Modeling and Simulation of Gene Regulatory Networks

Hidde de Jong

INRIA Grenoble - Rhône-Alpes

Hidde.de-Jong@inria.fr http://ibis.inrialpes.fr

UNIVERSITE JOSEPH FOURIER SCIENCES TECHNOLOGIE MEDECINE

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE | INRIA | centre de recherche GRENOBLE - RHÔNE-ALPES

INRIA Grenoble - Rhône-Alpes and IBIS

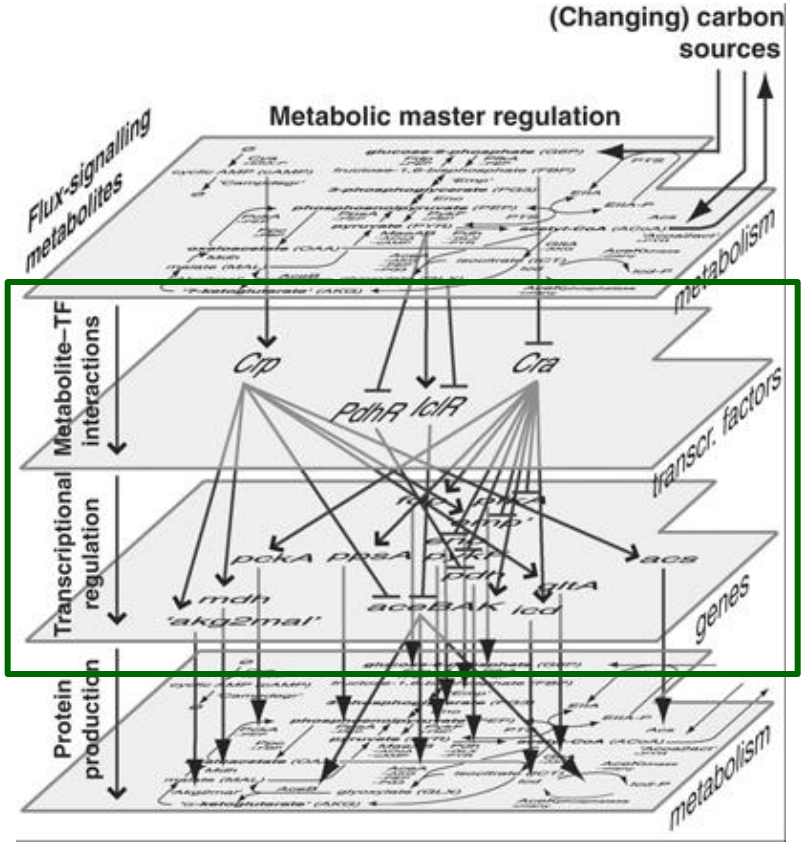
- IBIS: systems biology group of INRIA and Université Joseph Fourier/CNRS
  - Analysis of bacterial regulatory networks by means of models and experiments
  - Biologists, computer scientists, mathematicians, physicists, …

Overview

1. Gene regulatory networks in bacteria
2. Novel methods for measuring gene expression
3. **Quantitative modeling of gene regulatory networks**
   - **Ordinary differential equations**
   - **Stochastic master equations**
4. Qualitative modeling of gene regulatory networks
   - Piecewise-linear differential equations
5. Conclusions and perspectives

Gene regulatory networks

- **Gene regulatory networks** control changes in gene expression levels in response to environmental perturbations

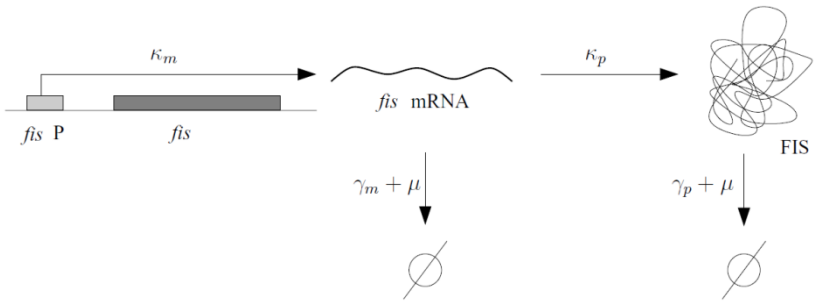


- Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

  Global regulators of transcription involved in glucose-acetate diauxie in *E. coli*

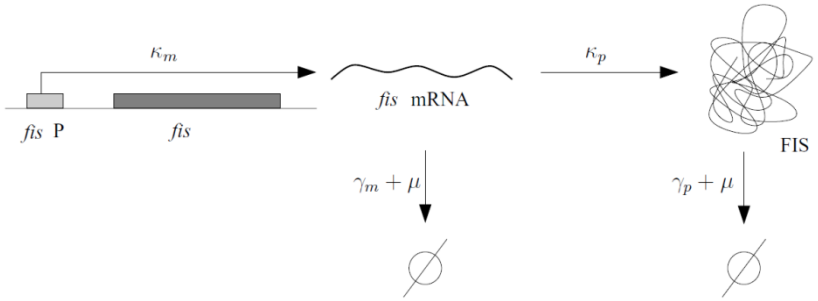
Kotte *et al*. (2010), *Mol. Syst. Biol*., 6: 355

Gene expression



- Typically, and simplifying quite a bit, **gene expression** in bacteria involves:
  - Transcription by RNAP (mRNA)
  - Translation by ribosomes (proteins)
  - Degradation of mRNA and protein

Regulation of gene expression



- Typically, and simplifying quite a bit, **regulation of gene expression** in bacteria involves:
  - Transcription regulation by transcription factors
  - Translational regulation by small RNAs
  - Regulation of degradation by proteases

Modeling of gene regulatory networks

- Different modeling formalisms exist, describing gene expression on different levels of detail

Stochastic master equations

Ordinary differential equations (ODEs)

Boolean networks

Smolen *et al.* (2000), *Bull. Math. Biol.*, 62(2):247-292

Hasty *et al.* (2001), *Nat. Rev. Genet.*, 2(4):268-279

de Jong (2002), *J. Comput. Biol.*, 9(1): 69-105

Szallassi et al. (2006), *System Modeling in Cellular Biology*, MIT Press

Bolouri (2008), *Computational Modeling of Gene Regulatory Networks*, Imperial College Press

Karleback and Shamir (2008), *Nat. Rev. Mol. Cell Biol.*, 9(10):770-80

Ordinary differential equation models

- Cellular concentration of proteins, mRNAs, and other molecules at time-point t represented by continuous variable $x_i(t) \in \mathrm{R}_{\geq 0}$
- Regulatory interactions, controlling synthesis and degradation, modeled by **ordinary differential equations**

$$\frac{d\boldsymbol{x}}{dt} = \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}),$$

where $\boldsymbol{x} = [x_1, \ldots, x_n]'$ and $\boldsymbol{f}(\boldsymbol{x})$ is **rate law**

