

In this work, we present a microfluidics-based feedback control strategy to quantitatively regulate gene expression from a Tetracycline-inducible promoter in mammalian cells. We show that it is indeed possible to force a population of cells, harbouring the inducible
promoter, to express a predetermined level of a protein of interest by automatically administering pulses of Tetracycline to the cells, whose duration is computed in real-time by a control algorithm. The ability to express a protein of interest at different levels, or in a time-varying fashion, from the same promoter would be a unique tool for several applications, including studying the effects of gene dosage in disease, probing the function of endogenous regulatory networks, and for synthetic biology applications.

Results and Discussion

Implementation of the control platform.

The control platform, described in Figure 1a, was set up to monitor in real-time fluorescent protein levels and to provide time-varying concentrations of an inducer molecule. A microfluidics device (7) was used to trap cells in a micro-metric chamber. Either culture medium or an inducer molecule can be supplied to the chamber by changing the relative heights between two syringes that are connected to it via capillary tubes, one containing just culture medium and the other culture medium plus the inducer molecule. The fluorescence produced by cells \( y \) is imaged at regular sampling times of 15 min by the microscope, and converted to a quantitative value \( y_m \) via an image segmentation algorithm, which computes the average fluorescence intensity across the cells in the image (8). Indeed, individual cells may express different levels of fluorescence due to biological noise, thus the choice of averaging fluorescence across the cells reduces the effect of intrinsic and extrinsic noise sources on the quantity to be controlled \( y_m \).

The measured fluorescence \( y_m \) is then compared to the desired fluorescence value \( r \), and a control algorithm (controller) computes the necessary amount of inducer molecule \( u \) to be administered to the cells in order to minimise the control error \( e = r - y_m \), and then moves the syringes accordingly.

In principle, the difference in hydrostatic pressure between the two syringes can be used
to provide any desired concentration of the inducer molecule by mixing the two fluids by properly adjusting the heights of the two syringes. However, small errors in positioning the syringes, which may be caused by the mechanical inaccuracies of motors, pulleys and belts supporting them, may accumulate over time leading to large discrepancies between the desired concentration of the molecule in the microfluidic cell chamber and its actual value. To overcome this problem, electrical motors with higher mechanical precision should be employed. Therefore, in order to minimise the possibility of technical errors, we decided to use a binary control input, that is, either no inducer is provided (i.e. just culture medium) or a single predetermined concentration of the inducer molecule.

The experimental model.

As a test-bed for assessing the feasibility of automatically controlling gene expression in mammalian cells, we chose the Tet-OFF system driving the expression of a destabilised yellow fluorescent reporter protein (d2EYFP) (9) in a monoclonal population of cells, as shown in Figure 2a. In the Tet-OFF system, d2EYFP is expressed under the control of the CMT-TET promoter(10, 11) composed of a minimal CMV promoter with seven tet-responsive operator motifs (tetO7), embedded in Chinese Hamster Ovary cells, constitutively expressing the Tetracycline-transactivator (tTA) protein (Figure 2). The CMV-TET promoter is the most commonly used and commercially available Tetracycline-responsive promoter engineered by manufacturers to have a high transcriptional activity in mammalian cells.

In cells grown in standard growth medium, tTA protein is able to bind the CMV-TET promoter causing d2EYFP to be maximally expressed. Upon addition of Tetracycline, or Doxycycline, to the culture medium, tTA detaches from the CMV-TET promoter thus preventing the expression of d2EYFP. This reporter protein has an half-life of 2 hours (8), which represents a good trade-off between quantum yield and rapid turnover. A reporter with a longer half-life could not have captured the dynamics of transcription, which occurs in a few
hours, while a short-lived reporter would not have yielded a robust fluorescence signal.

In the untreated population, fluorescence is present in all cells at detectable levels (Figure 2b). The growth rate in the microfluidic device is similar to the one obtained in standard culture dishes, and the ability of cells to express the fluorescent reporter is not affected under perfused conditions (Figure S1).

