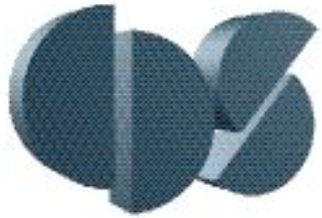
APh 161, Lecture 6: Feedback and Dynamics in Cells



Richard M. Murray
Control and Dynamical Systems

23 January 2007

Goals:

- Motivate the role of dynamics and feedback in the the cell
- Provide mathematical tools for analyzing and predicting the dynamics of transcriptional regulation in the cell
- Work through case study for the *lac* operon (model system)

Today: give overview of main concepts and approach

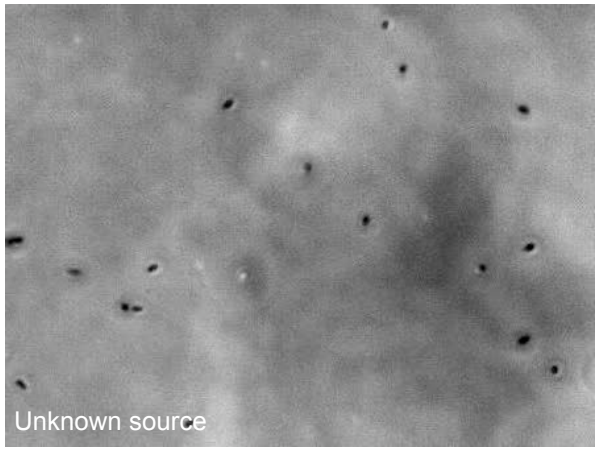
Wed: *optional* tutorial

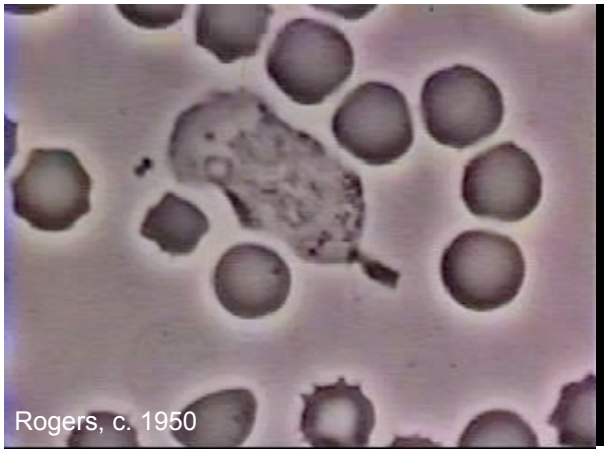
Thu: "stick in sand" + "fingers on keyboard"

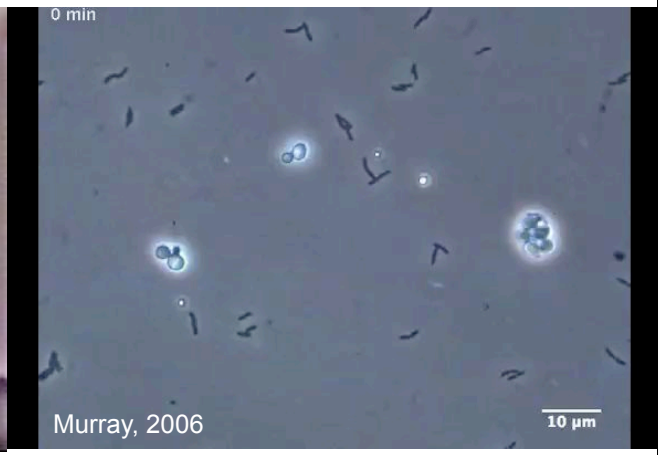
Reading

- Kondev, Phillips and Theriot, Physical Biology of the Cell, Chapter 19
- N. Yildirim, M. Santillan, D. Horiki and M. C. Mackey, "Dynamics and bistability in a reduced model of the *lac* operon". Chaos, 14(2):279-292, 2004.

Some Examples of Biological Dynamics and Feedback







Questions

- How do bacteria “decide” which way to swim (run/tumble) to find food?
- How does a neutrophil “sense, compute and actuate” to control its motion?
- What controls the stages and rate of the cell cycle?

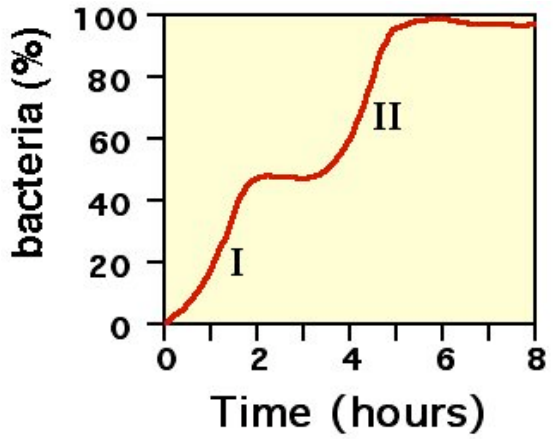
This week: develop tools for answering these questions through analysis of molecular scale dynamics and feedback

Approach

- Focus on transcriptional regulation, building on stat mech story
- Make use of modern tools from dynamical systems to quantify and classify dynamic behavior
- Combination of “stick in the sand” and “fingers on the keyboard” techniques

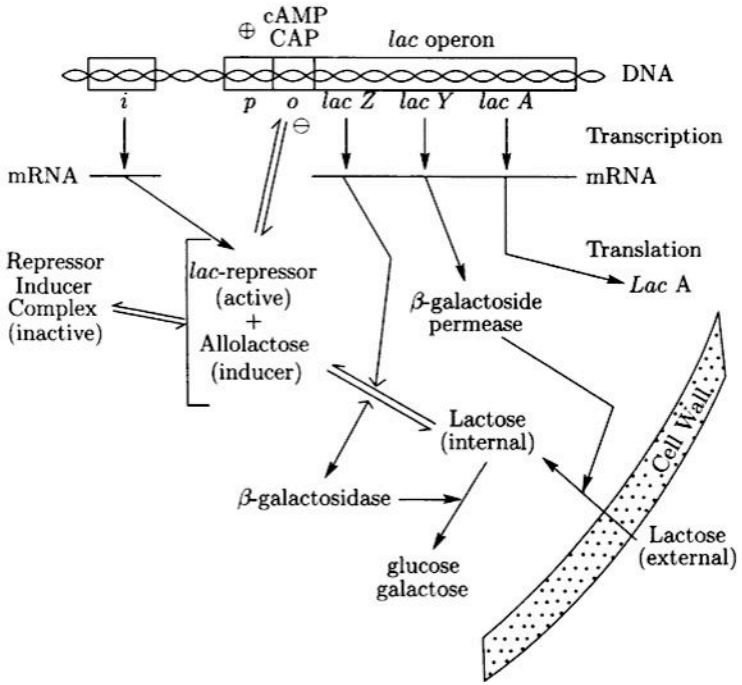
Model System: The Lac Operon

http://en.wikipedia.org/wiki/Image:Bacterial_growth_monod.png



Key idea (Jacob & Monod, 1950s): produce proteins when you need them

- Bacterial growth dependent on nutrient environment
- Lactose only consumed if glucose is *not* present
- Q: How does *E. coli* decide when to produce proteins?
- A: Lac "operon" (≈ control system)
- KPT07: "Lac operon is the hydrogen atom of regulation"

