Delay differential equation models of gene regulation

Matthew Macauley

Department of Mathematical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Algebraic Biology

Differential equations models of the *lac* operon

We will derive two ODE models of the *lac* operon: one with 3 variables, and another with 5 variables.

These models use Michaelis–Menten equations from mass-action kinetics.

Due to the time of transcription and translation, they will be delay differential equations.

They will also incorporate features of the operon such as:

- bistability
- dilution of protein concentration due to bacterial growth
- degredation (decay) of protein concentration
- time delays

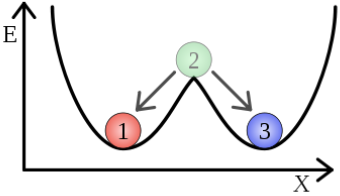
In general, bistable systems tend to have positive feedback loops in their "wiring diagrams" (variable dependancies).

A feedback loop with two negative interactions is considered positive.

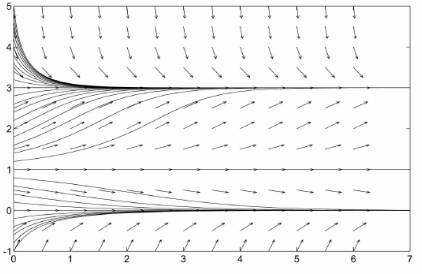
Bistability

A system is bistable if it has two stable steady-states.

Often, these are separated by an unstable steady-state.



From Wikipedia.



The *threshold ODE*: $y' = -ry\left(1 - \frac{y}{M}\right)\left(1 - \frac{y}{T}\right)$.

In the threshold model for population growth, there are three steady-states, $0 < T < M$:

- M = carrying capacity (stable),
- T = extinction threshold (unstable),
- 0 = extinct (stable).

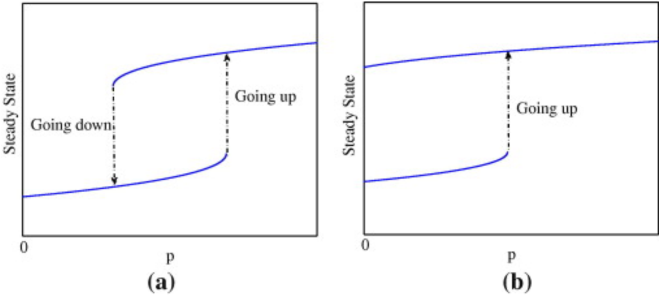
Types of bistability

The *lac* operon has been observed to exhibit bistability.

The expression level of the *lac* operon genes are either almost zero ("basal levels"), or very high (thousands of times higher). There's no "inbetween" state.

The exact level depends on the concentration of intracellular lactose. *Let's denote this parameter by p.*

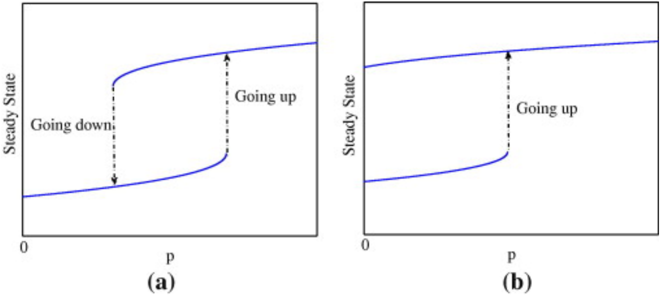
Now, let's "tune" this parameter. The result might look like the graph on the left.



This is reversible bistability. In other situations, it may be irreversible (at right).

Hysteresis

For reversible bistability, the *up-threshold* L_2 *of* p is higher than the *down-threshold* L_1 *of* p.



This is hysteresis: a dependence of a state on its current state *and* past state.

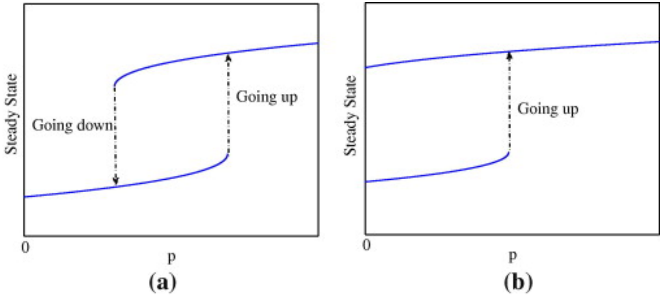
Weather example

Can we deduce what season it is just by the outdoor temperature at noon?

