# CELL0014 Report 2

Candidate number: LRMV8

University College London

## Introduction

Traditionally, in biology, we focus on the study of naturally occurring systems - attempting to focus on ever-smaller biological mechanisms; trying to decode and understand the tangled web of intertwined processes that living organisms are. Yet, an alternative and a rapidly growing field – synthetic biology – has a radically different approach — using man-designed biological systems which do not only present with an unimaginable number of potential applications for the future (much like the rise of 3D-printing is revolutionising the manufacturing industry), but also offer precious insight into what is required to construct robust, reliable, and predictable biological circuits and genetic networks. Nevertheless, designing *de novo* biological systems is not a straightforward feat, requiring a combination of engineering, mathematics, and both *in silico* and *in vivo* biology.

The Goodwin Oscillator, first theoretically described by Brian Goodwin in the 1960s, is regarded as the first and most simple of the *de novo* biological systems and has since been studied and implemented many times [1;2]. It comprises of a single self-repressing gene, which under the correct set of conditions results in oscillatory rises and falls in the concentration of the genes' product.

Extending the Goodwin Oscillator are repressilators — cyclical sets of one or more genes each of which is inhibiting its successor in the set $(1 \dashv 2 \dashv \ ... \ \dashv n \dashv 1)$ [3;2]. The term *Repressilator* was coined by Elowitz and Leibier (2000) [4] in their influential work describing the first *de novo* oscillatory mechanism *in vivo*. It uses a set of three regulatory genes — *lacI* from *Escherichia coli* (repressed by cI), *tetR* from the Tn10 transposon (repressed by LacI), and *cI* from the Bacteriophage $\lambda$ (repressed by TetR; see Fig. 1) [4]. In this report, I attempt to revisit the work of Elowitz and Leibier and verify their claims utilising both deterministic and stochastic modelling. First, however, it is paramount to understand the original design of the Repressilator.

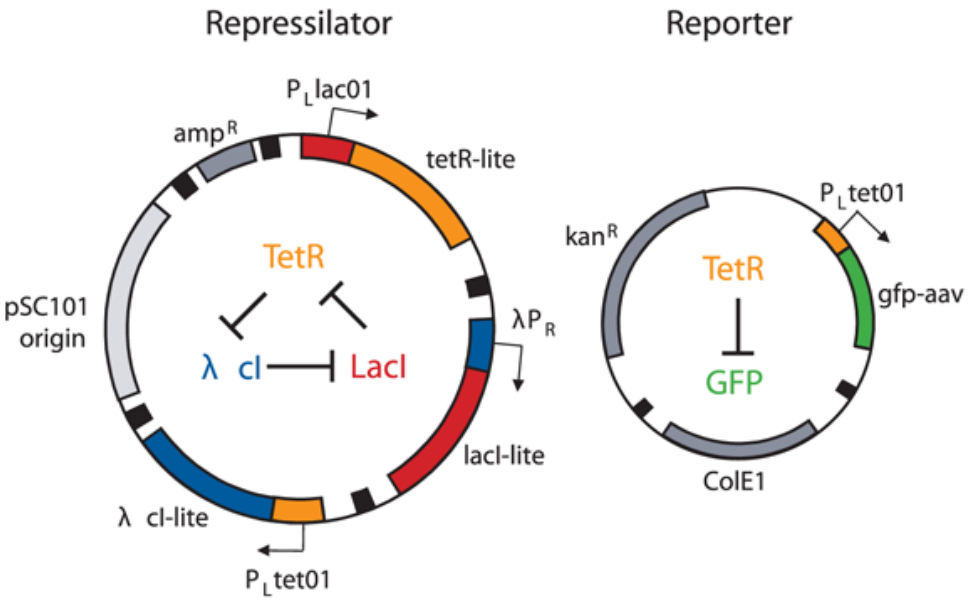


Fig. 1: **Repressilator design.** The Repressilator is composed of three regulatory genes, *lacI*, *cI*, and *tetR*, under the control of the $P_R$ promoter with the cI operator, the $P_L$ promoter with the TetR operator, and the $P_L$ promoter with the LacI operator, respectively. The genes are terminated with proteolysis targeting tags (as denoted by the "*lite*" suffix). In addition to the Repressilator, the system also includes the Reporter carrying the intermediate-stability GFP (*gfp-avv*) under the control of the $P_L$tetO1 promoter allowing to monitor the system state with fluorescent microscopy *in vivo*. The plasmids are equipped with the low-copy pSC101 and the high-copy ColE1 *E. coli* replication origins, and ampicillin and kanamycin resistance genes (allowing to select for cell containing only both plasmids), respectively. T1 terminators from the *E. coli rrnB* operon used to isolate the individual open reading frames are shown as black boxes. Gene network diagrams are displayed to illustrate the oscillatory negative-feedback loop circuit. Adapted from Elowitz & Leibier (2000) [4].

## The Repressilator

**Mathematical design:** Due to the contemporary limited understanding of biochemistry, the team set out to first create a simplistic ordinary differential equation (ODE) model of the biological circuit capturing only the key processes rather than modeling the precise system behaviour. The Repressilator comprises six coupled differential equations:

$$\begin{aligned} \frac{dm_i}{dt} &= -m_i + \frac{\alpha}{1+{p_j}^n} + \alpha_0 \\ \frac{dp_i}{dt} &= -\beta(p_i - m_i) \end{aligned} \qquad \begin{pmatrix} i = lacI,\ tetR,\ cI \\ j = cI,\ lacI,\ tetR \end{pmatrix},$$

where:

| | | |
|---|---|---|
| $m_i$ | = | mRNA concentration |
| $p_{i,j}$ | = | protein concentration |
| $\alpha_0$ | = | promoter "*leakiness*" (number of mRNA molecules transcribed from the promoter under saturating levels of repressor) |
| $\alpha + \alpha_0$ | = | number of mRNA molecules transcribed from the promoter produced in the absence of repressor |
| $\beta$ | = | ratio of protein decay rate to mRNA decay rate |
| $n$ | = | Hill coefficient (describing binding cooperativity); |

all parameters are identical for all three genes, except for their DNA-binding functions; time is rescaled in units of mRNA lifetime; protein concentrations are written in units of $K_M$ (number of repressor molecules required to reduce the rate of transcription by half); and mRNA concentrations are rescaled by their translation efficiency (average number of proteins produced per mRNA molecule)[4].

After exploring the parameter space (some of which I will engage in below), the authors describe a set of biologically sensible parameters:

| | | |
|---|---|---|
| $\alpha_0$ | = | $5 \times 10^{-4}$ $[s^{-1}]$ |
| $\alpha + \alpha_0$ | = | 0.5 $[s^{-1}]$ |
| average translational efficiency | = | 20 [proteins per transcript] |
| $n$ | = | 2 |
| protein $t_{1/2}$ | = | 10 [min] |
| mRNA $t_{1/2}$ | = | 2 [min], |

which result in oscillatory rises and falls of the mRNA and protein concentrations (see Fig. 2)[4].