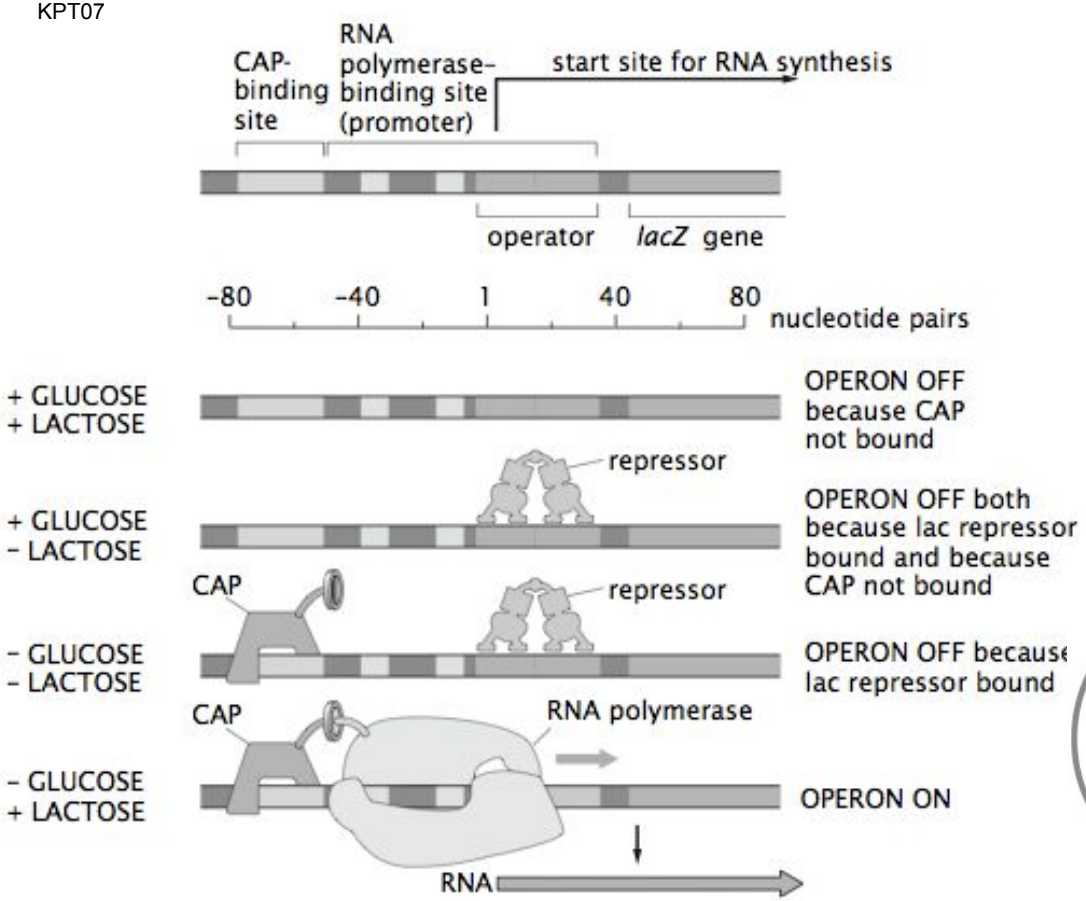


How does it work?

- Proteins for digesting lactose are controlled by binding of repressor and activator
- CAP (activator): recruits RNA polymerase when bound to cAMP, whose concentration is controlled by absence of glucose (+glucose ⇒ no cAMP)
- Repressor: blocks transcription by causing DNA looping unless it is bound to allolactose, a byproduct of lactose metabolism
- Feedback: if lactose present, then create proteins required to metabolise lactose, which turns on lac operon (required to matabolise lactose)

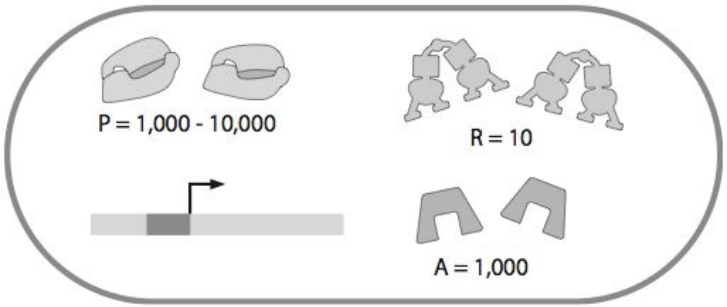
Operon Layout and Census

KPT07

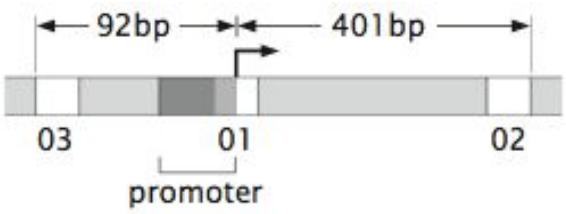


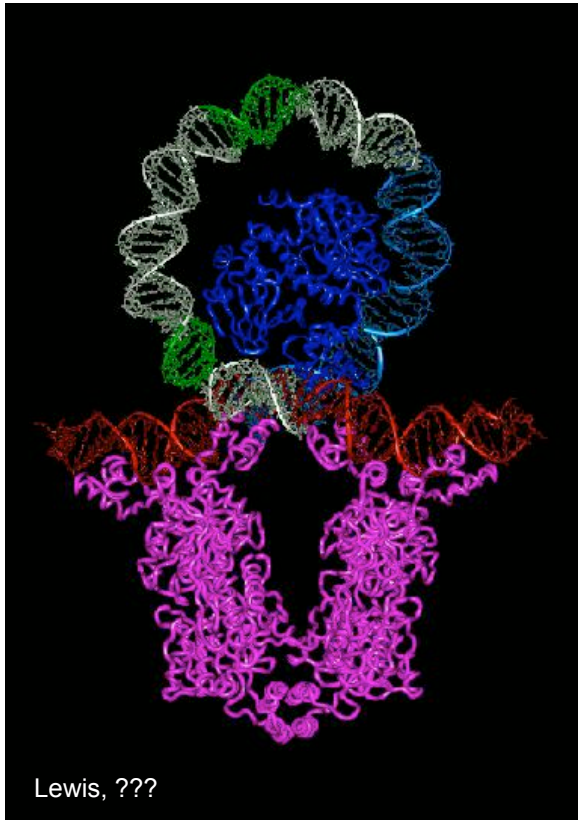
The regulatory landscape

- Repressor binds in a region that blocks the binding of RNAP
- If CAP is present (activated by cAMP, which is present in the absence of glucose), it recruits RNAP
- Can assign energies to each of these events and work out statistical mechanics (see text)



Repressor and DNA Looping




Lewis, ???