The response of this monoclonal population to varying concentrations of Doxycycline (from 1 ng/ml to 1 µg/ml) has been previously characterised (8). However, in our preliminary experiments we observed an unexpected effect of Doxycycline, i.e. the inability of cells to re-express fluorescence after Doxycycline washout (see Figure S2 and Supplementary Information for a discussion). This undesired effect was greatly mitigated by the use of Tetracycline, therefore we decided to use 100 ng/ml Tetracycline for our control experiments.

In order to assess whether automatic control of gene expression in mammalian cells was feasible, we set as control task the set-point regulation depicted in Figure 1d: after an initial calibration phase of 180 min, which is needed to quantify the initial level of fluorescence, cells are required to express 50% of their initial fluorescence value for up to 3,500 min.

**Control strategies: Relay and PI controllers**

The choice of a control algorithm is a trade-off between robustness and accuracy: the first is the ability to control different systems, albeit with a suboptimal performance; the latter minimises the control error ($e$). Obtaining an optimal accuracy is usually possible only with model-based control strategies. These approaches, however, require the identification of a mathematical model of the biological process being controlled, which may not always be available or readily identifiable.

We decided, therefore, to test two model-independent control strategies: (i) the relay controller, and (ii) the Proportional Integral controller.

The relay controller is the simplest type of model-independent control (12). It requires
only the computation of the control error at each sampling time \((kT)\), where \(T=15\) min, whose sign dictates which input will be provided to the cells. Specifically, cells will be treated with standard growing medium for the next 15 min, if \(e = (r - y_m) \geq 0\), or Tetracycline-containing medium otherwise (Methods and Supporting Information).

The drawback of this controller is that it does not take into account the dynamics of the system, therefore it can be slow in converging to the desired set-point and show marked oscillations around it. In addition, when the measured value \(y_m\) is very close to the reference, the control error can rapidly change sign, thus causing the control input to continuously switch between the two growing media (a phenomenon known as "chattering"). To avoid chattering, we added a 5% hysteresis interval to the controller (12), this corresponds to a percentage tolerance interval around the set-point, in which the algorithm ignores the control error value.

The Proportional-Integral (PI) controller is the most common technique used in practical applications. It computes the control input as the sum of two components: one proportional to the control error itself, and the other to its time integral (i.e. the past values of the error) (12). The proportionality constants \(K_p\) and \(K_i\) are called respectively proportional and integral gains, and their values were chosen by simple empirical rules (Methods and Supporting Information) (12). The PI controller produces a control input which can assume any positive or negative numerical value. However, cells can only be treated either with Tetracycline or standard medium. To solve this problem, we applied a Pulse Width Modulation (PWM) coding scheme. In PWM, at each sampling time, the control input is converted to a pulse of Tetracycline (ON) whose duration is proportional to the control input value. The Tetracycline pulse can vary from 0 min to 15 min (i.e. the sampling time), hence 0 min corresponds to the minimum control input value and 15 min to its maximum value. (Methods and Supporting Information).
**In silico** set-point control of gene expression

We decided to test beforehand the performance of our controllers *in silico*. We have previously derived a dynamical model of the monoclonal cell-line harbouring the Tet-OFF system used in this work\(^8\). Therefore, we employed this model to simulate the dynamics of d2EYFP in response to a time-varying input of Tetracycline (output). To simulate biological variability, we added Gaussian noise to the simulation output.

Figure 3 shows the simulated level of fluorescence in unstimulated conditions (Figure 3a), or following sustained administration of Tetracycline (Figure 3b).

Figure 3c-d shows the simulation’s result for the set-point control task when using the relay controller and the PI controller, respectively. In both cases, as expected, oscillations around the set-point are present. However, in the case of the PI controller, the amplitude of the oscillations tends to diminish over time. Figure 3

The relay and PI controllers do not require a quantitative model of the experimental system being controlled. The obvious advantage of this kind of control algorithms (i.e. no model is required), however, is also a disadvantage in terms of their performances, as evidenced by the oscillations around the set-point. Without a quantitative model, these control algorithms have no way to predict how gene expression will change in response to the input (Tetracycline- or standard-medium). Oscillations around the set-point are caused by this lack of predictive ability. On the contrary, if a quantitative model of the experimental system is available, a model-based control strategy can be used to predict the level of fluorescence in response to the input, and thus compute the optimal sequence of Tetracycline/standard medium pulses necessary to reach and maintain the desired fluorescence level.