- Kinetic theory of biochemical reactions provides well-established framework for specification of rate laws

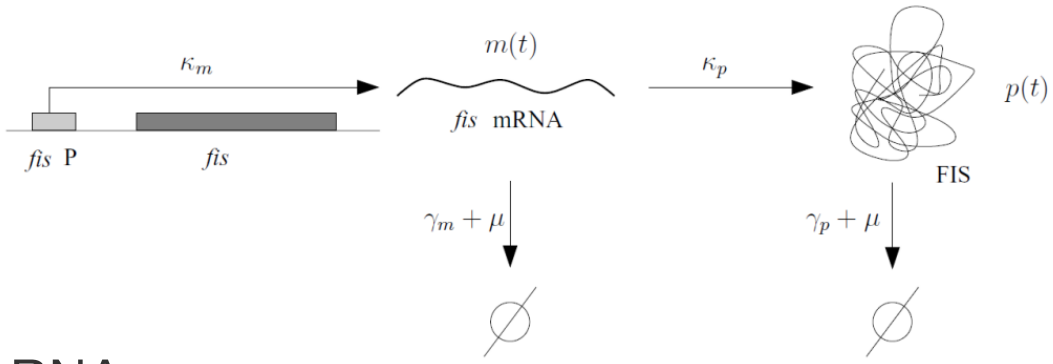
Heinrich and Schuster (1996), *The Regulation of Cellular Systems*
Cornish-Bowden (1995), *Fundamentals of Enzyme Kinetics*

Modeling of gene regulatory networks

- ODE model of gene expression, distinguishing **transcription** and **translation**

$$\dot{m} = \kappa_m - (\gamma_m + \mu)\, m$$

$$\dot{p} = \kappa_p\, m - (\gamma_p + \mu)\, p$$



$m(t) \geq 0$, concentration mRNA

$p(t) \geq 0$, concentration protein

$\kappa_m, \kappa_p > 0$, synthesis rate constants

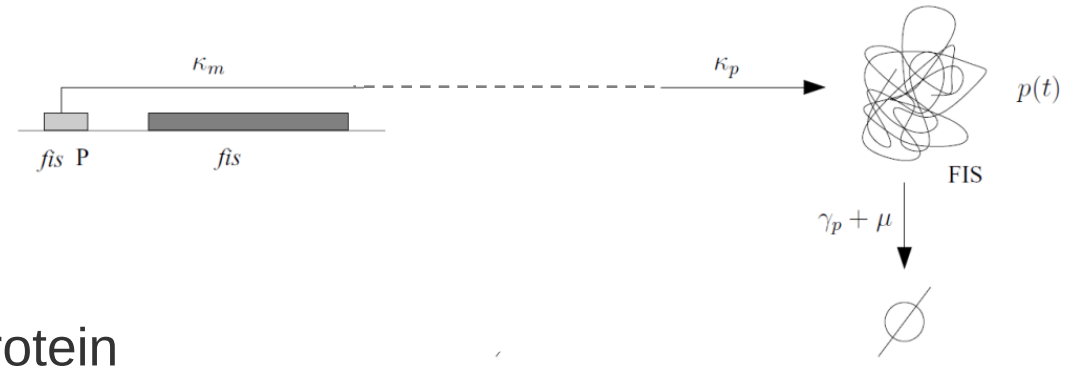
$\gamma_m, \gamma_p > 0$, degradation rate constants

$\mu(t) \geq 0$, growth rate

Modeling of gene regulatory networks

- ODE model of gene expression, collapsing **transcription** and **translation**

$$\dot{p} = \kappa_p - (\gamma_p + \mu)\, p$$



$p(t) \geq 0$, concentration protein

$\kappa_p > 0$, synthesis rate constants

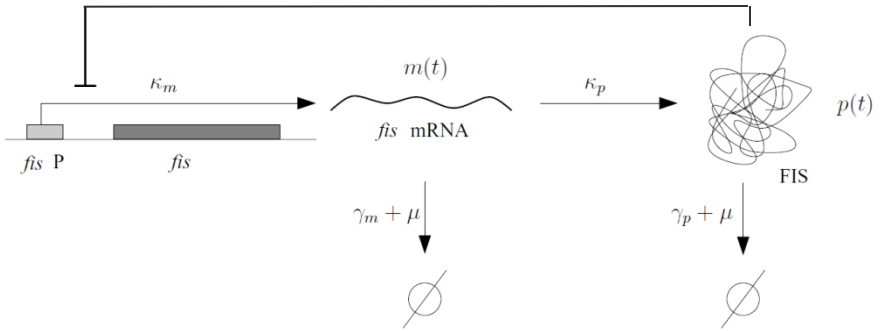
$\gamma_p > 0$, degradation rate constants

$\mu(t) \geq 0$, growth rate

Modeling of gene regulatory networks

- ODE model of gene expression, taking into account **regulation** of transcription

$$\dot{m} = \kappa_m \; f(p) - (\gamma_m + \mu)\, m$$
$$\dot{p} = \kappa_p\, m - (\gamma_p + \mu)\, p$$



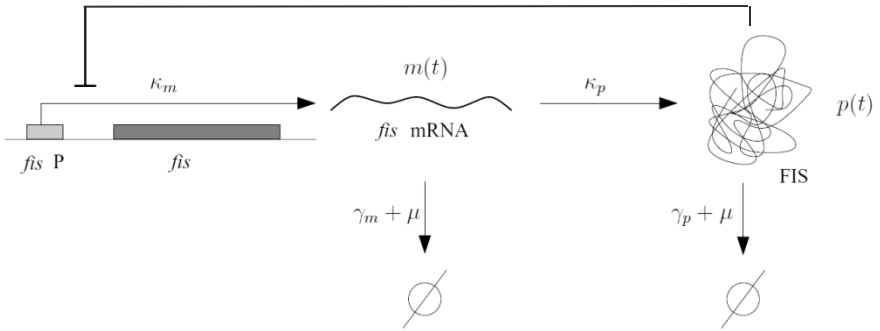
- Regulation function $f(p)$ describes modulation of synthesis rate by transcription factor

  Generalization to regulation on translational and proteolytic level

Modeling of gene regulatory networks

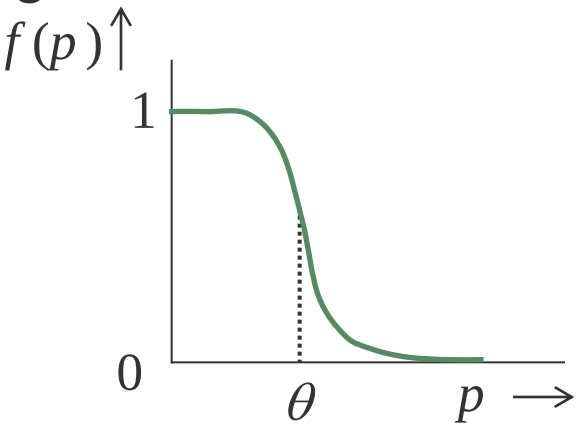
- ODE model of gene expression, taking into account **regulation** of transcription

$$\dot{m} = \kappa_m \; f(p) - (\gamma_m + \mu)\, m$$
$$\dot{p} = \kappa_p\, m - (\gamma_p + \mu)\, p$$