Repressor can act through DNA looping

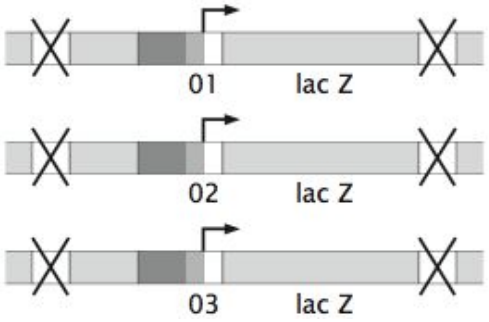
- Binding at secondary sites on DNA (eg, within *lacZ*)

Can use statistical mechanics to predict repression

- Analysis: compute Pbound of RNA for different situations

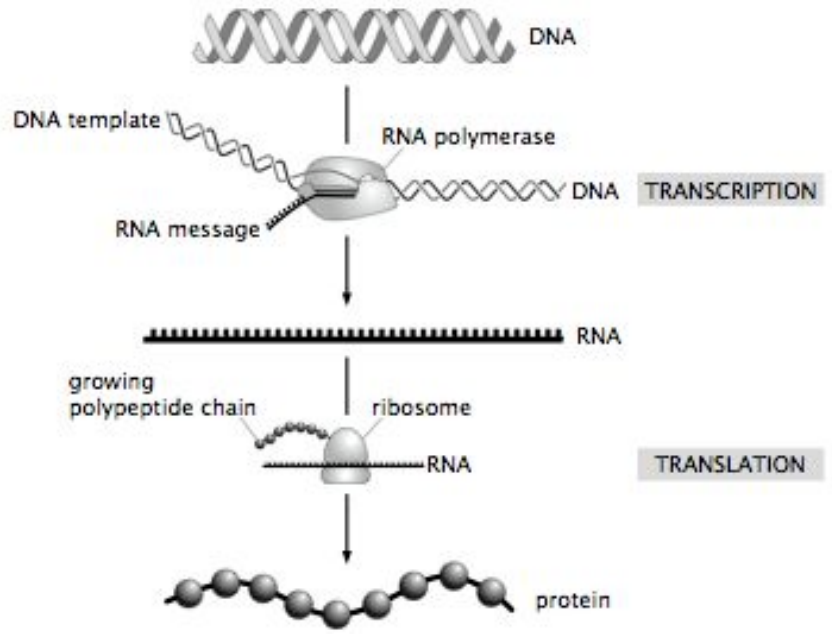
$$\text{Repression} = \frac{p_{\text{bound}}(R=0)}{p_{\text{bound}}(R)}$$

- Experiments: knock out secondary binding sites and see what happens
- Still get repression, with O1 being the strongest factor



repression		
50	R	900
200		4700
21		320
1.3		16

Rates of Transcriptional Regulation



Primary timescales:

- DNA production: 250-1000 bp/sec
- mRNA production: 10-30 bp/sec
- Protein production: 10-30 aa/sec

DNA, protein from KPT07, mRNA production from Vogel & Jensen

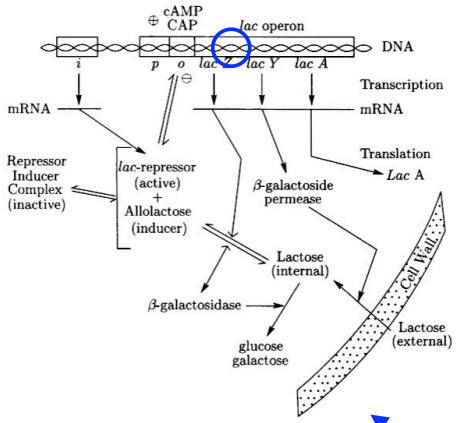
Other important rates

- mRNA half life: ~100 sec
- Protein half life: ~5 x 10^4 sec
- Protein diffusion (along DNA): up to 10^4 bp/sec

- Assume that activators and repressors reach equilibrium state much more quickly than other time scales

Half life times from Yildirim and Mackey, 2003; Protein diffusion from Blainey et al, PNAS 2006.

From Numbers to Equations: Dynamical Modeling

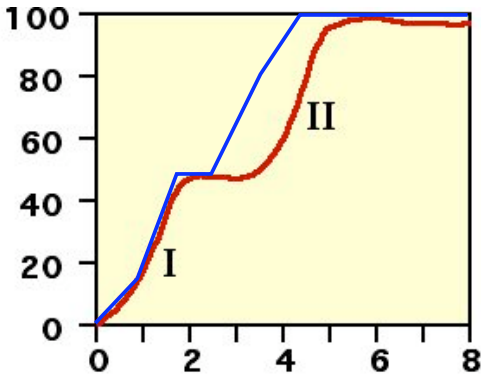


$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A_{\tau_M})^n}{K + K_1(e^{-\mu\tau_M} A_{\tau_M})^n} - \tilde{\gamma}_M M,$$

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

and

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



Modeling philosophy: Ask questions first, build model later

- Many different models possible for the same system; no such thing as “the model”
- The model you use depends on the questions you want to answer
- Never build a model without first posing the questions

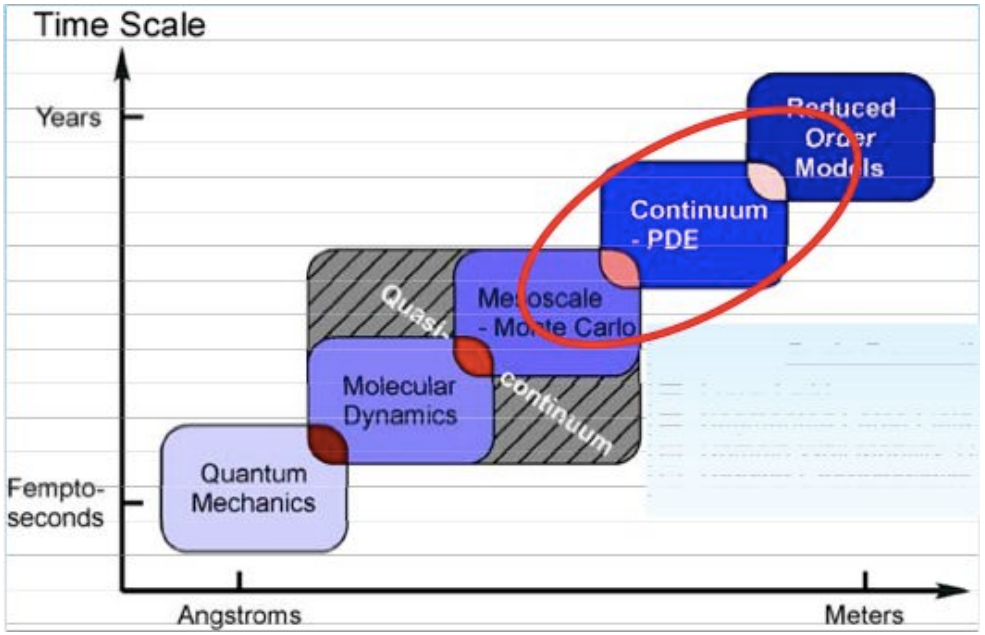
Analysis and design based on *models*

- A model provides a *prediction* of how the system will behave
- Feedback can give counter-intuitive behavior; models help sort out what is going on
- Models don’t have to be exact to be useful; they just need to help explain (and predict)