- If the outdoor temperature is $T < 40$, we know it's winter.
- If the outdoor temperature is $T > 90$, we know it's summer.
- But if the outdoor temperature is $T = 65$, we don't know whether it's spring or fall.

Hysteresis and the *lac* operon

If lactose levels are medium, then the state of the operon depends on whether or not a cell was grown in a lactose-rich environment.



Lac operon example

Let $[L] =$ concentration of intracellular lactose.

- If $[L] < L_1$, the operon is OFF.
- If $[L] > L_2$, the operon is ON.
- If $L_1 < [L] < L_2$, the operon might be ON or OFF.

In the region of bistability (L_1, L_2), one can find both induced and un-induced cells.

Modeling dilution in protein concentration due to bacterial growth

E. coli grows fast! It can double in 20 minutes. Thus, ODE models involving concentration can't assume volume is constant.

Let's define:

- $V =$ average volume of an *E. coli* cell.
- $x =$ number of molecules of protein X in that cell.

Assumptions:

- cell volume increases exponentially in time: $\frac{dV}{dt} = \mu V$.
- degradation of X is exponential: $\frac{dx}{dt} = -\beta x$.

The concentration of x is $[x] = \frac{x}{V}$. The derivative of this is (by the quotient rule):

$$\frac{d[x]}{dt} = \big(x'V - V'x\big)\frac{1}{V^2} = \big(-\beta xV - \mu Vx\big)\frac{1}{V^2} = -\big(\beta+\mu\big)\frac{x}{V} = -(\beta+\mu)[x].$$

Modeling of lactose repressor dynamics

Assumptions

- *Lac* repressor protein is produced at a constant rate.
- Laws of mass-action kinetics.

- *Repressor protein binds to allolactose*:

$$R + nA \underset{1}{\overset{K_1}{\rightleftharpoons}} RA_n \qquad \frac{d[RA_n]}{dt} = K_1[R][A]^n - [RA_n]$$

Assume the reaction is at equilibrium: $\frac{d[RA_n]}{dt} = 0$, and so $K_1 = \frac{[RA_n]}{[R][A]^n}$.

- *The repressor protein binds to the operator region if there is no allolactose*:

$$O + R \underset{1}{\overset{K_2}{\rightleftharpoons}} OR \qquad \frac{d[OR]}{dt} = K_2[O][R] - [OR].$$

Assume the reaction is at equilibrium: $\frac{d[OR]}{dt} = 0$, and so $K_2 = \frac{[OR]}{[O][R]}$.

Modeling of lactose repressor dynamics

Let O_{tot} = total operator concentration (a constant). Then, using $K_2 = \frac{[OR]}{[O][R]}$,

$$O_{tot} = [O] + [OR] = [O] + K_2[O][R] = [O](1 + K_2[R])\,.$$

Therefore, $\frac{[O]}{O_{tot}} = \frac{1}{1+K_2[R]}$. *"Proportion of free (unbounded) operator sites."*

Let R_{tot} be total concentration of the repressor protein (constant):

$$R_{tot} = [R] + [OR] + [RA_n]$$

Assume only a few molecules of operator sites per cell: $[OR] \ll \max\{[R], [RA_n]\}$:

$$R_{tot} \approx [R] + [RA_n] = [R] + K_1[R][A]^n$$

Eliminating $[RA_n]$, we get $[R] = \dfrac{R_{tot}}{1 + K_1[A]^n}$.

Now, the proportion of free operator sites is:

$$\frac{[O]}{O_{tot}} = \frac{1}{1 + K_2[R]} = \frac{1}{1 + K_2(\frac{R_{tot}}{1+K_1[A]^n})} \cdot \frac{1 + K_1[A]^n}{1 + K_1[A]^n} = \underbrace{\frac{1 + K_1[A]^n}{K + K_1[A]^n}}_{:=f([A])},$$

where $K = 1 + K_2 R_{tot}$.

Modeling of lactose repressor dynamics

Summary

The proportion of free operator sites is

$$\frac{[O]}{O_{tot}} = \underbrace{\frac{1 + K_1[A]^n}{K + K_1[A]^n}}_{:=f([A])}, \qquad \text{where } K = 1 + K_2 R_{tot}.$$

Remarks

- The function $f([A])$ is (almost) a Hill function of coefficient n.
- $f([A]=0) = \frac{1}{K} > 0$ "basal level of gene expression."
- f is *increasing* in $[A]$, when $[A] \geqslant 0$.
- $\lim_{[A]\to\infty} f([A]) = 1$ "with lots of allolactose, gene expression level is max'ed."