We performed an *in silico* simulation of a model-based control strategy (Model Predictive Control) for the set-point control task. The results are shown in Supplementary Figure S3, where it can be observed that oscillations have a much smaller amplitude, when compared to the relay and PI controllers, and quickly dampen out. Of course, the control performance of model-based control strategies is strictly dependent on the quality of the quantitative model.
of the experimental system being controlled. Therefore in this work we decided not to test this strategy experimentally.

**In vitro set-point control of gene expression**

The set-point control task, depicted in Figure 1d, was to reach and maintain a set point equivalent to 50% of the fluorescence value measured during the calibration phase. This value is obtained by imaging cells in the absence of Tetracycline for 180 minutes at 15 min intervals, and then calculating the time-averaged value of fluorescence across the images.

Figure 4a-b shows the average level of fluorescence *in vitro* in the cell population in unstimulated conditions (Figure 3a), or following sustained administration of Tetracycline (Figure 4b). In both cases results are quite comparable to the *in silico* simulations.

Figure 4c-d shows the results of the *in vitro* control experiments for both the relay and PI controllers. As predicted *in silico*, the system is unable to settle on the reference value, but it persistently oscillates around it for the whole length of the experiment (3500 min). The behaviour of the relay controller (Figure 4c) is similar to the *in silico* simulation (Figure 3c), exhibiting undamped oscillations around the set point. The time-lapse experiment reported in Figure 4c is reported in the Supplementary Movie 1.

In the case of the PI controller in Figure 4d, oscillations have a decreasing amplitude with time, as predicted *in silico*. One difference we observed, when compared to the corresponding *in silico* simulations (Figure 3d), is the longer initial delay following Tetracycline administration before fluorescence starts to decrease. Nevertheless, the controller is able to react accordingly, and the fluorescence value of the cells is slowly brought towards the desired set point. We believe that this slight difference in the initial dynamics is due to intrinsic variabilities caused by experimental conditions and cell batches. The time-lapse experiment reported in Figure 4d is reported in the Supplementary Movie 2.
Concluding remarks

Here, we presented a proof-of-principle study to demonstrate that microfluidics-based control of gene expression in mammalian cells is feasible. In order to do that, we chose two very simple “input-output” control strategies (PI and Relay) that do not require a detailed quantitative model of gene expression. Oscillations around the set-point are to be expected when using these controllers. For example, the Relay control algorithm checks at each sampling time whether the fluorescence has reached the desired value, if not then it will treat cells either with Tetracycline, to reduce fluorescence, or standard-medium, to increase it, for the whole duration of the next sampling interval. This may cause fluorescence to decrease (or increase) too much, so that at the next sampling interval, its value will be below (or above) the desired level, giving rise to oscillations around the set-point (Figure 4c-d). Oscillations around the set-point are well documented in the control literature, for example in application of “bang-bang” control (13), (14). This is an optimal control strategy to minimise the time required for a physical system to achieve the desired set-point.

Control performances can be improved using a model-based control algorithms, such as Model Predictive Control, in place of model-independent controllers used in this work, as suggested by our in-silico results (Figure S3). Specifically, using these advanced controllers, it might be possible to reduce or eliminate oscillations around the desired set-point and to force cells to track a desired time-varying fluorescence value.

Control of gene expression may be achieved also in a single cell, rather than in a cell population, thanks to the robustness to noise exhibited by feedback control strategies (12). In yeast cells, for example, single-cell control has been successfully demonstrated (5). However the major, but attainable, obstacle in the case of mammalian cells is the image segmentation algorithm to track single cells over time. Due to the low optical density of the mammalian cell membrane, different cells are mistakenly merged together during image segmentation. One way to solve the problem would be to mark the nucleus of the cells by an additional constitutively expressed fluorescent protein, in order to help the image segmentation algorithm
to recognise individual cells.