Dynamic Modeling Approaches

Possible approaches to modeling

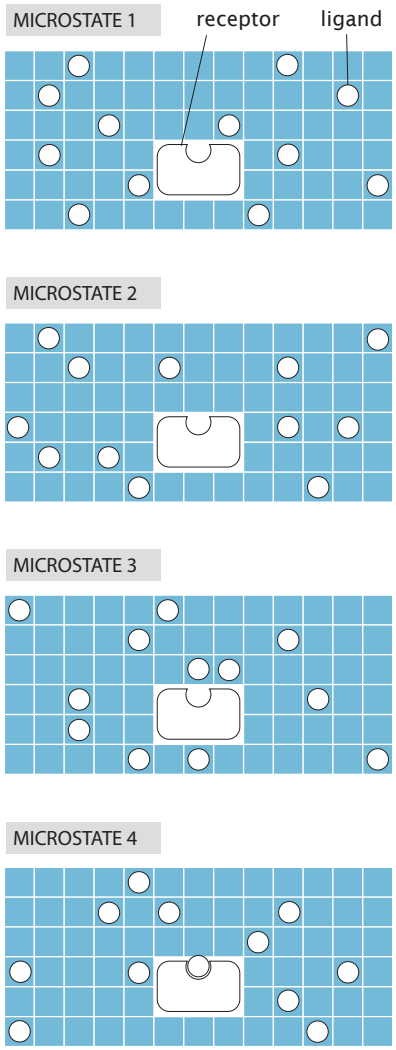
- Molecular dynamics - keep track of vibration of molecules and detailed reaction dynamics
- Monte Carlo/Stochastic simulation - extend ideas from statistical mechanics to include time
- Continuum/partial differential equations - keep track of evolution in space and time
- Reduced order models - ordinary differential equations that capture bulk properties

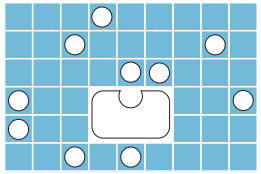
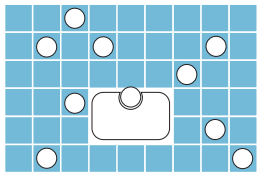


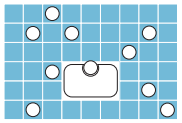
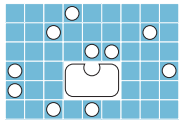
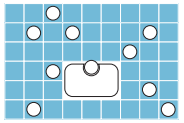
Choice of model depends on the questions you want to answer

- Modeling takes practice, experience and iteration to decide what aspects of the system are important to model at different temporal and spatial scales
- Good models make testable predictions and produce “surprising” results

Statistical Mechanics: The "Analytical Engine"



STATE	ENERGY	MULTIPLICITY	WEIGHT
	$L\varepsilon_{solution}$	$\frac{N!}{L!(N-L)!} \approx \frac{N^L}{L!}$	$\frac{N^L}{L!}\, e^{-\beta L \varepsilon_{solution}}$
	$(L-1)\varepsilon_{solution} + \varepsilon_{bound}$	$\frac{N!}{(L-1)!(N-L+1)!} \approx \frac{N^{L-1}}{(L-1)!}$	$\frac{N^{L-1}}{(L-1)!}\, e^{-\beta[(L-1)\varepsilon_{solution} + \varepsilon_{bound}]}$

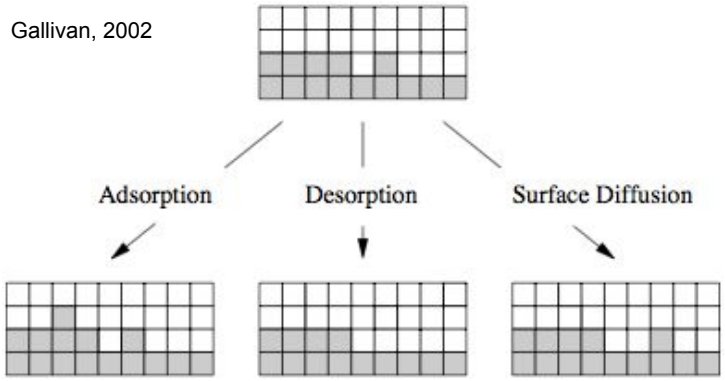
$$\text{pbound} = \frac{\sum_{\text{states}} \left(\text{  } \right)}{\sum_{\text{states}} \left(\text{  } \right) + \sum_{\text{states}} \left(\text{  } \right)}$$

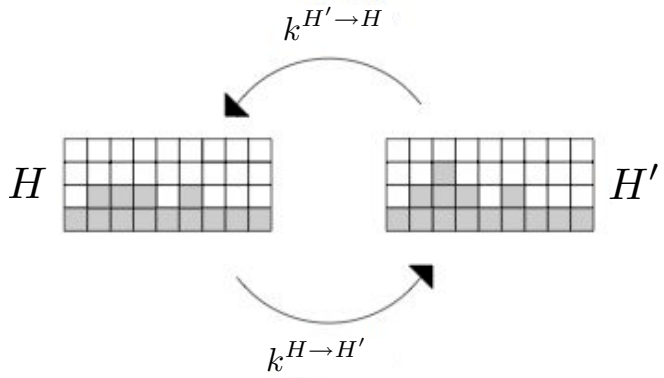
Provides an *equilibrium* view of the world

- Computes the "steady state" probability that a situation occurs
- Based on energy arguments; allows study of complex situations

The Master Equation: Detailed Events

Gallivan, 2002





$$\frac{d}{dt}P_H(t) = \sum_{H'} k^{H' \to H} P_{H'}(t) - \sum_{H'} k^{H \to H'} P_H(t)$$

Key idea: transition rates between microstates

- Enumerate micro-states corresponding to the system of interest
- Define the system "state" in terms of the individual probabilities of each microstate at each instant in time
- Dynamics are given by the rate of change of probability of each individual state

Transition rates depend on Boltzmann model

- Define the individual rates of transition based on the difference in energy between the states

$$k_{\mathrm{des},i} = \nu \exp\left(-\frac{E_{\mathrm{des},0} + i\Delta E}{k_b T}\right)$$