Modeling time-delays

The production of mRNA from DNA via transcription is not instantaneous; suppose it takes time $\tau > 0$.

Thus, the production rate of mRNA is not a function of allolactose at time t, but rather at time $t - \tau$.

Suppose protein P decays exponentially, and its concentration is $p(t)$.

$$\frac{dp}{dt} = -\mu p \quad \Longrightarrow \quad \int_{t-\tau}^{t} \frac{dp}{p} = -\mu \int_{t-\tau}^{t} dt\,.$$

Integrating yields

$$\ln p(t)\Big|_{t-\tau}^{t} = -\mu t\Big|_{t-\tau}^{t} dt = \ln \frac{p(t)}{p(t-\tau)} = -\mu[t - (t-\tau)] = -\mu\tau.$$

Exponentiating both sides yields $\frac{p(t)}{p(t-\tau)} = e^{-\mu\tau}$, and so

$$p(t) = e^{-\mu\tau}\, \underbrace{p(t-\tau)}_{:=p_\tau}\,.$$

A 3-variable ODE model of the *lac* operon

Consider the following 3 quantities, which represent *concentrations* of:

- $M(t) =$ mRNA,
- $B(t) = \beta$-galactosidase,
- $A(t) =$ allolactose.

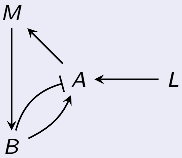
Assumption: Internal lactose (L) is available and is a parameter.

The model (Yildirim and Mackey, 2004)

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B$$

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A$$



These are *delay differential equations*, with discrete time delays due to the transcription and translation processes.

There should (?) be a self-loop  at M, B, and A, but we're omitting them for clarity.

A 3-variable ODE model of the *lac* operon

ODE for β-galactosidase (B)

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B,$$

Justification:

- $\tilde{\gamma}_B B = \gamma_B B + \mu B$ represents loss due to β-galactosidase degredation and dilution from bacterial growth.
- Production rate of β-galactosidase, is proportional to mRNA concentration.
- τ_B = time required for translation of β-galactosidase from mRNA, and $M_{\tau_B} := M(t - \tau_B)$.
- $e^{-\mu\tau_B} M_{\tau_B}$ accounts for the time-delay due to translation.

A 3-variable ODE model of the *lac* operon

ODE for mRNA (M)

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M$$

Justification:

- $\tilde{\gamma}_M M = \gamma_M M + \mu M$ represents loss due to mRNA degradation and dilution from bacterial growth.
- Production rate of mRNA [=expression level!] is proportional to the fraction of free operator sites,
$$\frac{[O]}{O_{tot}} = \frac{1 + K_1 A^n}{K + K_1 A^n} = f(A).$$
- τ_M = time required for transcription of mRNA from DNA, and $A_{\tau_M} := A(t - \tau_M)$.
- The term $e^{-\mu\tau_M} A_{\tau_M}$ accounts for the time-delay due to transcription.

A 3-variable ODE model of the *lac* operon

ODE for allolactose (A)

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A$$

Justification:

- $\tilde{\gamma}_A A = \gamma_A A + \mu A$ represents loss due to allolactose degredation and dilution from bacterial growth.
- The first two terms models the enzyme-substrate reactions involving the enzyme β-galactosidase.
  1. *Lactose into allolactose*:
     $$L + B \rightleftharpoons LB \longrightarrow A + B$$
     has solution $\dfrac{d[A]}{dt} = \dfrac{\alpha_A B[L]}{K_L + [L]}$.
  2. *Allolactose into glucose and galactose (both $C_6H_{12}O_6$)*:
     $$A + B \rightleftharpoons AB \longrightarrow B + Glu + Gal.$$
     has solution $\dfrac{d[Glu]}{dt} = \dfrac{d[Gal]}{dt} = \dfrac{\beta_A B[A]}{K_A + [A]} = -\dfrac{d[A]}{dt}$.

A 3-variable ODE model of the *lac* operon

Steady-state analysis

To find the steady states, we must solve the nonlinear system of equations:

$$\begin{aligned}
0 &= \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M \\
0 &= \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B \\
0 &= \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A
\end{aligned}$$

This was done by Yildirim et al. (2004). They set $L = 50 \times 10^{-3}$ mM, which was in the "bistable range."