- Regulation function $f(p)$ typically has **sigmoidal** form, accounting for cooperative nature of regulation

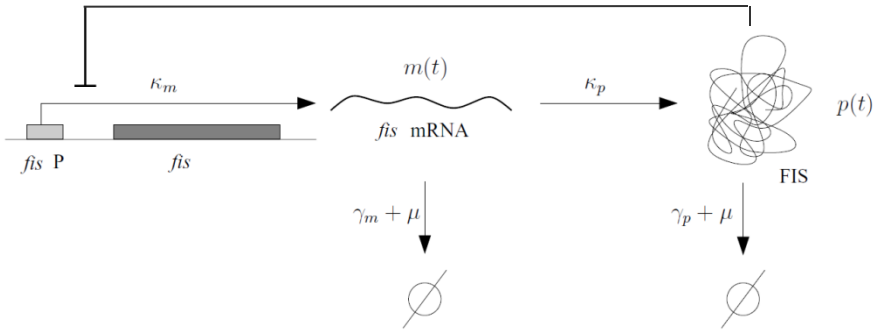
$$f(p) = \frac{\theta^n}{\theta^n + p^n}, \quad \theta > 0 \text{ threshold}, \quad n > 1 \text{ cooperativity}$$



Modeling of gene regulatory networks

- ODE model of gene expression, taking into account **regulation** of transcription

$$\dot{m} = \kappa_m \; f(p) - (\gamma_m + \mu)\, m$$
$$\dot{p} = \kappa_p\, m - (\gamma_p + \mu)\, p$$

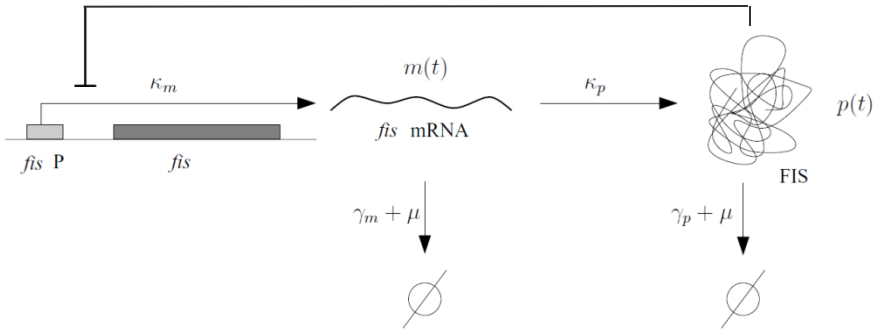


- Regulation function $f(p)$ typically has **sigmoidal** form, accounting for cooperative nature of regulation
- Implicit modeling assumptions:
  - Ignore gene expression machinery (RNA polymerase, ribosome)
  - Simplification of complex interactions of regulators with DNA to single response function

Modeling of gene regulatory networks

- ODE model of gene expression, taking into account **regulation** of transcription

$$\dot{m} = \kappa_m \; f(p) - (\gamma_m + \mu)\, m$$
$$\dot{p} = \kappa_p\, m - (\gamma_p + \mu)\, p$$



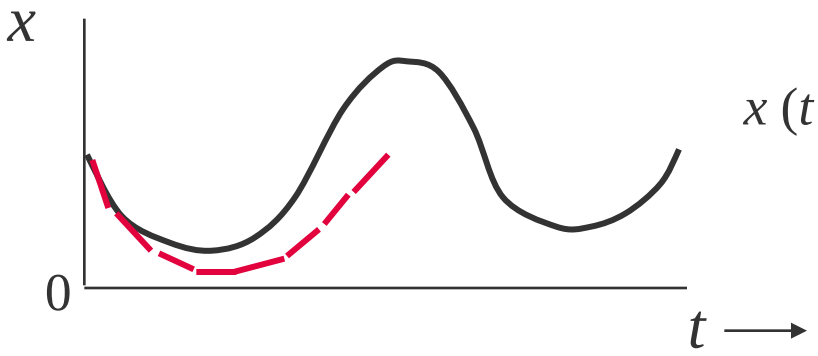
- Gene regulatory network has many genes with mutual regulatory interactions: model of coupled ODEs

Analysis and numerical simulation

- No analytical solution for most nonlinear differential equations
- **Dynamic systems theory** provides techniques for analysis of nonlinear differential equations, but usually not scalable
  - Phase portrait
  - Bifurcation analysis

Kaplan and Glass (1995), *Understanding Nonlinear Dynamics*

- Approximation of solution obtained by **numerical simulation**, given parameter values and initial conditions $\boldsymbol{x}(0) = \boldsymbol{x}^0$

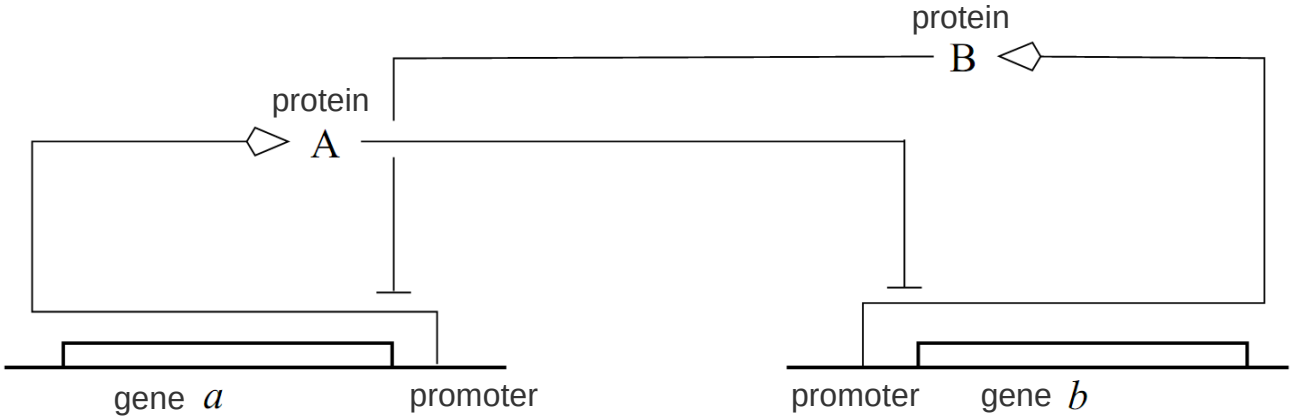


$$x(t + \Delta t) = x(t) + \int_t^{t+\Delta t} f(x)\, dt \approx x(t) + f(x)\, \Delta t$$