- v = vibrational frequency (10^{13})
- Assumes at most two species interacting at a time

Simulating the Master Equation for Chemical Reactions

Stochastic Simulation Algorithm (SSA)

- N states ($X_1, X_2, ..., X_N$) where X_i is the number of copies of molecule S_i in the system
- M reaction channels R_i that define changes in the state. $v_{ij}(X)$ = change in X_i for R_j
- Propensity function: $a_i(X)\,dt$ = probability that reaction i will occur in time interval dt
- "Gillespie algorithm": determine how many reactions occur in a given time step and execute them.
- Choose time steps to be small enough that propensity functions are roughly constant

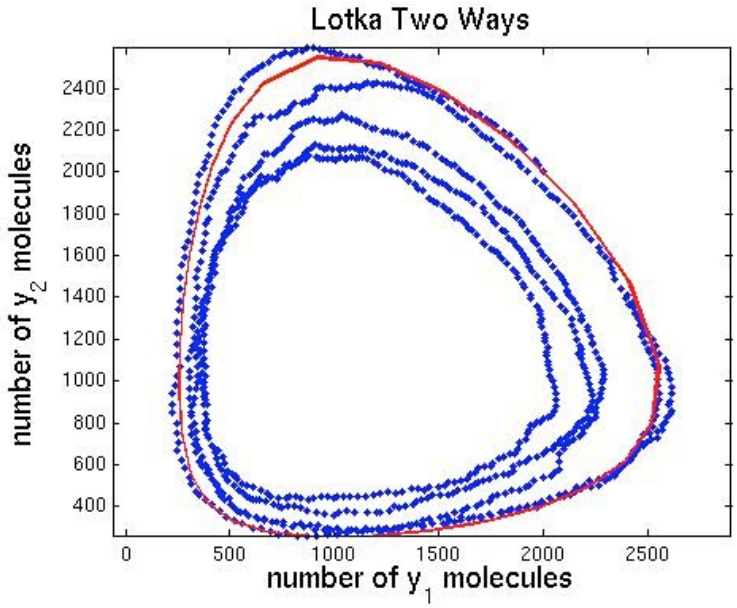
Example

$$X_\infty + Y_1 \rightarrow 2Y_1 \quad \text{with propensity } c_1$$
$$Y_1 + Y_2 \rightarrow 2Y_2 \quad \text{with propensity } c_2$$
$$Y_2 \rightarrow Z \quad \text{with propensity } c_3$$

- Propensity function: $a_1(X) = c_1 X_\infty Y_1$ to account for multiple copies of each species

Tools

- StochKit (Linda Petzold) - C++ libraries
- MATLAB - SimBiology (includes SSA)



http://www.caam.rice.edu/~caam210/reac/lec.html

Chemical Kinetics: The Law of Mass Action

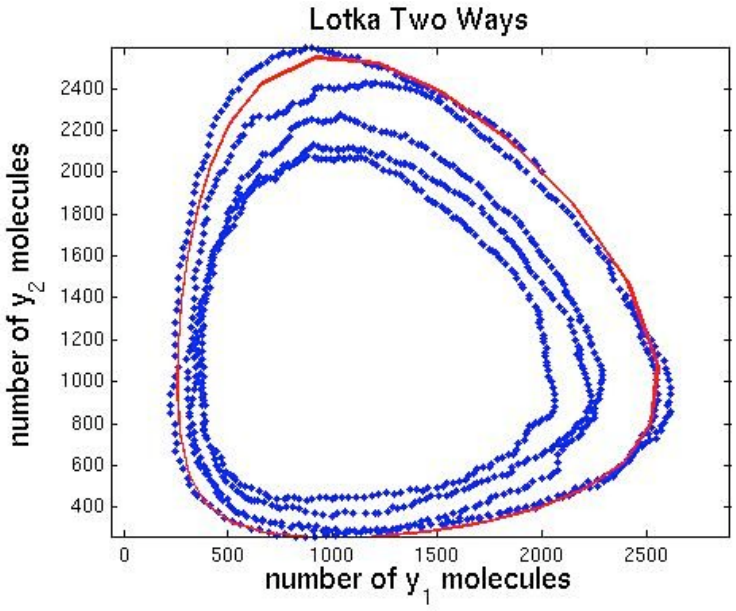
Alternative approach: keep track of concentrations

- If number of molecules is large, we can keep track of concentration of each species
- No longer track individual events; assume an average rate of events and use ODEs

$$X_\infty + Y_1 \rightarrow 2Y_1 \quad \text{with propensity } c_1$$
$$Y_1 + Y_2 \rightarrow 2Y_2 \quad \text{with propensity } c_2$$
$$Y_2 \rightarrow Z \quad \text{with propensity } c_3$$

$$d[Y_1]/dt = c_1[X_\infty][Y_1] - c_2[Y_1][Y_2]$$

$$d[Y_2]/dt = c_2[Y_1][Y_2] - c_3[Y_2]$$



Michaelis Menten kinetics

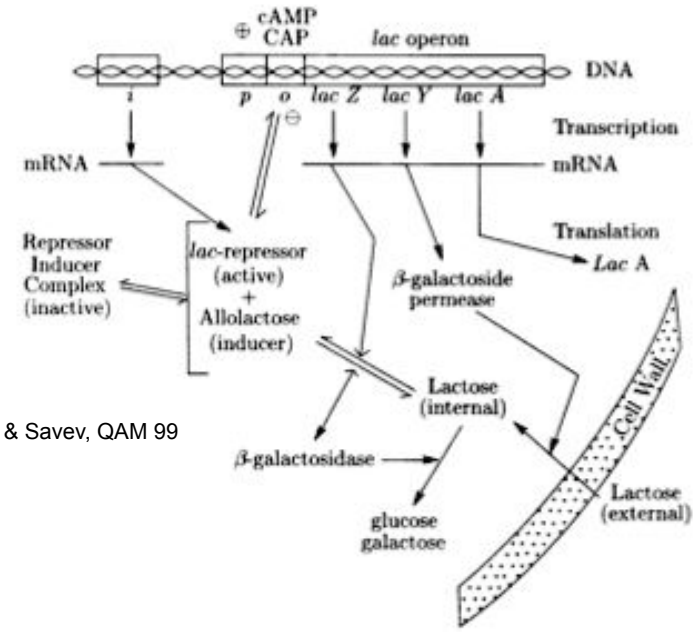
- Model for enzyme controlled reactions

$$E + S \rightleftarrows ES \rightarrow E + P$$

- Assume first reaction is fast compared to the second

$$K_d = \frac{[E][S]}{[ES]} \qquad \frac{d[P]}{dt} = k_p E_0 \frac{[ES]}{[E]+[ES]} = k_p E_0 \frac{[E][S]/K_d}{[E]+[E][S]/K_d} = V_{\max} \frac{[S]}{K_d + [S]}$$