Here we demonstrated that it is possible to regulate the expression of a reporter protein from the Tetracycline-inducible promoter in a population of mammalian cells using principles from automatic control engineering. To our knowledge, this is the first time automated control of gene expression from an inducible promoter has been achieved in mammalian cells.

The Tet-OFF system used in this work is orthogonal to all other genes in mammalian cells, has a very tight regulation and a high expression level (once induced) compared to other inducible systems in mammalian cells (11). In addition, it requires a drug (Doxycycline or Tetracycline) that is non-toxic and that can be administered systemically in mice and humans. Hence, the Tetracycline-inducible promoter is widely used in thousands of mammalian cell-lines, as well as animal models, in order to express a variety of trans-genes. Therefore our platform could be quickly implemented in any research laboratory to study already existing inducible cell lines, also derived from animal models, without the need of an ad hoc design of a genetic circuit.

Methods

Cell culture procedures

Chinese Hamster Ovary (CHO) tetO7-YFP cells(15) were maintained at 37°C in a 5% CO₂-humidified incubator, and cultured in α-MEM (Sigma) supplemented with 10% heat-inactivated TET system approved FBS (Life Technologies), 2 mM L-Glutamine and 100 U/ml Penicillin/Streptomycin (Gibco). Cells were passaged upon reaching of confluence, by rinsing the monolayer once with sterile filtered Phosphate-Buffered Saline (PBS, Gibco) and exposing cells to 0.25% Trypsin-EDTA for 1 minute. Cells are then resuspended and plated at a density of $10^4$ cells/cm².
Fluorescence time-lapse microscopy and control experiments

Image acquisition was performed using a Nikon Eclipse TI-E inverted epifluorescence microscope with a digital camera (Andor iXon897, Andor) and an incubation chamber (H201-OP R2, Okolab), a 40× objective (Obj. PlanFluor DLL 40× Ph2, Nikon), a 20× objective (Obj. CFI PF DLL 20× Ph1, Nikon), a TRITC HYQ filter (excitation 530 - 560 nm, emission 590 - 650 nm, Nikon) and a YELLOW GFP BP HYQ filter (excitation 490 - 510 nm, emission 520 - 550 nm, Nikon). Exposure time for experiments in chamber slides was set to 30 ms for the phase contrast images (with transmitted light lamp voltage was set to 3V), 300 ms for observation of d2EYFP, and 700 ms for observation of mCherry. Exposure time for microfluidics experiments was set to 286 ms for the phase contrast images (with transmitted light lamp voltage was set to 1.5V), 300 ms with EM gain set to 50 for observation of d2EYFP and Sulforhodamine tracker. Images were acquired every 15 minutes for all the experiments. Images were processed via a custom-made segmentation algorithm to quantify the average fluorescence expressed by the cell population (15). Temperature was maintained constantly at 37°C, and CO₂ concentration was set to 5% of the total air volume injected in the incubation chamber. Experiments were performed and images were extracted using the NIS-Elements AR v.3.22.14 software package and the Perfect Focus System (Nikon Instruments) to maintain the focal plane throughout the experiment. At the beginning of each experiment, cells were allowed to grow in untreated medium for three hours, to rule out any fluctuation in metabolism due to displacement of the device from the cell incubator to the incubation chamber of the microscope. Cells were then switched to Tetracycline-treated medium at the indicated concentrations for different time intervals. We quantified the mean fluorescence for each group of cells in a trap, and in order to rule out cell-to-cell variability, and we plotted the mean and standard deviation among all traps (i.e. replicates) for each experiment. For control experiments, tetO7-YFP cells were loaded into the microfluidic device as described in the Supplementary Information, where microfluidics device’s fabrication is also detailed. For calibration experiments, the timing of the exposure to the drug
(either Doxycycline or Tetracycline, according to the specific experiment) was set \textit{a priori}.
Each control experiment began with an initial calibration phase of 180 minutes in untreated medium. After the calibration phase the control experiment being and the duration of the exposure of the cells to Tetracycline was computed at each time point by the control algorithm (relay or PI). Fluorescence quantification was automatically performed in real time by a segmentation algorithm implemented in Matlab2010b, based on the one already published in (8). Details on the control algorithms design and implementation are described in the Supplementary Information.