Lambert (1991), *Numerical Methods for Ordinary Differential Equations*

Cross-inhibition network

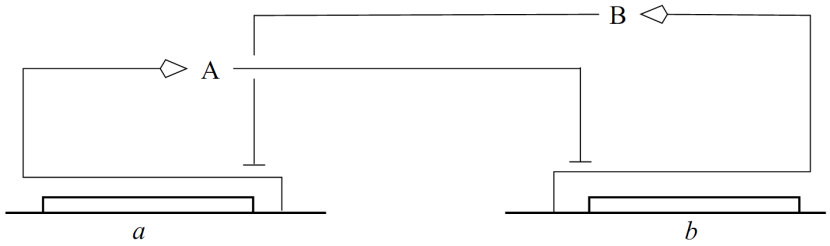
- **Cross-inhibition** network consists of two genes, each coding for transcription regulator inhibiting expression of other gene



- Cross-inhibition network is example of **positive feedback**, important for phenotypic differentiation (multi-stability)

Thomas and d'Ari (1990), *Biological Feedback*

ODE model of cross-inhibition network



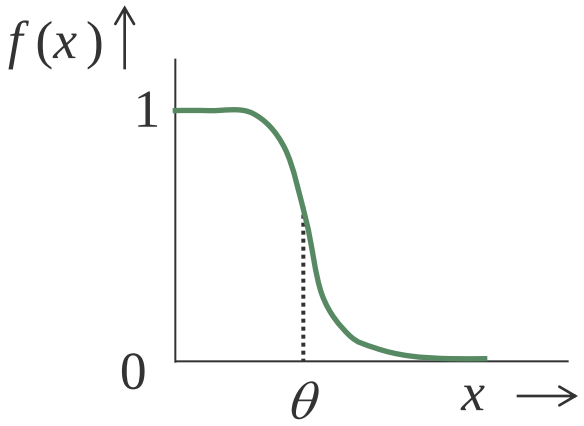
$$\dot{x}_a = \kappa_a f(x_b) - \gamma_a x_a$$
$$\dot{x}_b = \kappa_b f(x_a) - \gamma_b x_b$$

$x_a \geq 0$, concentration protein A

$x_b \geq 0$, concentration protein B

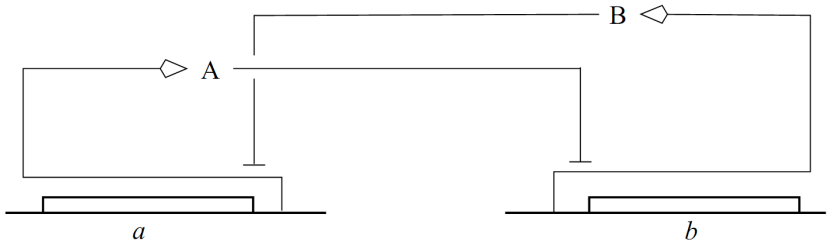
$\kappa_a, \kappa_b > 0$, production rate constants

$\gamma_a, \gamma_b > 0$, degradation rate constants



$$f(x) = \frac{\theta^n}{\theta^n + x^n}, \quad \theta > 0 \text{ threshold}, \quad n > 1 \text{ cooperativity}$$

ODE model of cross-inhibition network



$$\dot{x}_a = \kappa_a f(x_b) - \gamma_a x_a$$
$$\dot{x}_b = \kappa_b f(x_a) - \gamma_b x_b$$

$x_a \geq 0$, concentration protein A

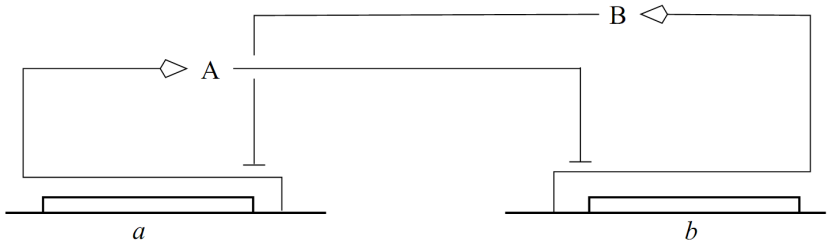
$x_b \geq 0$, concentration protein B

$\kappa_a, \kappa_b > 0$, production rate constants

$\gamma_a, \gamma_b > 0$, degradation rate constants

- Implicit modeling assumptions:
  - Ignore intermediate gene products (mRNA)
  - Ignore gene expression machinery (RNA polymerase, ribosome)
  - Simplification of complex interactions of regulators with DNA to single response function

ODE model of cross-inhibition network



$$\dot{x}_a = \kappa_a f(x_b) - \gamma_a x_a$$
$$\dot{x}_b = \kappa_b f(x_a) - \gamma_b x_b$$

$x_a \geq 0$, concentration protein A

$x_b \geq 0$, concentration protein B

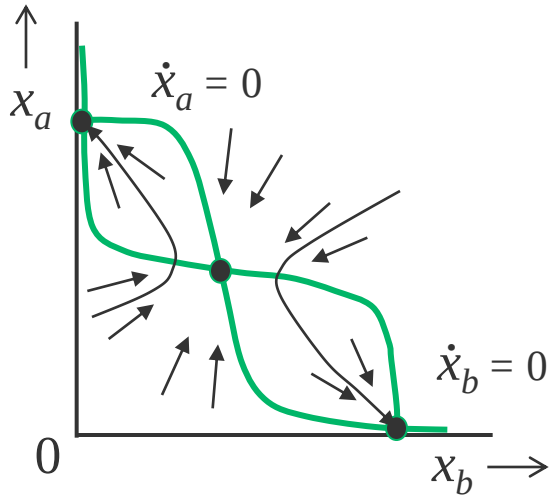
$\kappa_a, \kappa_b > 0$, production rate constants

$\gamma_a, \gamma_b > 0$, degradation rate constants

- Additional implicit modeling assumption:
  - Assume constant growth rate (and collapse with degradation)

Bistability of cross-inhibition network

- Analysis of **steady states** in phase plane



$$\dot{x}_a = 0 : \; x_a = \frac{\kappa_a}{\gamma_a} f(x_b)$$

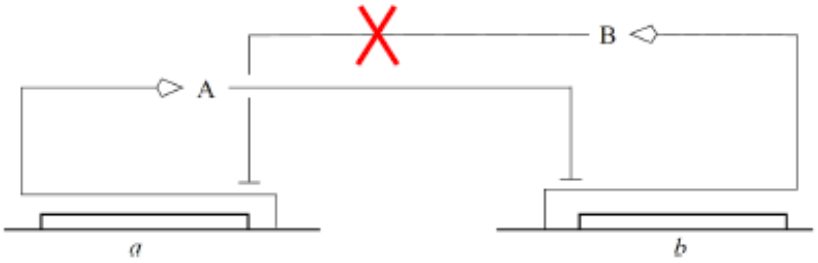
$$\dot{x}_b = 0 : \; x_b = \frac{\kappa_b}{\gamma_b} f(x_a)$$