They estimated the parameters through an extensive literature search.

Finally, they estimated $\mu = 3.03 \times 10^{-2}$ min^{-1} by fitting ODE models to experimental data.

Steady states	A^* (mM)	M^* (mM)	B^* (mM)	
I.	4.27×10^{-3}	4.57×10^{-7}	2.29×10^{-7}	basal (stable)
II.	1.16×10^{-2}	1.38×10^{-6}	6.94×10^{-7}	medium (unstable)
III.	6.47×10^{-2}	3.28×10^{-5}	1.65×10^{-5}	high (stable)

One-parameter bifurication diagram of the 3-variable ODE model

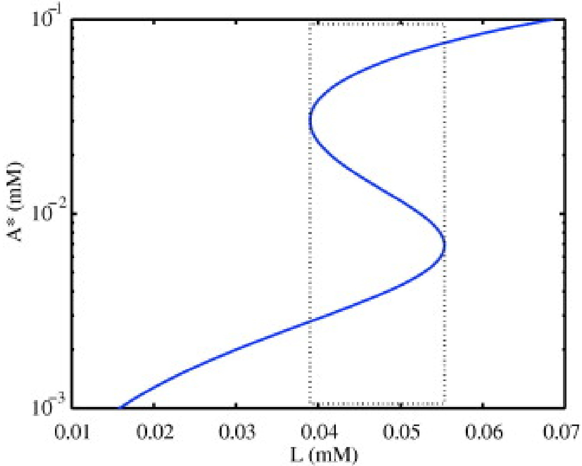


Figure: The fixed points of the allolactose concentration A^* in ODE model (6.47×10^{-2}, 1.16×10^{-2}, and 4.27×10^{-3}mM) as a function of the parameter L (lactose). For a range of L concentrations, there are 2 stable steady states, the phenomenon of bistability.

3-variable ODE model

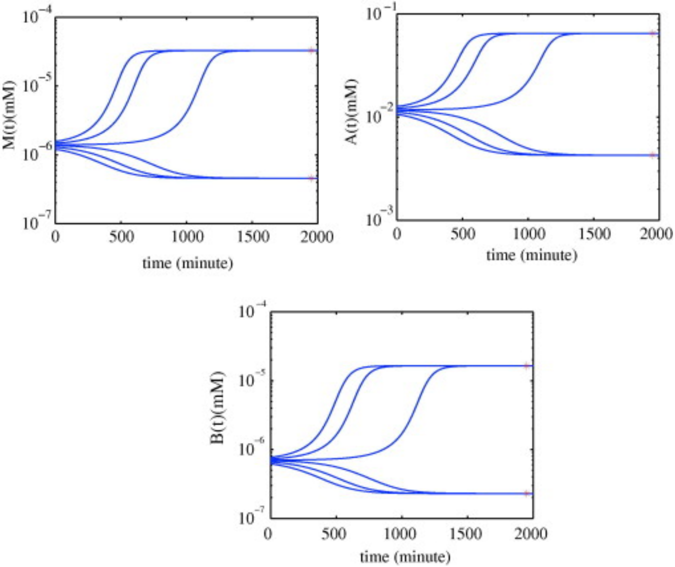


Figure: Numerical solutions of $M(t)$ (mRNA), $B(t)$ (β-galactosidase), and $A(t)$ (allolactose), using $L = 50 \times 10^{-3}$.

5-variable ODE model

Consider the following 5 variables, which represent *concentrations* of:

- $M(t) =$ mRNA,
- $B(t) = \beta$-galactosidase,
- $A(t) =$ allolactose.
- $P(t) =$ lac permease.
- $L(t) =$ intracellular lactose.

The model (Yildirim and Mackey, 2004)

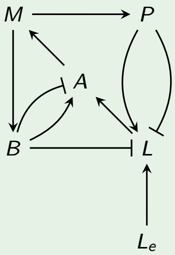
$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} + \Gamma_0 - \widetilde{\gamma}_M M$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \widetilde{\gamma}_B B$$

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \widetilde{\gamma}_A A$$

$$\frac{dP}{dt} = \alpha_P e^{-\mu(\tau_B+\tau_P)} M_{\tau_B+\tau_P} - \widetilde{\gamma}_P P$$

$$\frac{dL}{dt} = \alpha_L P \frac{L_e}{K_{L_e} + L_e} - \beta_{L_e} P \frac{L}{K_{L_e} + L} - \alpha_A B \frac{L}{K_L + L} - \widetilde{\gamma}_L L$$