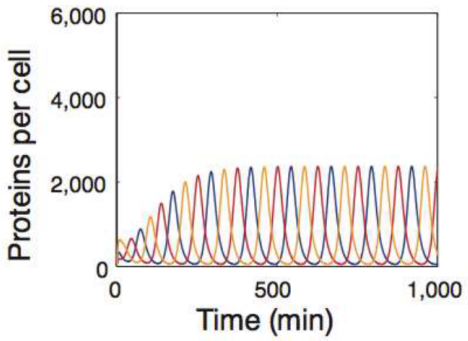


Fig. 2: **Repressilator protein concentration oscillations.** Color coding as per Fig. 1. Adapted from Elowitz & Leibier (2000)[4].

Mathematically, the oscillations are associated with a limit cycle and arise via a super-critical Hopf bifurcation for sufficiently high $n$ under the model proposed above[2;3;4]. The Hill function, in the context of transcriptional repression, represents the formation of repressor protein complexes or cooperative binding of repressor to the promoter[1]. This sigmoidal function describes the steepness of the repressional response (as discussed below). In other words, sufficiently boolean repression is required for the Repressilator to produce oscillatory behaviour[2]. Additionally, it is also required for a time delay to exist in the negative feedback loop for repressilators to produce oscillations[2;5].

Generally, repressilators were found to favour strong promoters with low leakiness, high translational efficiency, sufficiently nonlinear repression, the presence of a time delay in the system, and similar protein and mRNA decay rates. Additionally, models (including this document) usually assume all genes and their associated parameters to be identical[2], however, that can be biologically unpractical/unfeasible, and it has been shown that even unsymmetrical repressilator systems can produce oscillations[6]. Lastly, repressilator systems seem to prefer odd-numbered sets of genes, nevertheless, there is some evidence that oscillating solutions exist even for 6+ even-numbered sets and it might thus be only two- (which can form stable on/off switches) and four-membered sets of genes that are unable to produce oscillating behaviour[2;6].

**Biological design:** After exploring the circuit design *in silico*, the team set out to create the Repressilator *in vivo*. They resulted to utilising the $P_L$ promoter combined with LacI and TetR operator sequences to control the expression of *tetR* and *cI*, respectively, and the $P_R$ promoter containing the cI operator to control the expression of *lacI* (both promoters originating from the Bacteriophage $\lambda$; see Fig. 1) for their strong expression and tight repression[4].

Such constructed reading frames were cloned into a low-copy plasmid and transformed into the *E. coli* MC4100 $\Delta$(argF-lac)U169 strain with the Lac operon disabled to reduce interference between the Repressilator and the *E. coli* metabolism[4]. Additionally, Elowitz and Leibier also transformed their bacteria with a high-copy plasmid "*Reporter*", which was carrying the intermediate-stability green fluorescent protein variant (*gfp-aav*; with $t_{1/2} \approxeq 30 - 40$ min)[7] under the control of the $P_L$tetO1 to allow for a fluorescent readout of the circuit state (see Fig. 1)[4]. To reduce the relatively high protein half-life to be more comparable with the half-life of mRNA, as required by the model, the team also inserted ssrA tags (making it a target recognised by *E. coli* proteases) at the 3' ends of the genes[4].

As mentioned before, at the time of publication, many of the critical parameters were poorly known and thus it was not clear whether the system will indeed oscillate when tested *in vivo*. Conveniently, to the pleasure of the team, at least 40 % of the cells exhibited oscillatory blinking of green light radiating from the GFP regulated by the Repressilator

with peak-to-peak frequency of $160 \pm 40$ min (mean $\pm$ s.d.; notably, roughly threefold longer than the average cell-division time in the experiment)[4].

Additionally, the team also observed significant noise in the oscillatory behaviour, both, in-between mother-daughter and daughter-daughter cells with cells skipping and delaying oscillations[4]. Even though it was later shown that this erraticity was to a certain extent caused by the limited imaging technology available at the time[8], this chaotic behaviour can to a large degree be accounted to the inherent randomness and stochasticity of living organisms and the universe. For that reason, the researchers also explored a stochastic simulation of the Repressilator using the Gillespie algorithm (some of which I do myself later in this report) showing that the modelled system behaved not dissimilarly to the *in vivo* observations[4].

**Future improvements:** It is worth mentioning that since its publication in 2000, the design by Elowitz and Leibier has been subject to numerous investigations which yielded much insight into the dynamics of the system[2]. Most notably, study by Potvin-Trottier *et al.* (2016)[8] where the team set out to minimise noise in the original Repressilator, trying to achieve the best performace with the most minimalistic design possible.

Firstly, they showed that poor segregation of the high-copy Reporter caused significant fluctuations in the mother-daughter levels of GFP and just moving $P_L$tetO1 *gfp-aav* onto the more stable low-copy Repressilator alone lowered the fluorescence amplitude standard deviation from the original 78 % to 36 %[8].

Secondly, the team described another significant source of noise - the *E. coli* protein degradation machinery, which can be over-saturated when both the intermediate-stability GFP and the reporter proteins are being targeted for degradation via the ssrA-tags[8]. Removing the ssrA-tags from the reporter genes (or lowering the number of gene copies), it strongly improved the accuracy of oscillations[8].

Lastly, Potvin-Trottier *et al.* also showed that, the strength of TetR's repression is so high, that even an extremely low number of it's molecules can terminate promoter transcription, which subjects it's regulation under a large amount of stochastic noise. Increasing TetR's repression threshold, by adding additional "soaker" binding sites, greatly reduced the noise all steps in the circuit[8].

## Analysis

Alongside discussing the Repressilator's design and relevance, this document also aims to revisit the original work by Elowitz and Leibier [4] and reconfirm their claims and provide additional insight into the system's behaviour.

Expanding the original ODEs, we arrive at set of equations:

$$\begin{aligned}
\partial^{+} m_i(t,\, t{+}1) &= k_m \frac{K^n}{K^n + P_j(t)^n} + k_{m0} \\
\partial^{+} P_i(t,\, t{+}1) &= k_p * m_i(t) \\
\partial^{-} m_i(t,\, t{+}1) &= k_{dm} * m_i(t) \\
\partial^{-} P_i(t,\, t{+}1) &= k_{dp} * P_i(t) \qquad \begin{pmatrix} i = lacI,\, tetR,\, cI \\ j = cI,\, lacI,\, tetR \end{pmatrix}, \\
\Delta m_i(t,\, t{+}1) &= \partial^{+} m_i(t,\, t{+}1) - \partial^{-} m_i(t,\, t{+}1) \\
\Delta P_i(t,\, t{+}1) &= \partial^{+} P_i(t,\, t{+}1) - \partial^{-} P_i(t,\, t{+}1) \\
m_i(t{+}1) &= m_i(t) + \Delta m_i(t,\, t{+}1) \\
P_i(t{+}1) &= P_i(t) + \Delta P_i(t,\, t{+}1)
\end{aligned}$$

where[i]:

| | | |
|---|---|---|
| $t$ | = | time [min] |
| $K$ | = | 40 (repression $K_M$, number of repressor molecules required to reduce expression by 50 %) |
| $n$ | = | 2 (Hill coefficient) |
| $P_{i,j}(t)$ | = | number of protein molecules at time $t$ |
| $m_i(t)$ | = | number of mRNA molecules at time $t$ |
| $k_m$ | = | 30 [$\text{min}^{-1}$] (maximum rate of transcription; $\alpha + \alpha_0$) |
| $k_{m0}$ | = | 0.03 [$\text{min}^{-1}$] (promoter "leakiness; $\alpha_0$) |
| $k_p$ | = | 6.931 [$\text{min}^{-1}$] (translational efficiency) |
| $k_{dm}$ | = | 0.3466 [$\text{min}^{-1}$] (rate of mRNA decay) |
| $k_{dp}$ | = | 0.06931 [$\text{min}^{-1}$] (rate of protein decay), |

assuming saturated transcription and translation, which can be used to model the mRNA and protein levels across time. In the chapters below, I discussed how I used these equations implemented in Python to study the Repressilator system. All code used to conduct this analysis is attached as part of the Supplementary Information.