- System is **bistable**: two stable and one unstable steady state.
- For almost all initial conditions, system will converge to one of two stable steady states (**differentiation**)
- System returns to steady state after small perturbation

Switching in cross-inhibition network

- Transient perturbation may cause irreversible **switch** from one steady state to the other

  Temporary disable one of the inhibitors



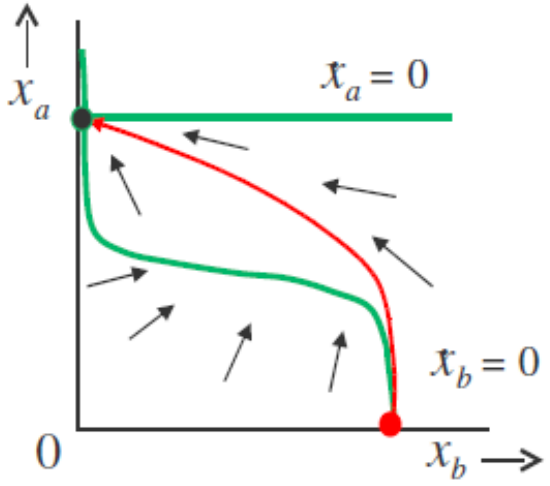
$$x_a = \kappa_a - \gamma_a x_a$$

$$x_b = \kappa_b f(x_a) - \gamma_b x_b$$

Switching in cross-inhibition network

- Transient perturbation may cause irreversible **switch** from one steady state to the other

System evolves to new steady state



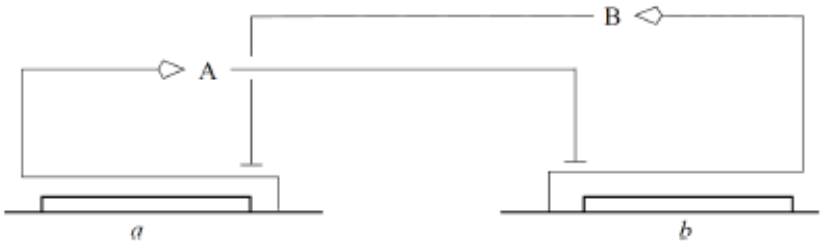
$$x_a = 0 : \; x_a = \frac{\kappa_a}{\gamma_a}$$

$$x_b = 0 : \; x_b = \frac{\kappa_b}{\gamma_b} f(x_a)$$

Switching in cross-inhibition network

- Transient perturbation may cause irreversible **switch** from one steady state to the other

  Enable again inhibitor



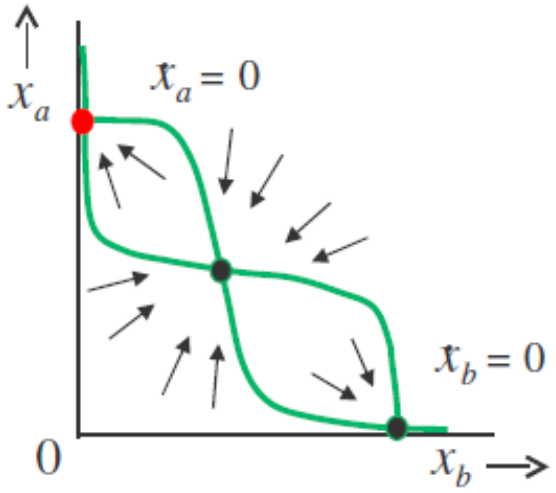
$$x_a = \kappa_a f(x_b) - \gamma_a x_a$$

$$x_b = \kappa_b f(x_a) - \gamma_b x_b$$

Switching in cross-inhibition network

- Transient perturbation may cause irreversible **switch** from one steady state to the other

  System remains in new steady state



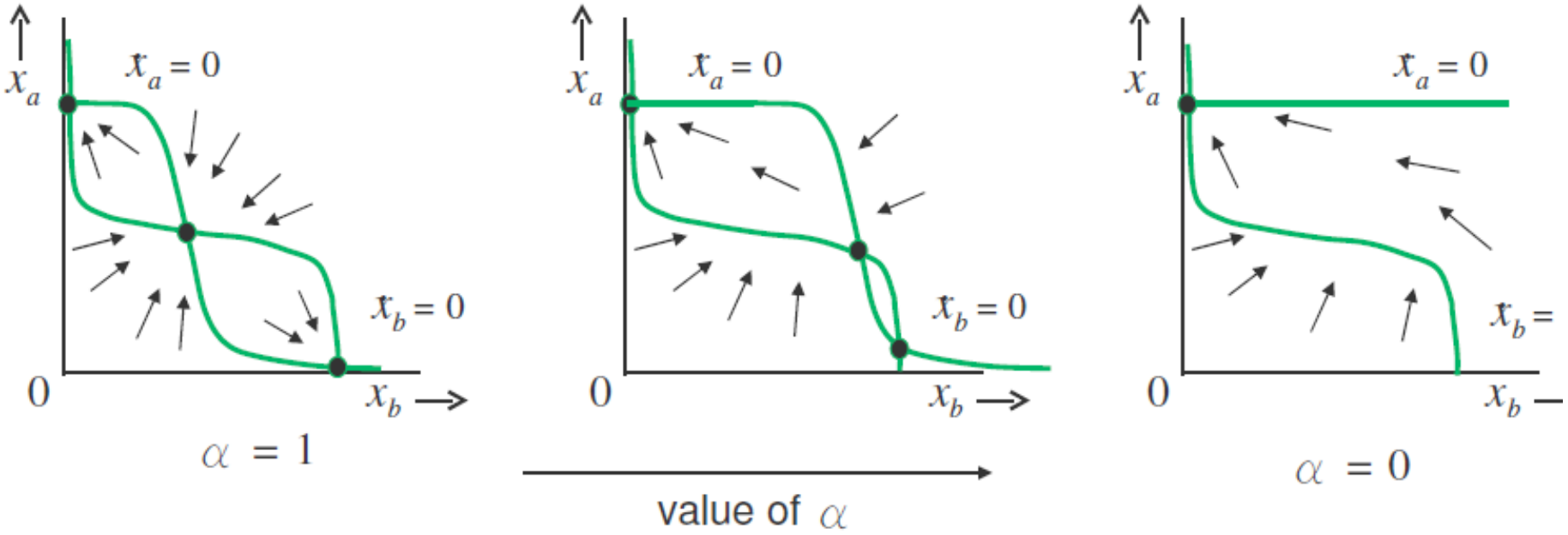
$$x_a = 0: \; x_a = \frac{\kappa_a}{\gamma_a} f(x_b)$$

$$x_b = 0: \; x_b = \frac{\kappa_b}{\gamma_b} f(x_a)$$

Bifurcation in cross-inhibition network

- Switching of cross-inhibition network can be interpreted as sequence of **bifurcations**, induced by change in parameter