5-variable ODE model

$$\begin{aligned}
\frac{dM}{dt} &= \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} + \Gamma_0 - \tilde{\gamma}_M M \\
\frac{dB}{dt} &= \alpha_B e^{-\mu\tau_B} M_{\tau_B} - \tilde{\gamma}_B B \\
\frac{dA}{dt} &= \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{A}{K_A + A} - \tilde{\gamma}_A A \\
\frac{dP}{dt} &= \alpha_P e^{-\mu(\tau_B + \tau_P)} M_{\tau_B + \tau_P} - \tilde{\gamma}_P P \\
\frac{dL}{dt} &= \alpha_L P \frac{L_e}{K_{L_e} + L_e} - \beta_{L_e} P \frac{L}{K_{L_e} + L} - \alpha_A B \frac{L}{K_L + L} - \tilde{\gamma}_L L
\end{aligned}$$

ODEs for M, B, A, and P

- The only difference in the ODE for M is the extra term Γ_0 which describes the basal transcription rate ($L_e = 0$).
- The ODEs for B and A are the same as in the 3-variable model.
- The ODE for P is very similar to the one for B:
  - production rate of *lac* permease $\propto$ mRNA concentration, with a time-delay.
  - the 2nd term accounts for loss due to degredation and dilution.

5-variable ODE model

ODE for lactose (L)

$$\frac{dL}{dt} = \alpha_L P \frac{L_e}{K_{L_e} + L_e} - \beta_{L_e} P \frac{L}{K_{L_1} + L} - \alpha_A B \frac{L}{K_L + L} - \tilde{\gamma}_L L,$$

Justification:

- The first term models the transport of lactose by *lac* permease **into** the cell:

$$L_e + P \rightleftharpoons PL_e \longrightarrow P + L$$

- The second term models the transport lactose by *lac* permease **out** of the cell:

$$L + P \rightleftharpoons PL \longrightarrow P + L_e$$

- The 3rd term describes the reaction of Lactose into allolactose catalyzed by β-galactosidase:

$$L + B \rightleftharpoons LB \longrightarrow A + B$$

- the 4th term accounts for loss due to degredation and dilution.

A 5-variable ODE model

To find the steady states, we set $M' = A' = B' = L' = P' = 0$ and solve the resulting nonlinear system of equations.

This was done by Yildirim et al. (2004). They set $L_e = 50 \times 10^{-3}$ mM, in the "bistable range."

They also estimated the parameters through an extensive literature search.

Finally, they estimated $\mu = 2.26 \times 10^{-2}$ min^{-1} by fitting the ODE models to experimental data.

Fixed point	A^* (nM)	M^* (mM)	B^* (mM)	L^* (mM)	P^* (mM)
High (stable)	3.10×10^{-1}	5.80×10^{-4}	3.92×10^{-4}	2.30×10^{-1}	8.09×10^{-3}
Med (unstable)	2.64×10^{-2}	7.58×10^{-6}	5.13×10^{-6}	2.06×10^{-1}	1.05×10^{-4}
Low (stable)	7.85×10^{-3}	2.48×10^{-6}	1.68×10^{-6}	1.69×10^{-1}	3.46×10^{-5}

One-parameter bifurication diagram of the 5-variable ODE model

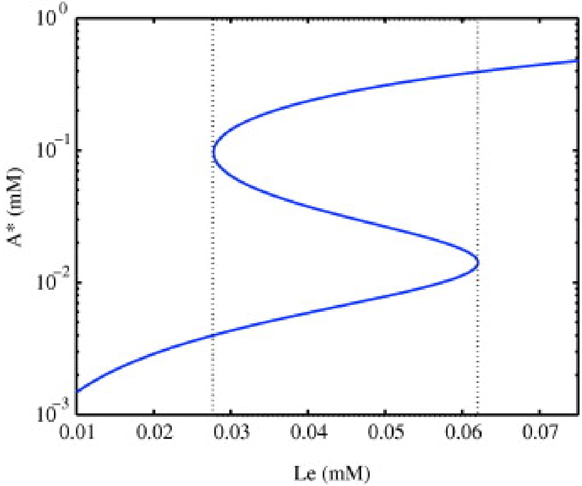


Figure: The fixed points of the allolactose concentration A^* in ODE model (3.10×10^{-1}, 2.64×10^{-2}, and 7.85×10^{-3}mM) as a function of the parameter L_e (external lactose). For a range of L concentrations, there are 2 stable steady states, the phenomenon of bistability.