[i] All parameter values have been converted from the original Repressilator parameters described above (see Supplementary Information for detailed calculations).

**Task 1**

First, let us consider a simplified scenario with only two genes out of the three regulatory genes – *lacI* and *tetR*:

$$P_{\text{lacI}} \dashv P_{\text{tetR}}$$

where we are concerned only with two out of the six ODEs describing the Repressilator ($m_i$ & $P_i$; $i = tetR$, $j = lacI$). Assuming constant levels of LacI, we can examine the steady-state (once the system has reached it; $t_{\text{MAX}} = 1000$ [min]) concentration of TetR under different parameters — namely, the Hill coefficient. This will allow us to see how different levels of cooperativity affect the repression kinetics between LacI and TetR under the model described above[ii] (see Fig. 3).

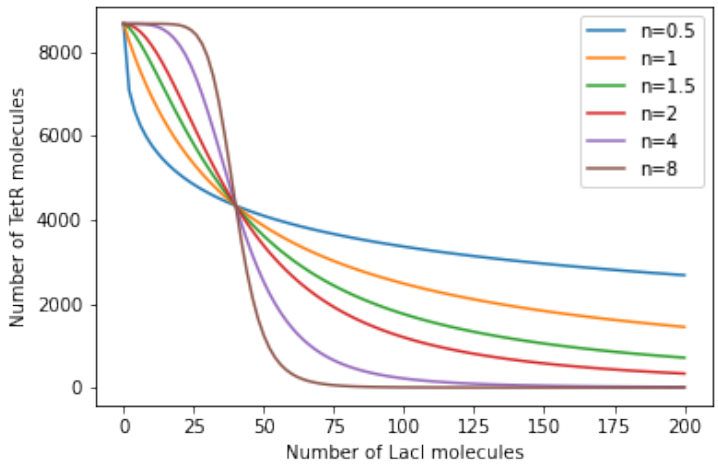


Fig. 3: **Characteristics of repression kinetics in the isolated system of $P_{\text{lacI}} \dashv P_{\text{tetR}}$.** With growing $n$ (Hill coefficient), repression of TetR production by LacI becomes more sigmoidal and boolean. Sufficiently high $n$ ($\gtrapprox 1.64$; see Fig. 9) is required to produce a Hopf bifurcation and for the system to oscillate. All curves converge at LacI = 40 ($K$; concentration of LacI required to reduce *tetR* expression by 50%).

I have tested $n$ ranging from 0.5 (negative cooperativity), 1.0 (no cooperativity), to 8.0 (high positive cooperativity) — a biologically relevant range (in the context of transcriptional repression, it is almost impossible to reach Hill coefficients over 4[1]). In the simulation we can observe that with the Hill coefficient increasing, the repression curve of $P_{\text{lacI}} \dashv P_{\text{tetR}}$ grows

[ii] Initial conditions: $m_{tetR}(t=0) = 0$, $P_{\text{TetR}}(t=0) = 5$

steeper - decreasing the number of molecules required to switch the gene expression state (while $K$ stays fixed). Biologically, this behaviour could be explained either by repressor protein multimerisation (increasing the affinity of biding to the operator, or by cooperative binding (where repressor protein binding to the operator increases the binding affinity for other repressor molecules [1].

### Task 2

Now, let us consider the full Repressilator circuit as described above[iii]. When simulated for 1000 min, the system exhibited stable and indefinite (as they are associated with a limit cycle; see below) oscillatory rises and falls in the regulatory-proteins' concentrations consistent with the results reported by Elowitz and Leibier in 2000 (see Fig. 2 and Fig. 4) [4].

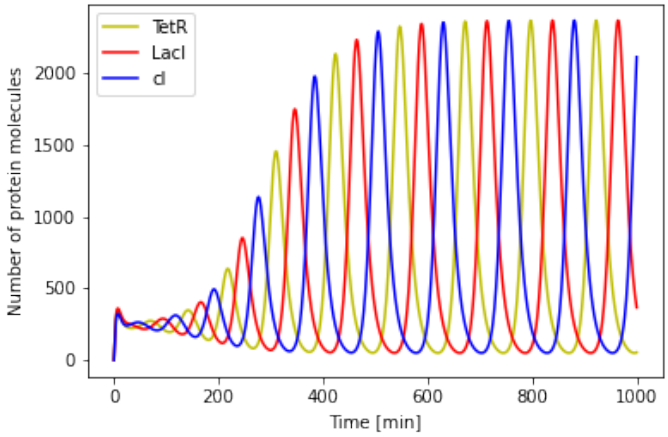


Fig. 4: **Repressilator ODE simulation results.** It is clear, that the Repressilator system as described as above does indeed produce an oscillatory behaviour of the regulatory protein concentrations *in silico*. After a transient period, after the system reaches the limit cycle, the peak amplitude $\approxeq 2500$ molecules, and the peak-to-peak period $\approxeq 125$ min. Colour coding as per Fig. 1.

In addition to the limit cycle oscillations, the system can also enter a steady state as per the assumption that all gene parameters are identical. This steady state arises when the initial conditions for all three genes are also identical ($P_{\mathrm{LacI}}(t{=}0) = P_{\mathrm{TetR}}(t{=}0) = P_{\mathrm{cI}}(t{=}0) \ \wedge \ m_{lacI}(t{=}0) = m_{tetR}(t{=}0) = m_{cI}(t{=}0)$) as shown in Fig. 5. However, should even only a minuscule difference exist between the three sets of initial conditions (for each gene) exist, the system will eventually reach the limit cycle and produce oscillations as shown in Fig. 6.

[iii] Initial conditions (used throughout the analysis, unless stated otherwise): $m_{lacI}(t{=}0) = 0$, $m_{tetR}(t{=}0) = 0$, $m_{cI}(t{=}0) = 0$, $P_{\mathrm{LacI}}(t{=}0) = 0$, $P_{\mathrm{TetR}}(t{=}0) = 5$, and $P_{\mathrm{cI}}(t{=}0) = 0$

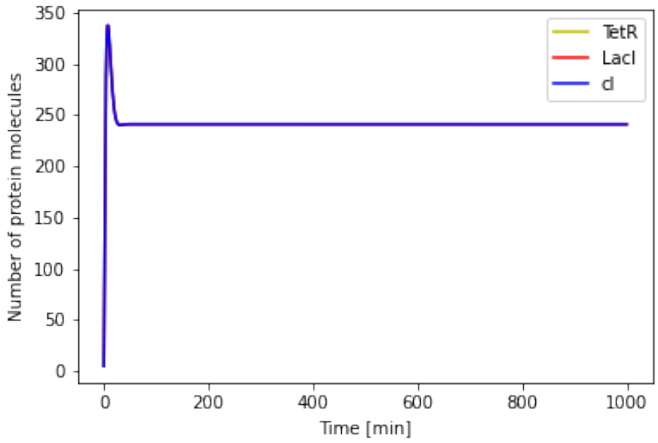


Fig. 5: **Repressilator ODE simulation of steady-state.** Instead of producing oscillations, the system enters a steady state ($P_{\text{ALL}} \approxeq 240$ [molecules]) when the initial conditions for all three genes are identical. Note that only the curve representing the cI concentration is visible, as all plotted concentrations are identical and overlap. Initial conditions: $m_{lacI}(t{=}0) = 0$, $m_{tetR}(t{=}0) = 0$, $m_{cI}(t{=}0) = 0$, $P_{\text{LacI}}(t{=}0) = 5$, $P_{\text{TetR}}(t{=}0) = 5$, and $P_{\text{cI}}(t{=}0) = 5$. Colour coding as per Fig. 1.