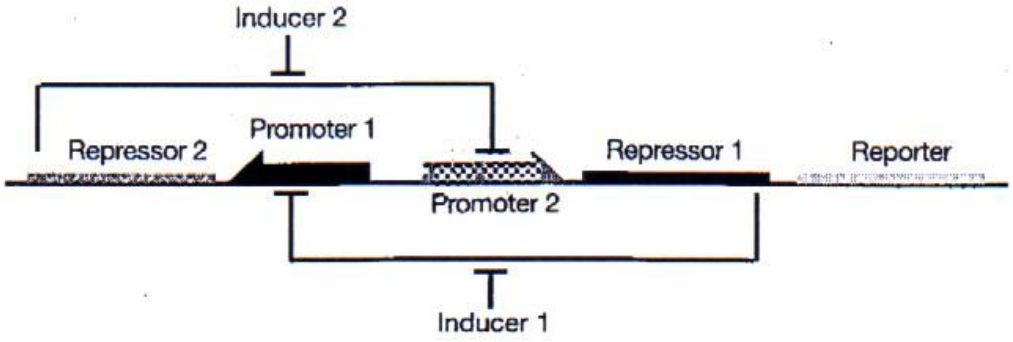
$$x_a = \kappa_a f(\alpha\ x_b) - \gamma_a x_a$$
$$x_b = \kappa_b f(x_a) - \gamma_b x_b$$

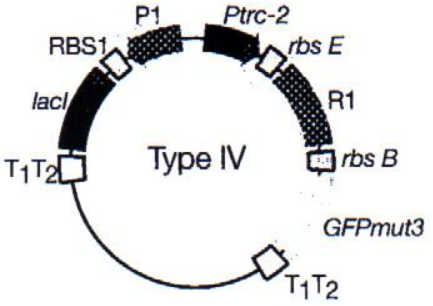


Construction of cross inhibition network

- Construction of cross inhibition network *in vivo*

Gardner *et al.* (2000), *Nature*, 403(6786): 339-342





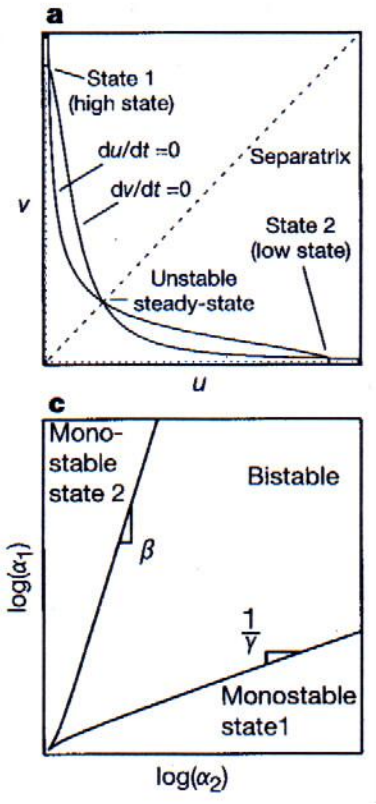
- Differential equation model of network

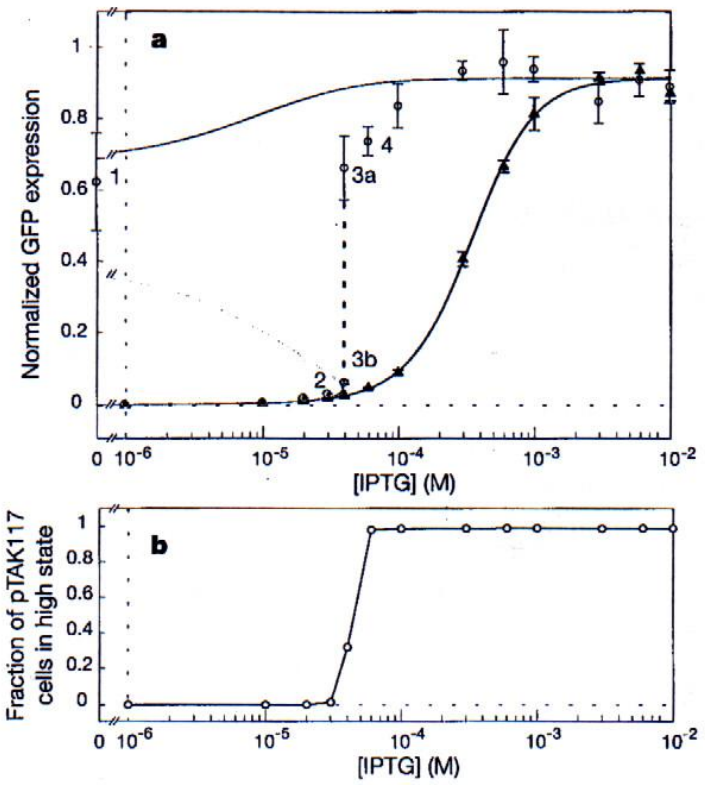
$$\dot{u} = \frac{\alpha_1}{1 + v^{\beta}} - u \qquad \dot{v} = \frac{\alpha_2}{1 + u^{\gamma}} - v$$

Experimental test of model

- Experimental test of mathematical model (bistability and hysteresis)