Yildirim-Mackey Model for the Lac Operon



Mahaffy & Savev, QAM 99

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

$$\frac{dB}{dt} = \alpha_B e^{\mu\tau_B} M(t-\tau_B) - \tilde{\gamma}_B B$$

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

Questions:

- In the absence of glucose, what concentration of lactose is required for the *lac* operon to become "active"?
- Focuses on "bistability": *lac* operon has two stable equlibrium points:
  - low lactose: machinery off
  - high lactose: machinery on

Model

- Ordinary differential equation for rates of transcription, translation and degradation of β-galactosidase (β-gal) and allolactose
- Assume levels of lactose outside the cell is constant and level of permease (from *lacY*) is constant to simplify the model
- Takes into account time delays in producing proteins (RBS transcription + protein translation)

Model Derivation: β-Gal Production

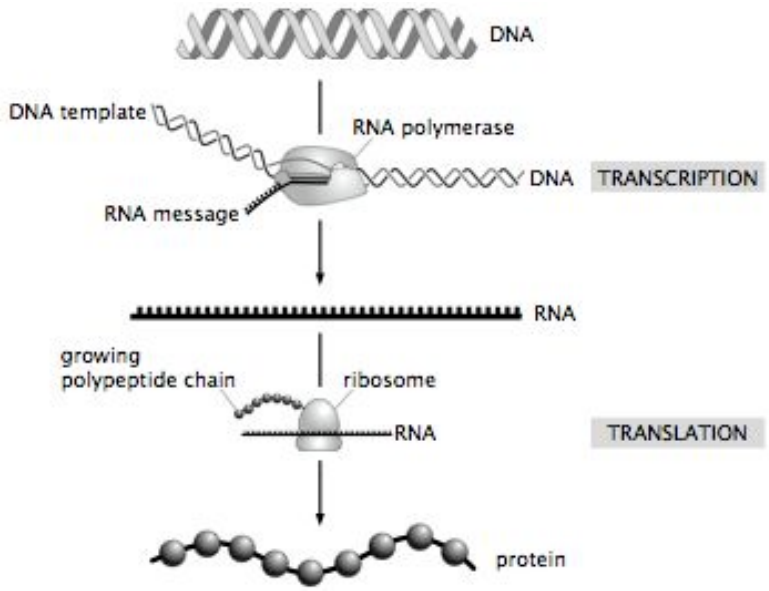
$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A(t-\tau_m))^n}{K + K_1(e^{-\mu\tau_M} A(t-\tau_m))^n} - \bar{\gamma}_M M$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M(t-\tau_B) - \tilde{\gamma}_B B$$

M = lacZ mRNA concentration

B = β-gal concentration

A = allolactose concentration



mRNA production

- Production rate related p_{bound} via a modified Hill function, ala

$$\frac{dM}{dt} = \alpha\big([1 - p_{\text{bound}}(A)] + \epsilon\big) - \gamma M$$

- RNA degradation via exponential decay
- Account for time delay in translation of RBS via τ_M :
  - Use allolactose concentration, *A*, at time t - τ_M
  - Exponential factor to account for dilution due to cell division

Protein production

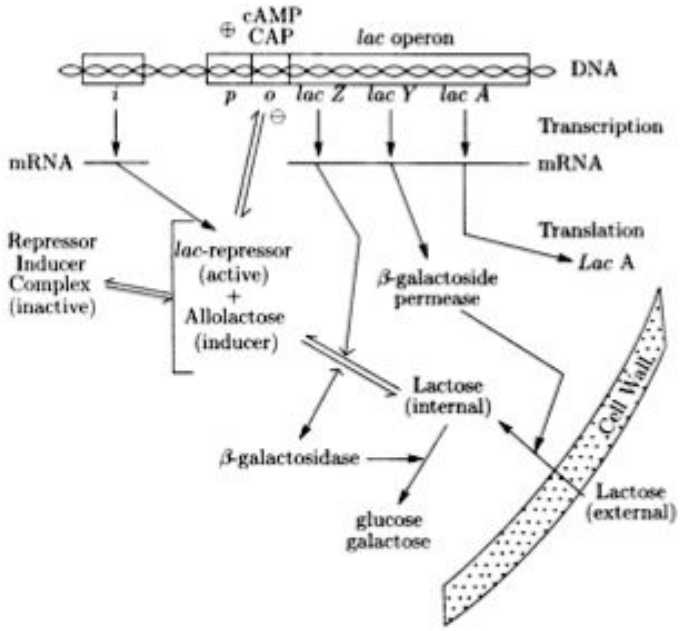
- Assume rate of production is proportional to amount of mRNA
- Include protein degradation via exponential decay
- Add time delay to account for time to produce functional protein, τ_B

Model Derivation: Allolactase Dynamics

$$\frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L}^{①} - \beta_A B \frac{A}{K_A + A}^{②} - \tilde{\gamma}_A A^{③}$$

$$\frac{dL}{dt} = \alpha_L P \frac{L_e}{K_{L_e} + L_e}^{①} - \beta_{L_1} P \frac{L}{K_{L_1} + L}^{②} - \beta_{L_2} B \frac{L}{K_{L_2} + L}^{③} - \tilde{\gamma}_L L^{④}$$

$$\frac{dP}{dt} = \alpha_P e^{-\mu \tau_P} M(t - \tau_P)^{①} - \tilde{\gamma}_P P^{②}$$



A = allolactose concentration

L = internal lactose concentration

P = permease concentration

B = β-gal concentration

Allolactose

1. Converted from lactose with Michaelis Menten-like kinetics (Huber et al)
2. Converted back to glucose and galactose
3. Degradation

Lactose (internal)

1. Transported to interior of cell by permease
2. Loss back to external environment
3. Converted to allolactose by β-gal
4. Degradation

Permease

1. Produced by *lacY* gene (after delay)
2. Degradation

Determining the Constants

Parameter	Value
$\mu_{\max}$	3.47×10^{-2} min^{-1}
$\bar{\mu}$	3.03×10^{-2} min^{-1}
α_M	997 nM-min^{-1}
α_B	1.66×10^{-2} min^{-1}
α_A	1.76×10^{4} min^{-1}
γ_M	0.411 min^{-1}
γ_B	8.33×10^{-4} min^{-1}
γ_A	1.35×10^{-2} min^{-1}
n	2
K	7200
K_1	$2.52\times10^{-2}(\mu\,\mathrm{M})^{-2}$
K_L	0.97 mM
K_A	1.95 mM
β_A	2.15×10^{4} min^{-1}
τ_M	0.10 min
τ_B	2.00 min

Yildirim, Santillan, Horike and Mackey:

- μ - dilution rate, based on 20 minute cell division time
- α_x - production rate, based on steady state values
- γ_x - decay rate, based on half life experiments
- τ_M - time delay to produce RBS, based on RNA elongation rates
- τ_B - time delay to translate protein, based on protein length and translation speed
- n - Hill coefficient (no justification!)
- K - based on basal rate of production (Yagil & Yagil)
- K_1 - based on dissociation constant (Yagil & Yagil)
- K_x - measured by Wong, Gladney and Keasling (97)
- β_A - loss of allolactase, through conversion to glucose and galactose. Measured by Hubert et al (75)

Note: repressor binding model is pretty ad hoc...