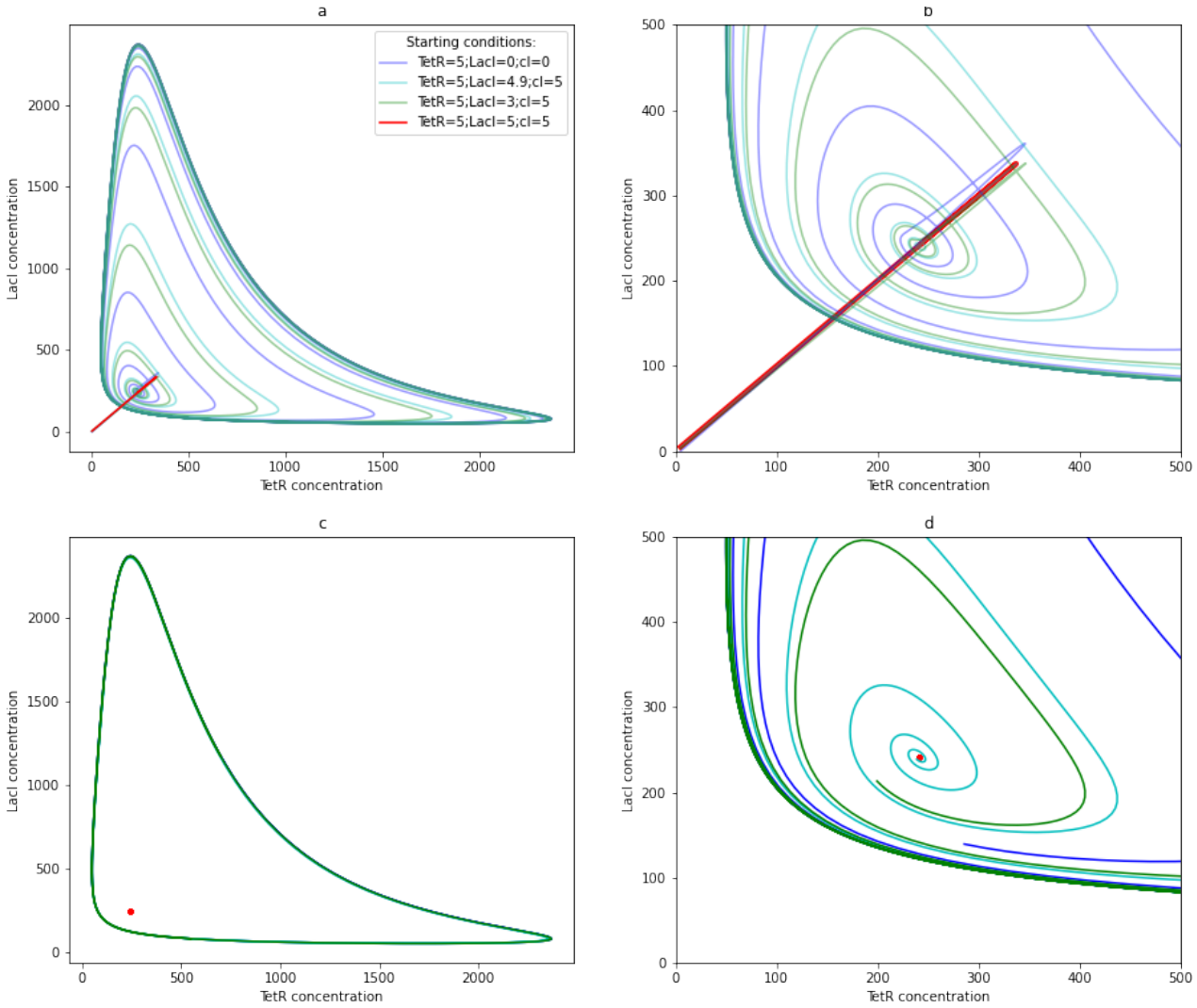


Fig. 6: **Repressilator ODE Phase plot.** This plot is showing the relation of TetR and LacI concentrations, nevertheless, the system comprises of 3 genes. To be able to completely understand the system behaviour, a 3D representation is required (See Supplementary Information). **(*a*)** Repressilator simulations ($t_{\mathrm{MAX}} = 1000$ min) with different starting protein concentrations are shown (all mRNA concentrations = 0). Regardless the starting conditions (unless they are on the limit cycle), the system will promptly collapse close near the steady state and, unless started with all concentrations identical (in which case the system will enter the steady state), the system will start approaching the limit cycle. The steady state simulation is shown in red; its linear trajectory represents the system behaviour in Fig. 5 where the protein concentrations reach the steady state after a short period of transient rise and fall, with all concentrations being identical. **(*b*)** Showing simulations as in ***a***, but cropped to better show the steady state a identical simulation approaches to the limit cycle. **(*c*)** Showing simulations as in ***a*** with the initial transient behaviour (first 800 min of simulation) removed - showing that all simulations except for the steady state (shown in red) are associated with the limit cycle. **(*d*)** Showing simulations as in ***a*** with the initial 200 min (once the steady state simulation arrived at the steady state) removed to show that while, the initial conditions might be almost identical and the system will come very close to the steady state, it will always eventually reach the limit cycle (the closer the initial conditions to being identical, the longer it will take for the system to reach the limit cycle; see Supplementary Information).

### Task 3

In previous Tasks, we have examined how the Hill coefficient affects the repression profile, and how the Repressilator produces oscillations. Here, I will demonstrate how the Hill affects the system behaviour.

When we simulate the Repressilator with all parameters as outlined above, except for setting the Hill coefficient to a sub-critical level ($\lessapprox 1.64$; see Fig.9) for the Hopf bifurcation to arise ($n = 1.3$), the system will not produce oscillations, but will rather settle at a steady state ($P_{\text{ALL}} \approxeq 420$ [molecules]; see Fig. 7). Extending this approach and simulating over a range of $n$ (see Fig. 8), shows what a drastic effect the Hill coefficient has on the behaviour of the Repressilator. When $n$ becomes super-critical, as $n$ increases, so does the oscillation peak amplitude (see Fig. 9), while the peak-to-peak period is unchanged (data not shown).