Gardner *et al.* (2000), *Nature*, 403(6786): 339-342

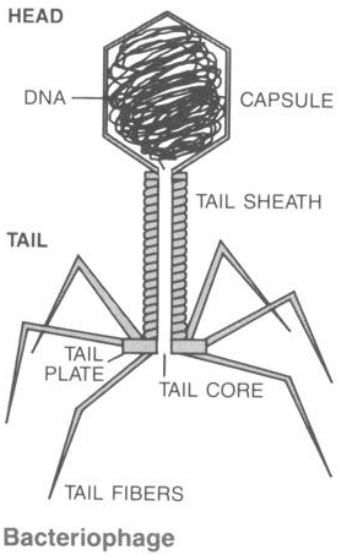


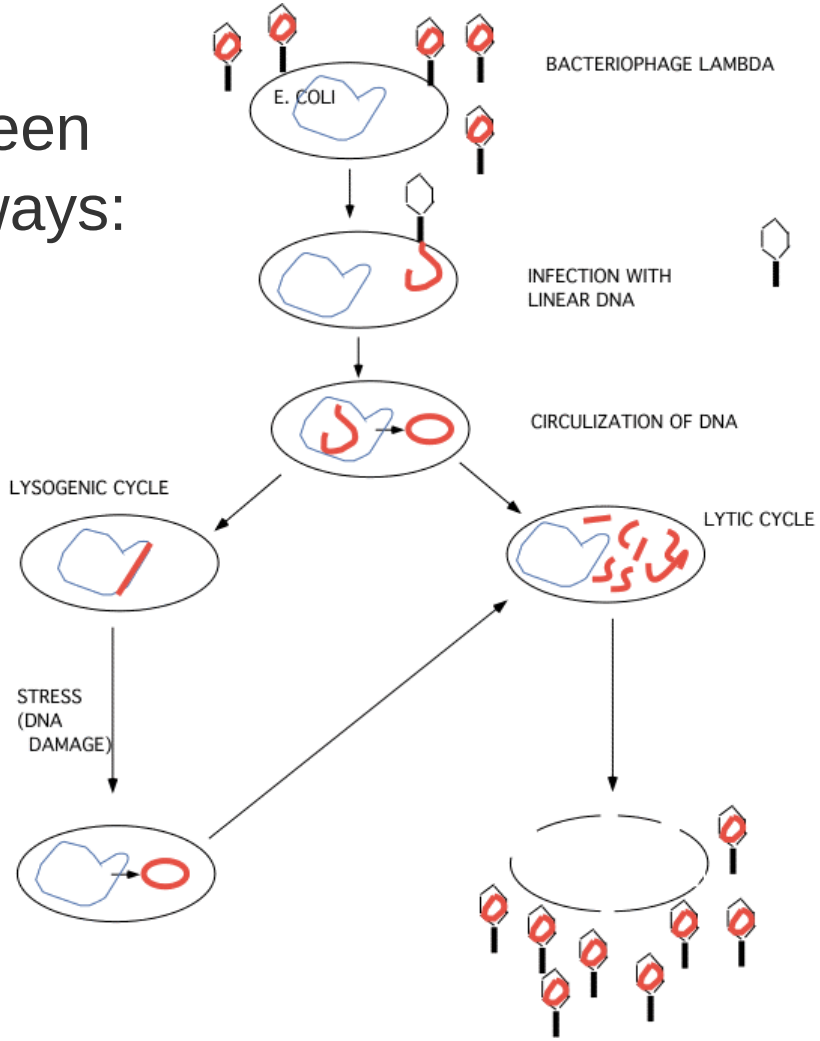


Bacteriophage λ infection of *E. coli*

- Response of *E. coli* to phage λ infection involves decision between alternative developmental pathways: **lysis** and **lysogeny**

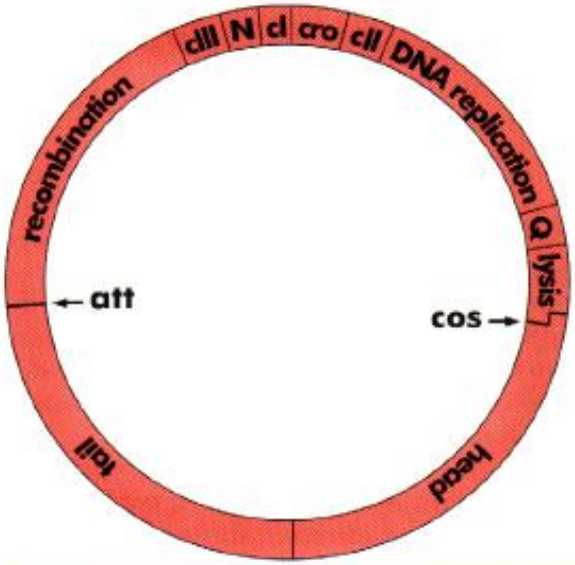
Ptashne, *A Genetic Switch*, Cell Press,1992

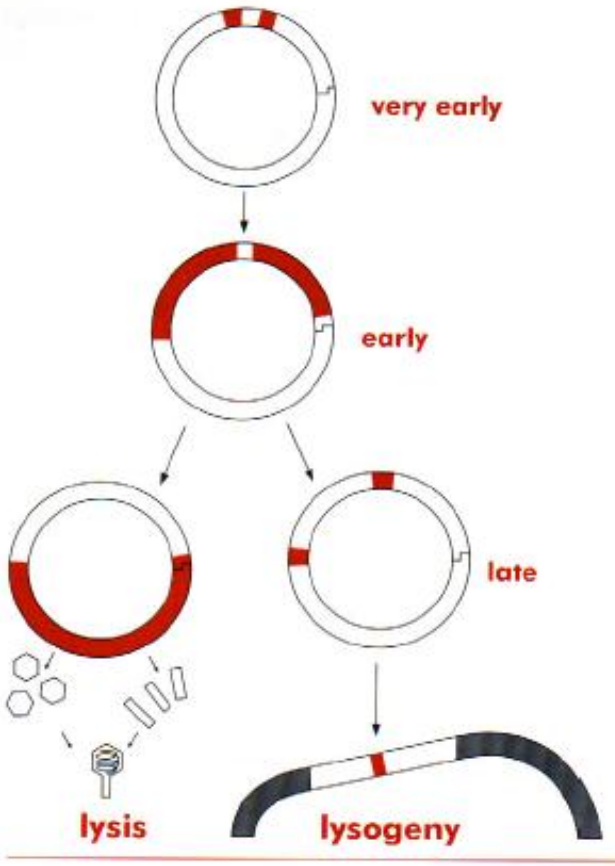




Bistability in phage λ

- Lytic and lysogenic pathways involve different patterns of gene expression

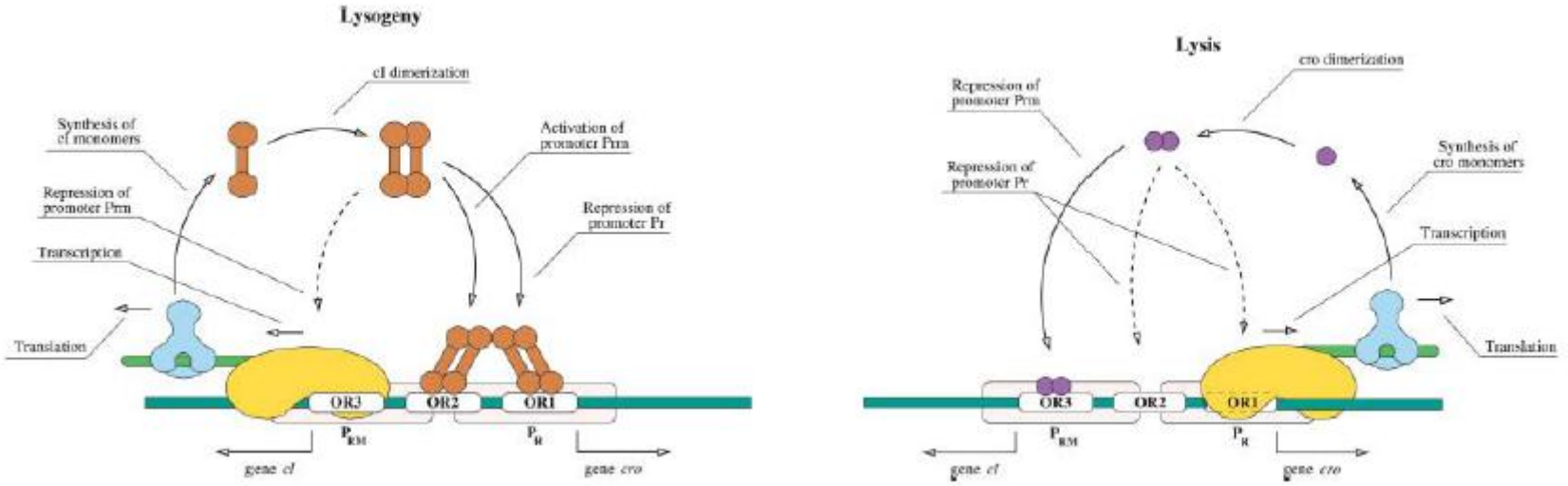




Ptashne, *A Genetic Switch*, Cell Press,1992

Control of phage λ fate decision

- Cross-inhibition feedback plays key role in establishment of lysis or lysogeny, as well as in induction of lysis after DNA damage