5-variable ODE model

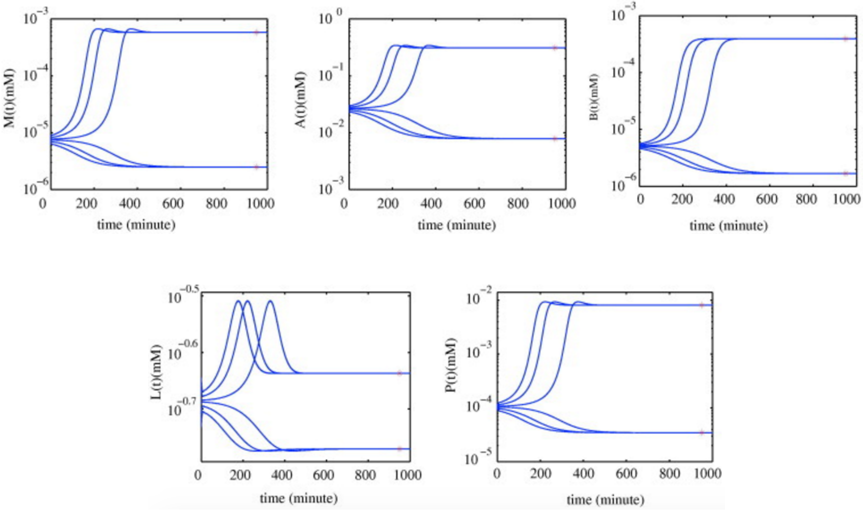


Figure: Numerical solutions of mRNA, β-galactosidase, allolactose, *lac* permease, and lactose concentrations, using $L_e = 50 \times 10^{-3}$.

A model of the arabinose (*ara*) operon (Yildirim, 2012)

Variables. $A(t)$, $E(t)$, and $F(t)$ are concentrations of intracellular arabinose, *araE* mRNA, and *araFGH* mRNA, respectively.

Constants.

- A_e is concentration of extracellular arabinose.
- γ_A, γ_E, and γ_F are degradation rates.
- μ describes loss of concentration due to cell growth.
- V_E, V_F, V_{m_E} V_{m_F}, and K_E, K_F, K_{m_E}, and K_{m_F} arise from Michaelis-Menten functions.

Model.

$$\begin{aligned}
A'(t) &= \frac{A_e V_E E(t)}{K_E + A_e} + \frac{A_F V_F F(t)}{K_F + A_e} - (\mu + \gamma_A)A(t) \\
E'(t) &= \alpha_E + \frac{V_{m_E}(A(t))^n}{K_{m_E}^n + (A(t))^n} - (\mu + \gamma_E)E(t) \\
F'(t) &= \alpha_F + \frac{V_{m_F}(A(t))^n}{K_{m_F}^n + (A(t))^n} - (\mu + \gamma_F)F(t).
\end{aligned}$$

A model of the tryptophan (*trp*) operon (Santillán/Mackey, PNAS 2001)

Model.

$$\begin{aligned}
O_F'(t) &= \frac{K_r}{K_r + R_A(T)}\big(\mu O - k_p P[O_F(t) - O_F(t-\tau_p)e^{-\mu\tau_p}]\big) - \mu O_F(t) \\
M_F'(t) &= k_p P O_F(t-\tau_m)e^{-\mu\tau_m}(1 - A(T)) - k_\rho\rho\big[M_F(t) - M_F(t-\tau_p)e^{-\mu\tau_\rho}\big] - (k_d D + \mu)M_F(t) \\
E'(t) &= \frac{1}{2}k_\rho\rho M_F(t-\tau_e)e^{-\mu\tau_e} - (\gamma+\mu)E(t) \\
T'(t) &= KE_A(E,T) - G(T) + F(T,T_{ext}) - \mu T(t)
\end{aligned}$$

$$A(t) = b(1 - e^{-T(t)/c}), \qquad R_A(T) = \frac{T(t)}{T(t) + K_t}R, \qquad G(T) = g\,\frac{T(t)}{T(t) + K_g}.$$

$$E_A(E,T) = \frac{K_i^{n_H}}{K_i^{n_H} + T^{n_H}(t)}E(t), \qquad F(T,T_{ext}) = d\,\frac{T_{ext}}{e + T_{ext}[1 + T(t)/f]}.$$

A model of the tryptophanse (*tna*) operon (Orozco-Gómez et al., 2019)

Variables. $A(t)$, $B(t)$, and $W(t)$ are concentrations of tryptophanase (TnaA), the TnaB permease, and intracellular tryptophan.