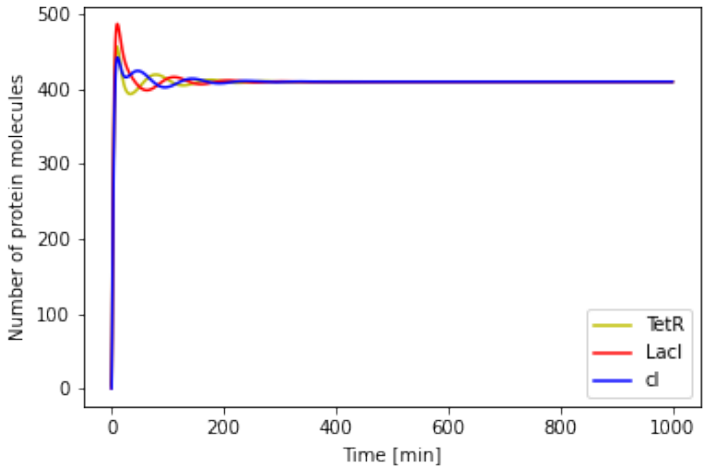


Fig. 7: **Repressilator ODE simulation with sub-critical $n$.**

Additionally, in this section, I will also examine how additional system parameters — translational efficiency ($k_p$; see Fig. 10) and promoter leakiness ($k_{m0}$; $\alpha_0$; see Fig. 11). It can be seen the relationship between the rate of translation and the peak concentration amplitude resembles a $\Gamma$ distribution, with its peak $\approxeq 12$. This can be interpreted as the fact that, too low of an efficiency is not able to produce enough protein before the next oscillation, however, too high of a translational efficiency and the genes expression will be silenced a sufficient amount of mRNA can be produced. In the context of promoter leakiness, the protein amplitude decreases drastically with growing leakiness, and as $k_{m0}$ approaches 0.068, the system seizes to oscillate. This observation can likely be explained by the fact that the background expression of the Repressilator is so high, that the promoters are constantly under saturating levels of repressors.

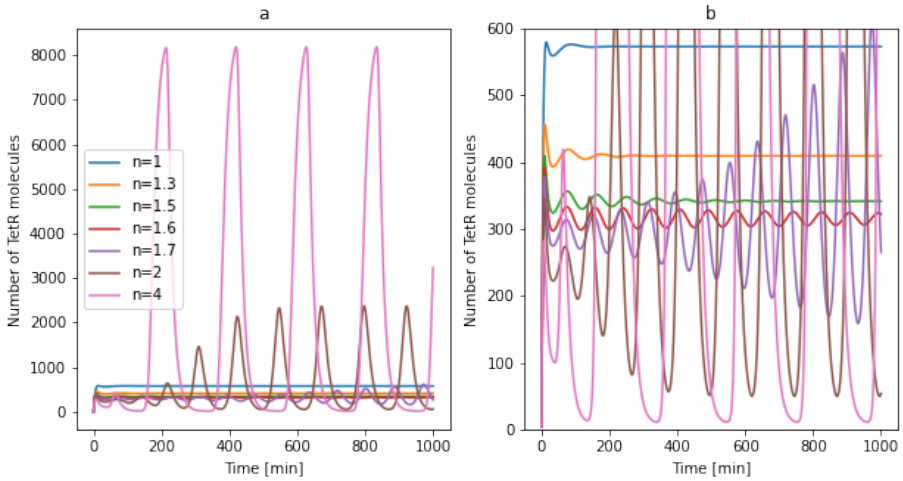


Fig. 8: **Repressilator ODE simulation over range of $n$. (*a*)** With increasing $n$, the system behaviour changes. When $n$ rises above the super-critical level ($\gtrapprox 1.64$; see Fig. 9) and a Hopf bifurcation arises in the system, oscillations emerge. The repressilator was simulated over a biologically relevant range of Hill coefficients, as stipulated above. **(*b*)** Showing simulations as in ***a***, but cropped to better illustrate the fine transition from sub- to super-critical $n$.

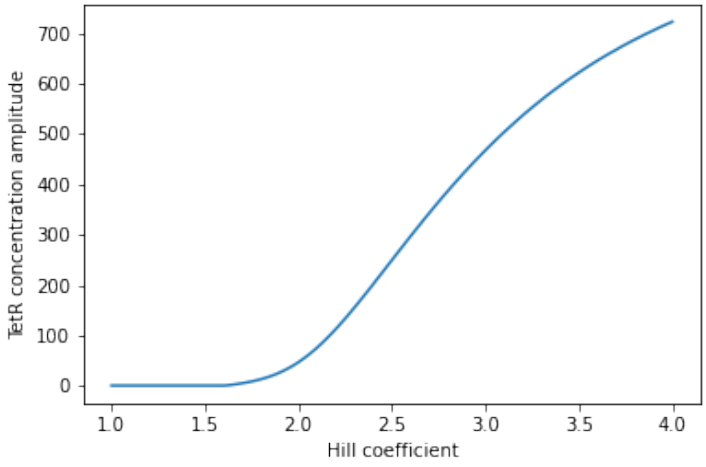


Fig. 9: **Protein peak amplitude in relation to changing $n$.** When $n$ crosses the super-critical level ($\gtrapprox 1.64$), oscillations occur in the system and with increasing $n$, the peak concentration amplitude rises as well (data only for TetR shown, but the relationship of $n$ to peak amplitude is identical for all three Repressilator genes).

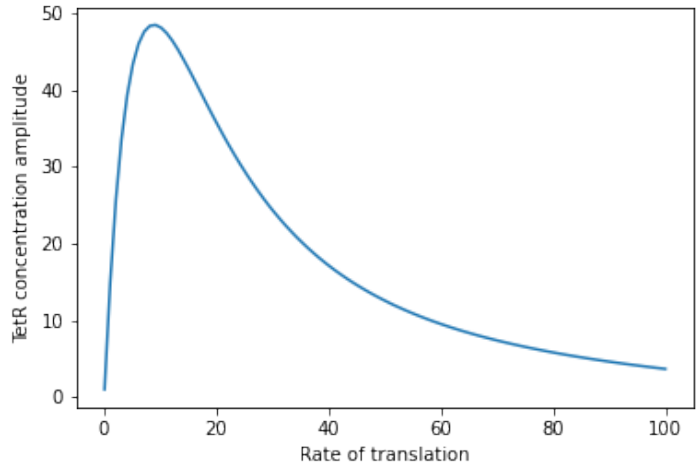


Fig. 10: **Repressilator ODE simulation over range of $k_p$.** Data only for TetR shown, but the relationship of $k_p$ to peak amplitude is identical for all three Repressilator genes.

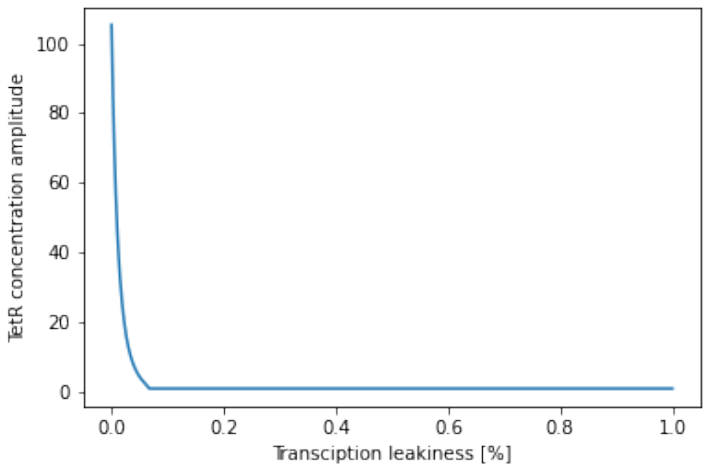


Fig. 11: **Repressilator ODE simulation over range of $k_{m0}$.** Data only for TetR shown, but the relationship of $k_{m0}$ to peak amplitude is identical for all three Repressilator genes.