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A(t-\tau_m))^n}{K + K_1(e^{-\mu\tau_M} A(t-\tau_m))^n} - \bar{\gamma}_M M \qquad \frac{dA}{dt} = \alpha_A B \frac{L}{K_L + L} - \beta_A B \frac{a}{K_A + A} - \tilde{\gamma}_A A$$

Equilibrium Analysis

General approach

- Nonlinear ODE with parameters p

$$\dot{x} = f(x, p) \qquad x = (M, B, A)$$

- Equilibrium point: $f(x^*, p) = 0$. Value depends on parameters $x^*(p)$
- Equilibrium point is (asymptotically) stable if solutions starting near x^* converge to x^*
- Check stability through linearization:

$$z = x - x^*$$

$$\dot{z} = \left[\left. \frac{\partial f}{\partial x} \right|_{(x^*, p)} \right] z + \text{h. o. t}$$

- Equilibrium point is stable if eigenvalues of the linearization have negative real part (⇒ solutions decay to zero at rate Re(λ))
- Can gain insight into dynamics by plotting value and stability of $x^*(p)$ [bifurcation diagram] and stable regions of p [stability plot]

Lac operon case

- Assume internal lactose concentration constant ⇒ can eliminate L, P dynamics
- Non-dimensionalize equations (rescale):

$$\frac{dm}{d\hat{t}} = \hat{\alpha}_M f(a_{\hat{\tau}_M}) - \hat{\gamma}_M m,$$

$$\frac{db}{d\hat{t}} = \hat{\alpha}_B e^{-\hat{\mu}\hat{\tau}_B} m_{\hat{\tau}_B} - \hat{\gamma}_B b,$$

$$\frac{da}{d\hat{t}} = \hat{\alpha}_A h(l) b - \hat{\beta}_A g(a) b - \hat{\gamma}_A a.$$

$$\tau = \tau_M + \tau_B, \quad a = \sqrt[n]{K_1} A,$$

$$\hat{t} = \frac{t}{\tau}, \quad b = \sqrt[n]{K_1} B,$$

$$\hat{\tau}_M = \frac{\tau_M}{\tau}, \quad m = \sqrt[n]{K_1} M,$$

$$\hat{\tau}_B = \frac{\tau_B}{\tau}, \quad l = \sqrt[n]{K_1} L,$$

$$\hat{\mu} = \mu\tau,$$

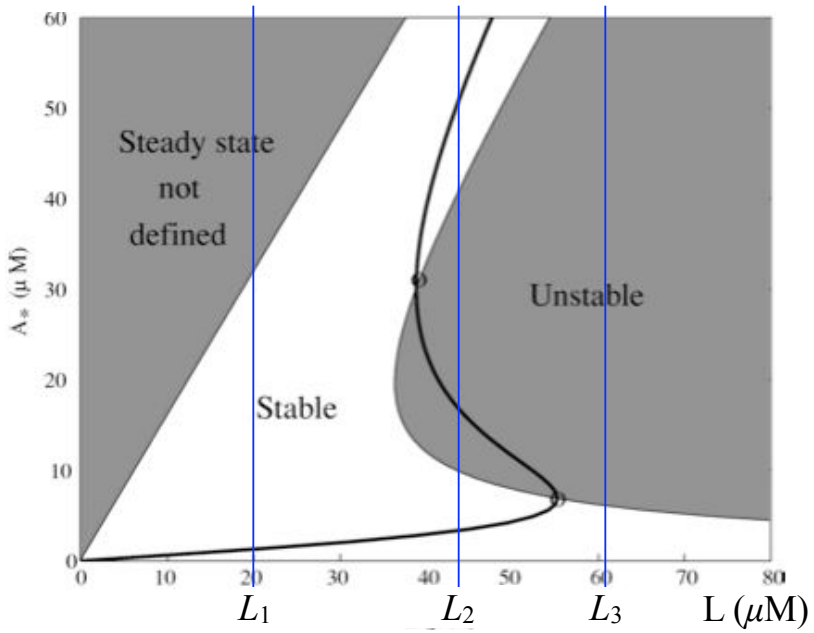
- Compute equilibrium points (solve nonlinear equations):

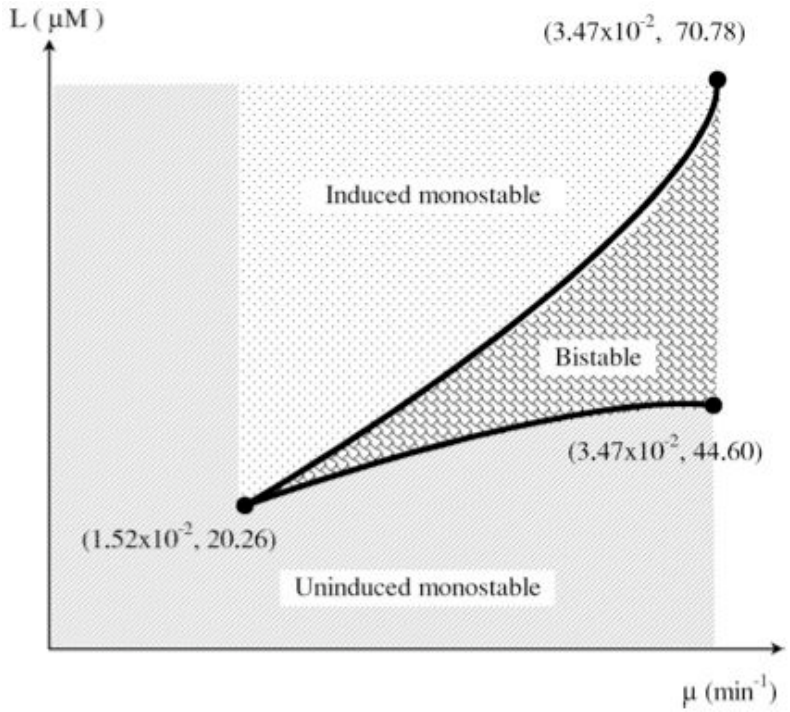
$$m_* = \frac{\hat{\alpha}_M}{\hat{\gamma}_M} f(a_*) \qquad f(a_*) = \Theta \frac{a_*}{h(l) - \frac{\hat{\beta}_A}{\hat{\alpha}_A} g(a_*)}$$

$$b_* = \frac{\hat{\gamma}_A a_*}{\hat{\alpha}_A h(l) - \hat{\beta}_A g(a_*)}$$

- Linearize & compute evalues (w/ delay...)