Constants.

- W_e and G_e are concentrations of extracellular tryptophan and glucose.
- k_A and k_B are rate constants from mass-action kinetics.
- γ_A and γ_B model protein degradation; μ models dilution from cellular growth.
- P_A is a sigmoidal function that accounts for catabolite repression.

Model.

$$\begin{aligned} A' &= k_A P_G(G_e) P_W(W) - (\gamma_A + \mu)A \\ B' &= k_B P_G(G_e) P_W(W) - (\gamma_B + \mu)B \\ W' &= (\alpha + \beta B)W_e - (\delta + \epsilon A P_A(G_e, W_e) + \mu)W. \end{aligned}$$

This model suggests that glucose and tryptophan regulate TnaA via a common signaling pathway.

Experimental results suggest that it exhibits bistability; this model provides further evidence.

A Boolean model (I. Deal et al., 2023) of this operon also showed bistabilty.

DDE mathematical models of biological systems is transdisplinary!

The researchers involved in this work have diverse backgrounds in math, science, and engineering.

Necmettin Yildirim is an applied mathematician.

Michael Mackey has a PhD in Physiology and Biophysics. He and **Leon Glass** (PhD Chemistry) developed the the Mackey-Glass equations that model blood cells:

$$\frac{dP(t)}{dt} = \frac{\beta_0 \theta^n}{\theta^n + P(t-\tau)^n} - \gamma P(t), \qquad \text{and} \qquad \frac{dP(t)}{dt} = \frac{\beta_0 \theta^n P(t-\tau)}{\theta^n + P(t-\tau)^n} - \gamma P(t).$$

Moisés Santillán has PhD in Physics.

Some of the co-authors of the *tna* operon model paper are in Biomedial Engineering and Physics.

- The Mackey-Glass equations were published in *Science*.
- The *trp* operon model was published in *Proc. Natl. Acad. Sci*.
- *lac* operon models (Yildirim/Mackey, and Yildirim et al.) were published in *Biophys J.* and *Chaos* and *J R Soc Interface*.
- The *ara* operon model was published in *Mol. Biosyst*.

References

- M.C. Mackey, L. Glass, Oscillation and chaos in physiological control systems. *Science*. **197**(4300) 287–289, 1977.
- M. Santillán and M.C. Mackey. Dynamic regulation of the tryptophan operon: a modeling study and comparison with experimental data. *Proc. Natl. Acad. Sci.* **98**(4), 1364–69, 2001.
- N. Yildirim and M.C. Mackey. Feedback regulation in the lactose operon: a mathematical modeling study and comparison with experimental data. *Biophys. J.* **84**(5):2841–51, 2003.
- N. Yildirim, M. Santillan, D. Horike, and M.C. Mackey. Dynamics and bistability in a reduced model of the lac operon. *Chaos*. **14**(2) 279–292, 2004.
- M. Santillán, M.C. Mackey, and E.S. Zeron. Origin of bistability in the *lac* operon. *Biophys. J.* **92**(11), 3830–42, 2007.
- M. Santillán and M.C. Mackey. Quantitative approaches to the study of bistability in the *lac* operon of *Escherichia coli*. *J. Royal Soc. Interface*. **5**(suppl_1), S29–39, 2008.
- N. Yildirim. Mathematical modeling of the low and high affinity arabinose transport systems in *Escherichia coli*. *Mol. Biosyst.* **8**(4), 1319–24, 2012.
- R. Robeva and N. Yildirim. *Bistability in Lactose Operon of Escherichia coli: A Comparison of Differential Equation and Boolean Network Models*. Chapter 2 (pp. 37–74) of R. Robeva and T. Hodge. Mathematical Concepts and Methods in Modern Biology. Academic Press, 2013.
- D.I. Orozco-Gómez, J.E. Sosa-Hernández, O.A. Gallardo-Navarro, J. Santana-Solano, and M. Santillán. Bistable behaviour and medium-dependent post-translational regulation of the tryptophanase operon regulatory pathway in *Echerichia coli*. *Scientific Reports* **9**(5451), 13 pp., 2019.