**Task 4**

After having explored the Repressilator parameter space, let us examine behaviour of the system when we include the expression of GFP as controlled on the Reporter. This requires the addition of two differential equations into our system — $m_{gfp}$ and $P_{\mathrm{GFP}}$:

$$\begin{aligned}
\partial^{+} m_{gfp}(t,\, t{+}1) &= k_m \frac{K^n}{K^n + P_{\mathrm{TetR}}(t)^n} + k_{m0} \\
\partial^{+} P_{\mathrm{GFP}}(t,\, t{+}1) &= k_p * m_{gfp}(t) \\
\partial^{-} m_{gfp}(t,\, t{+}1) &= k_{dm} * m_{gfp}(t) \\
\partial^{-} P_{\mathrm{GFP}}(t,\, t{+}1) &= k_{dp_{\mathrm{GFP}}} * P_{\mathrm{GFP}}(t) \\
\Delta m_{gfp}(t,\, t{+}1) &= \partial^{+} m_{gfp}(t,\, t{+}1) - \partial^{-} m_{gfp}(t,\, t{+}1) \\
\Delta P_{\mathrm{GFP}}(t,\, t{+}1) &= \partial^{+} P_{\mathrm{GFP}}(t,\, t{+}1) - \partial^{-} P_{\mathrm{GFP}}(t,\, t{+}1) \\
m_{gfp}(t{+}1) &= m_{gfp}(t) + \Delta m_{gfp}(t,\, t{+}1) \\
P_{\mathrm{GFP}}(t{+}1) &= P_{\mathrm{GFP}}(t) + \Delta P_{\mathrm{GFP}}(t,\, t{+}1),
\end{aligned}$$

where all parameters are identical as outlined above except for GFP $t_{1/2} = 60$ [min] ($k_{dp_{\mathrm{GFP}}} = 0.01155$ [min$^{-1}$]) to reflect its lower degradation rate compared to the Repressilator proteins. The simulation (see Fig. 12) shows results comparable with those seen originally by Elowitz and Leibier[4].

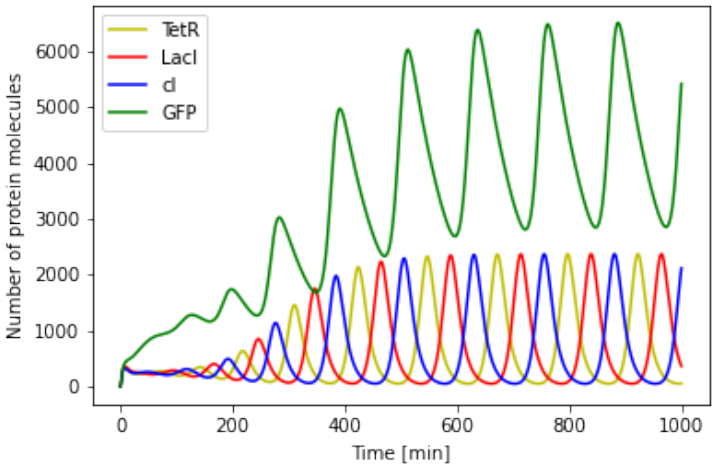


Fig. 12: **Repressilator ODE simulation with GFP.** The system simulation produces oscillations in protein concentration for both the Repressilator genes and the *gfp*. Notably, GFP minimal amplitude corresponds to the of cI as predicted by the model (both being controlled by a TetR operator). Additionally, it can be seen that the GFP concentration never returns to 0 as it does with the rest of the genes. This is due GFP's comparably longer $t_{1/2}$. Colour coding as per Fig. 1.

### Task 5

Let us consider options for finer control of the Repressilator. In this section, I introduce IPTG induction and control of the protein degradation rate in the system. This could be useful to gain a finer control over the system and for added potential functionality.

**IPTG Induction** IPTG can be used to "*soak*" LacI and thus control the expression levels of *tetR*. To introduce the IPTG induction into the system, we need to modify the ODE describing the production of *tetR* mrNA as follows:

$$\partial^{+} m_{tetR}(t,\ t{+}1) \ = \ k_m \frac{K^n}{K^n + (X * P_{\mathrm{LacI}}(t)^n)} + k_{m0}$$

where $X$ represents the proportion of available LacI not-bound to IPTG (i.e. IPTG concentration). Simulating this system with different concentrations of IPTG lends us insight into how it affects the Repressilator. With rising levels of IPTG (see Fig. 13 and 14), we can see both amplitude and frequency of TetR oscillations plummet and eventually reach a steady state ($T \approxeq 125$ [min]; $P_{\mathrm{TetR}} \approxeq 9000$ [molecules]) as we remove any affect of LacI onto the *tetR* promoter with saturating levels of IPTG.

**Protein Degradation Rate** To introduce the protein degradation rate modulation into the system, we need to modify the ODE describing protein degradation as follows:

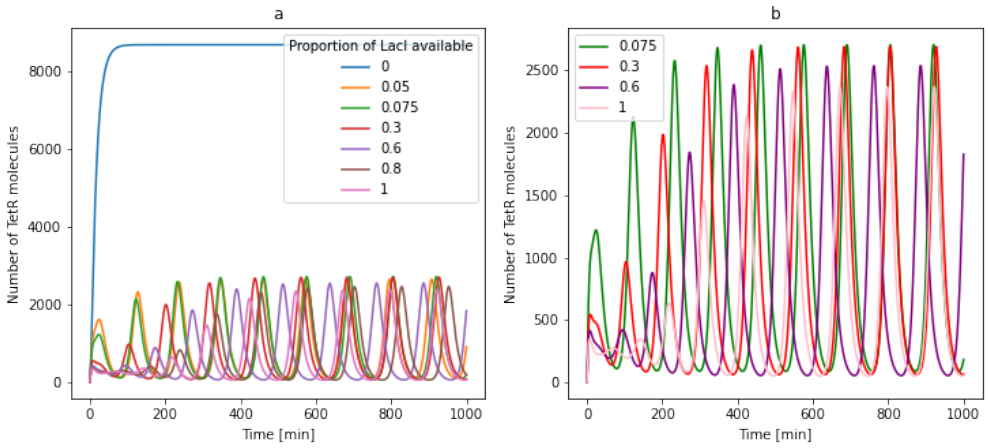


Fig. 13: **Repressilator ODE simulation over a range of IPTG concentrations. (*a*)** Overlay of Repressilator simulations at different IPTG concentrations. **(*b*)** Showing a selection of simulations as in ***a*** to better show the effect of lower IPTG concentrations on the Repressilator.

$$\partial^{-} P_i(t,\ t{+}1) \ = \ Y * k_{dp} * P_i(t) \qquad \big(i \ = \ lacI,\ tetR,\ cI\big)$$

where $Y$ represents the repression modulation factor ($Y = 0.167 \rightarrow t_{1/2_{\mathrm{P}}} = 60$ [min], $Y = 5 \rightarrow t_{1/2_{\mathrm{P}}} = 2$ [min]). From the simulation results (see Fig. 15 and 16), we can infer that while protein degradation does not affect the final oscillation period (however, it critically affects how much time is it going to take the system until it reaches the limit cycle), it certainly affects the oscillation amplitude — after a short period of the system reaching its maximum amplitude, the quicker the protein degradation the lower the amplitude, as not enough protein is able to build up before it gene expression is silenced.