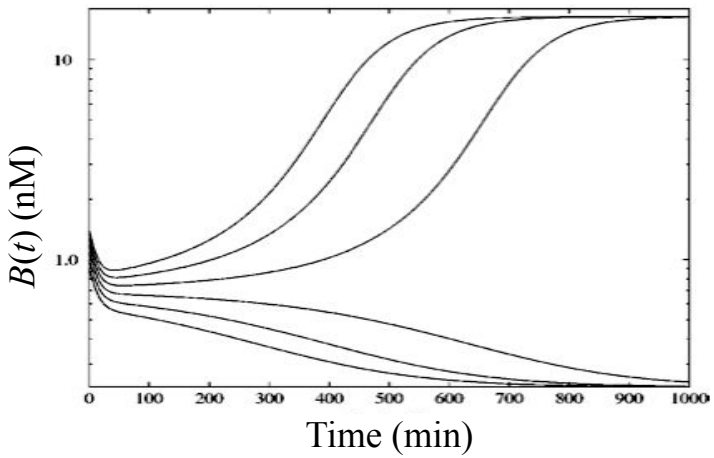
Some Predictions



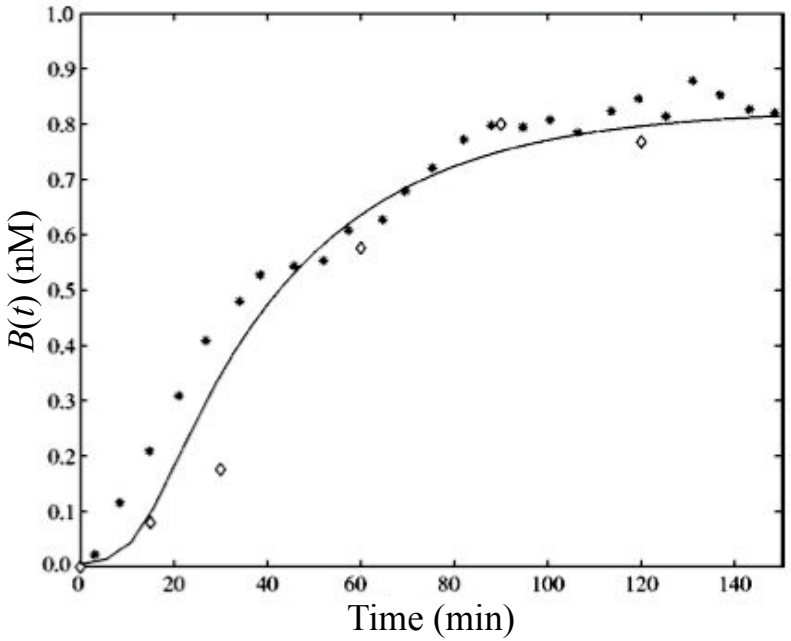


Bistable behavior (saddle node bifurcations)

- Can have single or multiple equilibrium points depending on parameters
- Bifurcation plot: change in stability versus params
  - Note: ossible hysteresis from saddle node
- Parametric stability plot: stability regions
- Simulations: nearby initial conditions can lead to different steady state solutions
- Use to predict behavior (for future experiments)

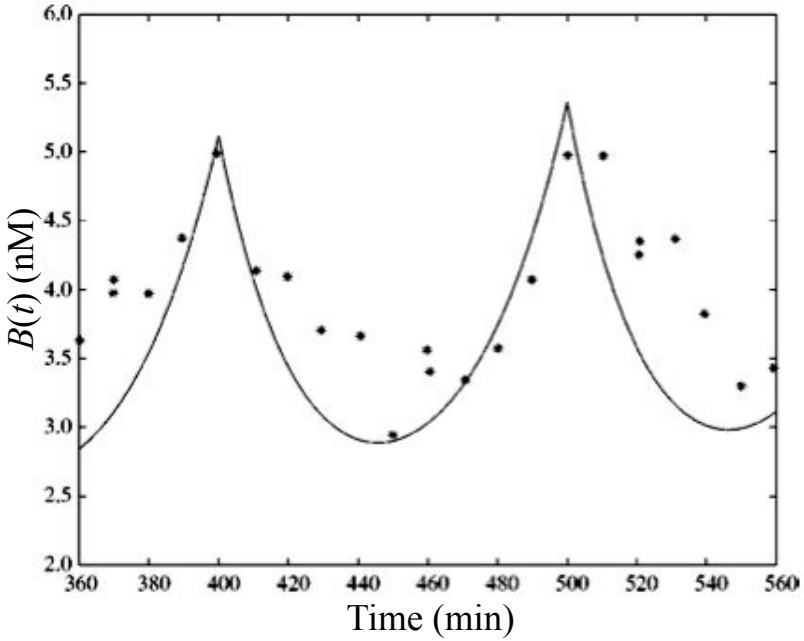


Comparison to Experiment



β-gal activity for L_e = 8 x 10^{-2} mM

- Experimental data from Knorre (1968) for *E. coli* ML30 (◊) and Pestka et al. (1984) for *E. coli* 294 (●)
- Model simulation using constants from Table 1 (slide 16) with μ = 2.26 x 10^{-2} min^{-1} and γ_x (= ??) fit to data



Oscillation in β-gal w/ phosphate feeding

- Periodic phosphate feeding from Goodwin (1969)
- Simulation used μ = 2.26 x 10^{-2} min^{-1} and γ_x (= ??). Other parameters unchanged.
- Q: how should we assess these data?

Questioning the Model

Do time delays matter?

- *lacZ* transcription:
  - Half life (γ_M) ≈ 1.73 min
  - τ_M = 0.1 min; $\mu\tau_M$ = 3 x 10^{-3}
  - exp(-μ τ_M) = 0.997
- β-gal production:
  - Half life (γ_B) ≈ 900 min
  - τ_B = 2 min (cf *lacZ* half life)
  - exp(-μ τ_B) = 0.942

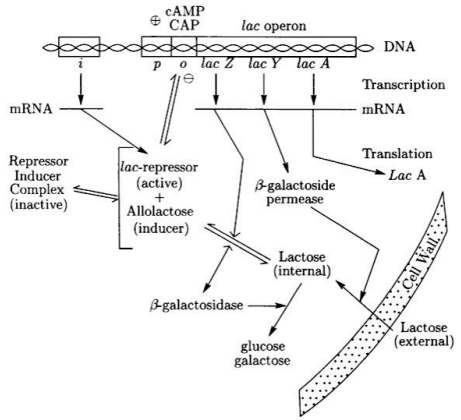
$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1(e^{-\mu\tau_M} A(t-\tau_m))^n}{K + K_1(e^{-\mu\tau_M} A(t-\tau_m))^n} - \bar{\gamma}_M M$$

$$\frac{dB}{dt} = \alpha_B e^{\mu\tau_B} M(t-\tau_B) - \tilde{\gamma}_B B$$

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

Do we learn anything new from the model?

Can we use the model for prediction, design, ???



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

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

and

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