**Associated content**

**Supporting Information**

Additional text and figures referenced in this article.

**Supplementary Methods**

Microfluidic device fabrication protocol and control algorithms (PI, Relay, MPC).

**Supplementary Movies**

Movies of control experiments reported in this article.

**Acknowledgements**

We thank Prof. Jeff Hasty and his lab for providing the design of the microfluidic device and details on how to use it.
Funding sources

This work was supported by a Human Frontier Science Program HFSP RGP0020/2011 to DdB and by the Italian Fondazione Telethon.
Figures Legends

Figure 1: Control platform and control algorithms. (a) Cells are grown in a microfluidic device within a temperature- and CO\textsubscript{2}-controlled environment under an inverted fluorescence microscope. Images are acquired every 15 minutes (y), and quantification of fluorescence is performed via an image-segmentation algorithm. The measured level of fluorescence (y\textsubscript{m}) is compared to the desired reference value (r). The control algorithm computes the input u to be provided to the cells and moves the syringes containing either untreated medium or Tetracycline-treated medium. (b) Block scheme of a Relay controller. The output from the system (Plant) is subtracted to the reference to obtain the error e; the error is processed by the Relay into a binary input u. (c) Block scheme of a PI-PWM controller. The error is calculated as in b, the control input \( \hat{u} \) is proportional to the control error and to its time-integral, a Pulse Width Modulation block is used to transform \( \hat{u} \) into a binary input u. (d) Schematic representation of the control task. The high steady state of fluorescence (black line) is calculated by averaging the first 180 minutes of acquisition (calibration phase), then the reference (blue line) is calculated as 50% of the high steady state; the control task is to keep the measured level of fluorescence on the reference for 3500 minutes.

Figure 2: The experimental model: tetO\textsubscript{7}-YFP monoclonal cell line. (a) The Tetracycline transactivator (tTA) is constitutively expressed by a CMV promoter. The tTA binds the CMVTET promoter harbouring seven Tet Responsive Elements (tetO\textsubscript{7}) upstream of a minimal CMV promoter, thus driving the transcription of a destabilised EYFP (d2EYFP). Doxycycline, or Tetracycline, bind the tTA protein and prevent its binding to the CMVTET promoter. (b) Representative phase contrast (left panel) and fluorescence (right panel) images of tetO\textsubscript{7}-YFP cells in a microfluidic device grown in the absence of Tetracycline.

Figure 3: In silico set-point control experiments. The red line under each graph represents the respective Tetracycline input simulated. The green line represents the simulated fluorescence output. (a) Simulation of the model when no Tetracycline is provided to cells. (b) Simulation of the model when a saturating concentration of Tetracycline is applied to the cells. (c,d) Simulations of the Relay (c) and PI (d) control algorithms when the reference set-point value (cyan line) is equal to 50% of the fluorescence level exhibited by cells in the absence of Tetracycline.
Figure 4: *In vitro* set-point control experiments. The red line under each graph represents the respective Tetracycline input provided to cells in the microfluidic device. The green line represents the measured mean fluorescence of the imaged fields (mean and standard deviation of 3 to 8 replicates). (a) Fluorescence exhibited by the cells in the presence of growing medium with no Tetracycline. (b) Fluorescence exhibited by the cells in the presence of 100 ng/ml Tetracycline in the growing medium. (c,d) Results of the Relay (c) and PI (c) control experiments for the set-point control task (cyan line) where the desired fluorescence is equal to 50% of the initial fluorescence exhibited by cells during the calibration phase of 180 min.

References


6. Darzacq, X., Shav-Tal, Y., de Turris, V., Brody, Y., Shenoy, S. M., Phair, R. D., and