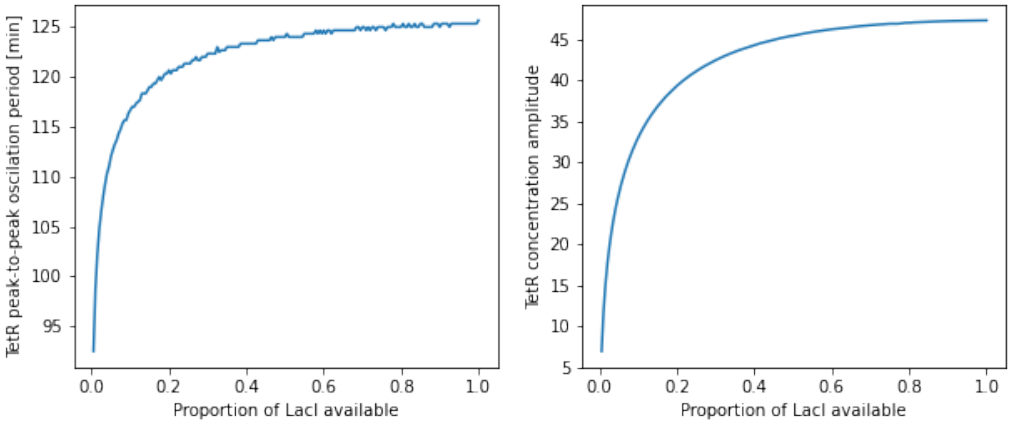


Fig. 14: **Repressilator protein oscillatory period and amplitude over a range of IPTG concentrations.** $t_{\text{MAX}} = 10000$ [min]. (*left* is staggered due to an imperfect period sampling)

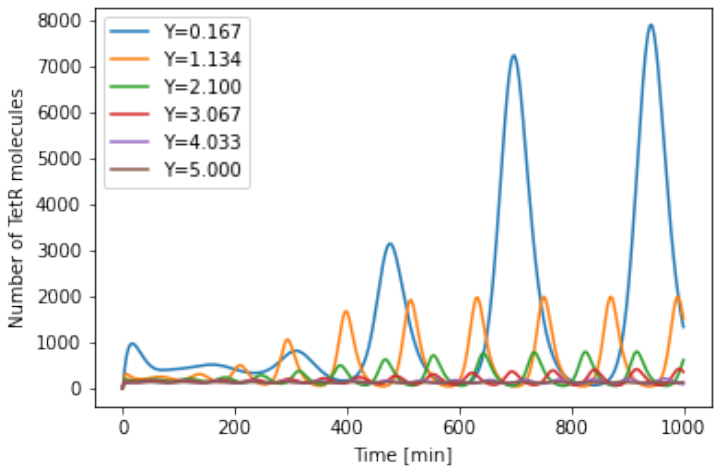


Fig. 15: **Repressilator ODE simulation over a range of $Y$ levels.**

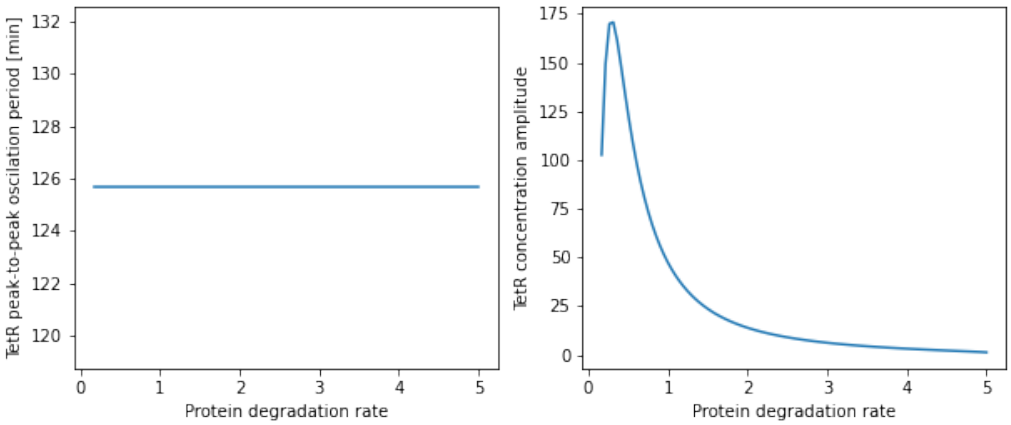


Fig. 16: **Repressilator protein oscillatory period and amplitude over a range of $Y$ levels.** $t_{\text{MAX}} = 10000$ [min].

### Task 6

Until now, I have explored the Repressilator using only deterministic ODE simulations. Sadly, however, the nature of living organisms is far from deterministic. Stochastic fluctuations, which can critically influence the behaviour of the system, need to be accounted for and expected. Therefore, we will now model the Repressilator using the stochastic Gillespie algorithm (see Supplementary Information).

Simulating the Repressilator in this fashion results in the proteins' concentration oscillating with variable amplitude and period (see Fig. 17), consistently with results reported by Elowitz and Leibier[4].

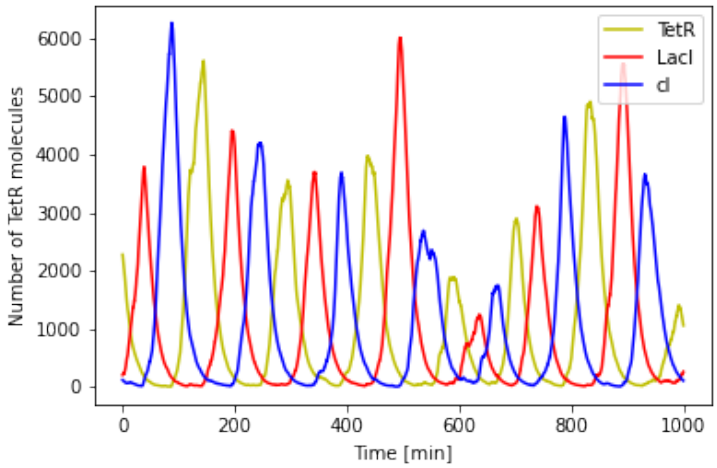


Fig. 17: **Repressilator stochastic simulation.** Randomness seed = 42. Colour coding as per Fig. 1.

By overlaying several random simulations of the Repressilator (see Fig. 18), we can see that the different simulations can drastically differ from one another. To get a better idea of the system behaviour, we can perform *de-facto* a Monte-Carlo approximation of the distributions in peak height and peak-to-peak period (see Fig. 19). Doing so, reveals that the average simulation of the Repressilator behaves very closely to its deterministic alternative, demonstrating that the system can produce oscillations even stochastically driven scenarios (such as an *in vivo* study).

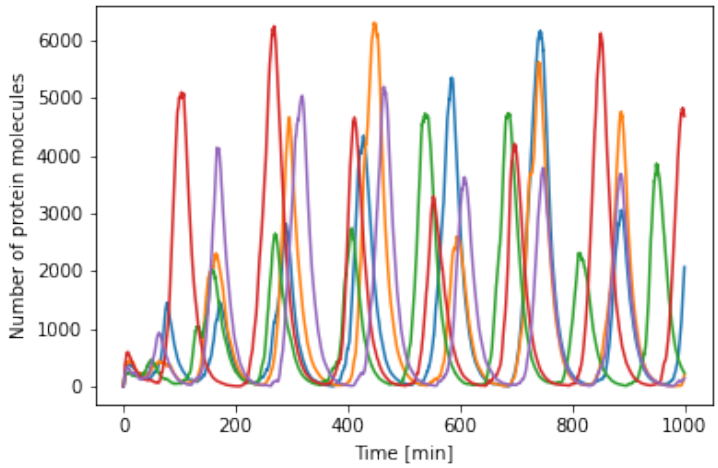


Fig. 18: **Multiple Repressilator stochastic simulations.** Multiple stochastic simulations of the Repressilator shown overlaid. Randomness seeds = 42, $6.67408 * 10^{-11}$, 299792.458, 554.984319180, $6.02214076 * 10^{23}$. Colour coding as per Fig. 1.

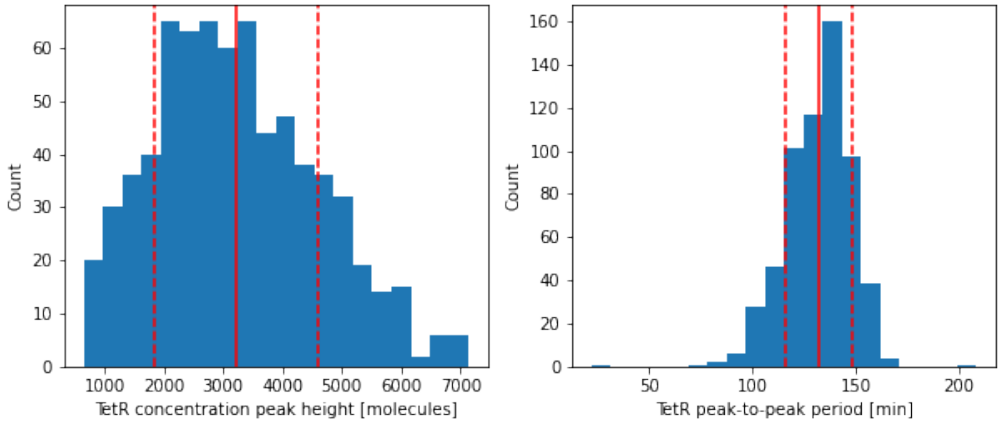


Fig. 19: **Oscillation period and peak height distributions across 100 stochastic simulations.** See Supplementary Information for simulations used for this analysis. Mean ± *s.d.* shown in red ($3221.6 \pm 1370.9$ [molecules], $132.0 \pm 16.1$ [min], respectively). For both, peak height and peak-to-peak period, the distribution observed resembles normal distribution – showing our successful stochastic simulations.

## Conclusion

In this report, I have discussed the Repressilator developed by Elowitz and Leibier around the year 2000 [4] — a foundational publication in synthetic biology. I have reviewed the original research and its future improvements, and introduced analysis of the Repressilator design, producing valuable insight into the circuits behaviour, which would be virtually impossible to produce merely with *in vitro* and *in vivo* experiments. *In silico* modelling can not only save vast amounts of time and resources by elevating the need for certain wet-lab experiments, but also allows us (especially when combined with synthetic biology) to gain much finer insight into how live organisms work with granularity unachievable *ex silico*.

Nevertheless, while being an extremely useful tool, computational modelling techniques are far from perfect and need to be always corroborated with other supporting evidence. It is especially important to pay attention to, in biology, ever-present, nonlinearities and stochastic randomness as (as also demonstrated in this report) they can/do dramatically influence the system behaviour, in this expertise from nonlinear dynamics, statistics is crucial.

Since the original Repressilator, scientists have developed countless new circuits capable of controlling cellular growth, image detection, event counting, and made significant progress in enhancing the circuit stability and predictability [2;9]. Hopefully, this and further effort will lead to a future where we are able to construct on-demand biological circuits with off-the-shelf items and program microorganisms much like computers in the present. This will bring unimaginable opportunities including *in vivo* drug synthesis and delivery, industrial synthesis, or climate control.

# Bibliography


[1] **Gonze D.** and **Abou-Jaoudé W.** (**2013**) *The Goodwin Model: Behind the Hill Function.* PLoS ONE, volume 8(8); p. 69 573. ISSN 19326203. https://doi.org/10.1371/journal.pone.0069573.

[2] **Purcell O.**, **Savery N.J.**, **Grierson C.S.**, *et al.* (**2010**) *A comparative analysis of synthetic genetic oscillators.* Journal of The Royal Society Interface, volume 7(52); pp. 1503–1524. ISSN 1742-5689. https://doi.org/10.1098/rsif.2010.0183.

[3] **Müller S.**, **Hofbauer J.**, **Endler L.**, *et al.* (**2006**) *A generalized model of the repressilator.* Journal of Mathematical Biology, volume 53(6); pp. 905–937. ISSN 03036812. https://doi.org/10.1007/s00285-006-0035-9.

[4] **Elowitz M.B.** and **Leibier S.** (**2000**) *A synthetic oscillatory network of transcriptional regulators.* Nature, volume 403(6767); pp. 335–338. ISSN 00280836. https://doi.org/10.1038/35002125.

[5] **Gonze D.** and **Ruoff P.** (**2020**) *The Goodwin Oscillator and its Legacy.* Acta Biotheoretica; pp. 1–18. ISSN 15728358. https://doi.org/10.1007/s10441-020-09379-8.

[6] **Strelkowa N.** and **Barahona M.** (**2010**) *Switchable genetic oscillator operating in quasi-stable mode.* Journal of the Royal Society Interface, volume 7(48); pp. 1071–1082. ISSN 17425662. https://doi.org/10.1098/rsif.2009.0487.

[7] **Andersen J.B.**, **Sternberg C.**, **Poulsen L.K.**, *et al.* (**1998**) *New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria.* Applied and Environmental Microbiology, volume 64(6); pp. 2240–2246. ISSN 00992240. https://doi.org/10.1128/aem.64.6.2240-2246.1998.

[8] **Potvin-Trottier L.**, **Lord N.D.**, **Vinnicombe G.**, *et al.* (**2016**) *Synchronous long-term oscillations in a synthetic gene circuit.* Nature, volume 538(7626); pp. 514–517. ISSN 14764687. https://doi.org/10.1038/nature19841.

[9] **Danino T.**, **Mondragón-Palomino O.**, **Tsimring L.**, *et al.* (**2010**) *A synchronized quorum of genetic clocks.* Nature, volume 463(7279); pp. 326–330. ISSN 00280836. https://doi.org/10.1038/nature08753.