Santillán, Mackey (2004), *Biophys. J.*, 86(1): 75-84

Simple model of phage λ fate decision

- Differential equation model of cross-inhibition feedback network involved in phage λ fate decision

  mRNA and protein, delays, thermodynamic description of gene regulation

$$\frac{d[M_{cI}]}{dt} = k^{q}_{cI}[O_R] f^{q}_{RM}([CI_2]_{\tau_M}, [CI_2]_{\tau_M}) + k^{s}_{cI}[O_R] f^{s}_{RM}([CI_2]_{\tau_M}, [Cro_2]_{\tau_M}) - (\gamma_M + \mu)[M_{cI}],$$

$$\frac{d[M_{cro}]}{dt} = k_{cro}[O_R] f_R([CI_2]_{\tau_M}) - (\gamma_M + \mu)[M_{cro}],$$

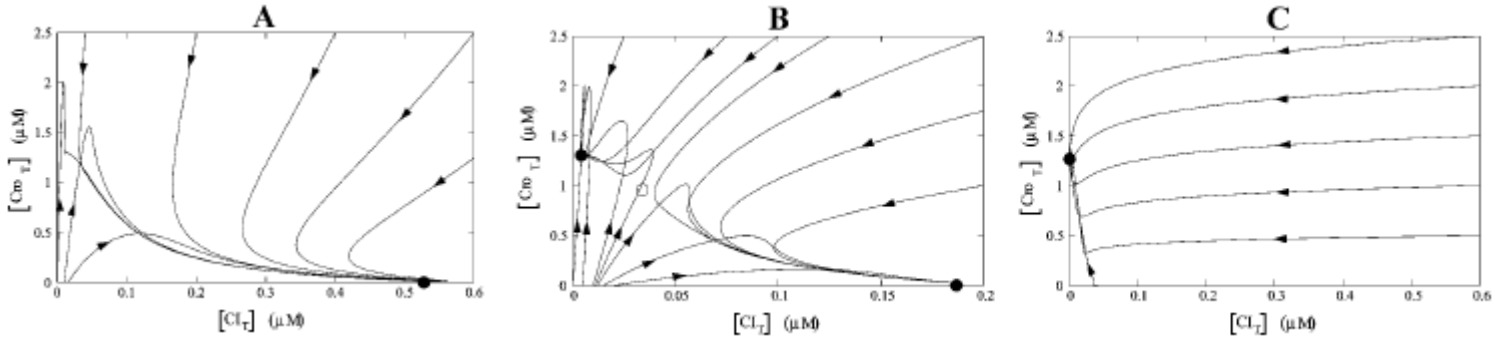
$$\frac{d[CI_T]}{dt} = \upsilon_{cI}[M_{cI}]_{\tau_{cI}} - (\gamma_{cI} + \mu)[CI_T],$$

$$\frac{d[Cro_T]}{dt} = \upsilon_{cro}[M_{cro}]_{\tau_{cro}} - (\gamma_{cro} + \mu)[Cro_T].$$

Santillán, Mackey (2004), *Biophys. J.*, 86(1): 75-84

Analysis of phage λ model

- Bistability (lysis and lysogeny) only occurs for certain parameter values
- Switch from lysogeny to lysis involves bifurcation from one monostable regime to another, due to change in degradation constant



Santillán, Mackey (2004), *Biophys. J.*, 86(1): 75-84

Extended model of phage λ infection

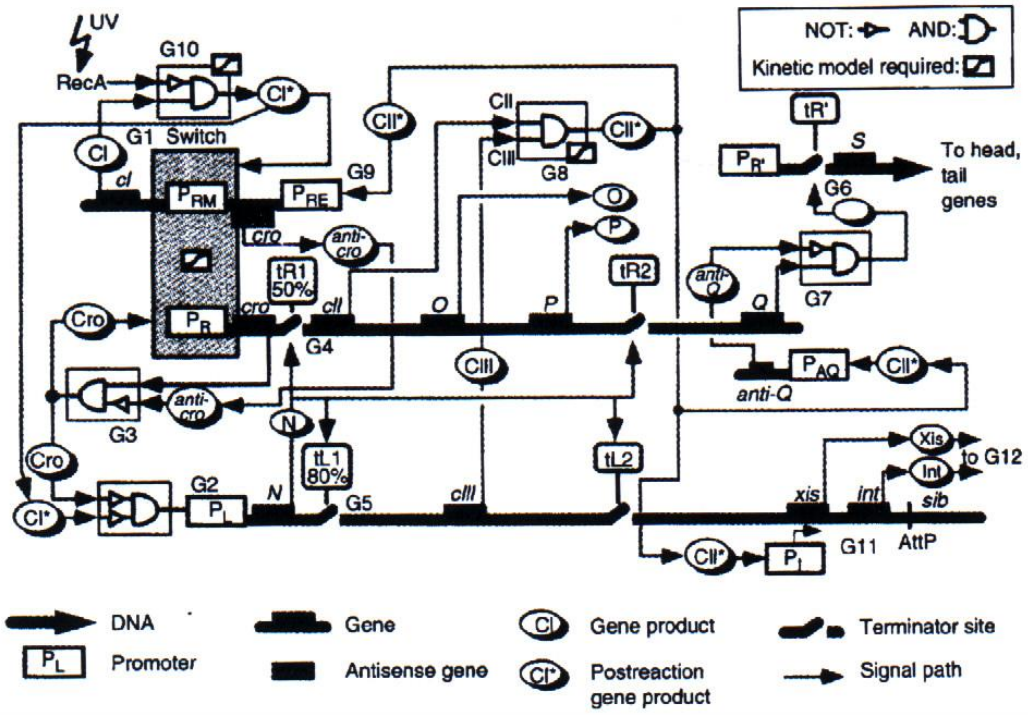
- ODE model of the **extended network** underlying decision between lysis and lysogeny

  Role of other regulatory proteins (CII, N, Q, …)

McAdams and Shapiro (1995), *Science*, 269(5524): 650-656

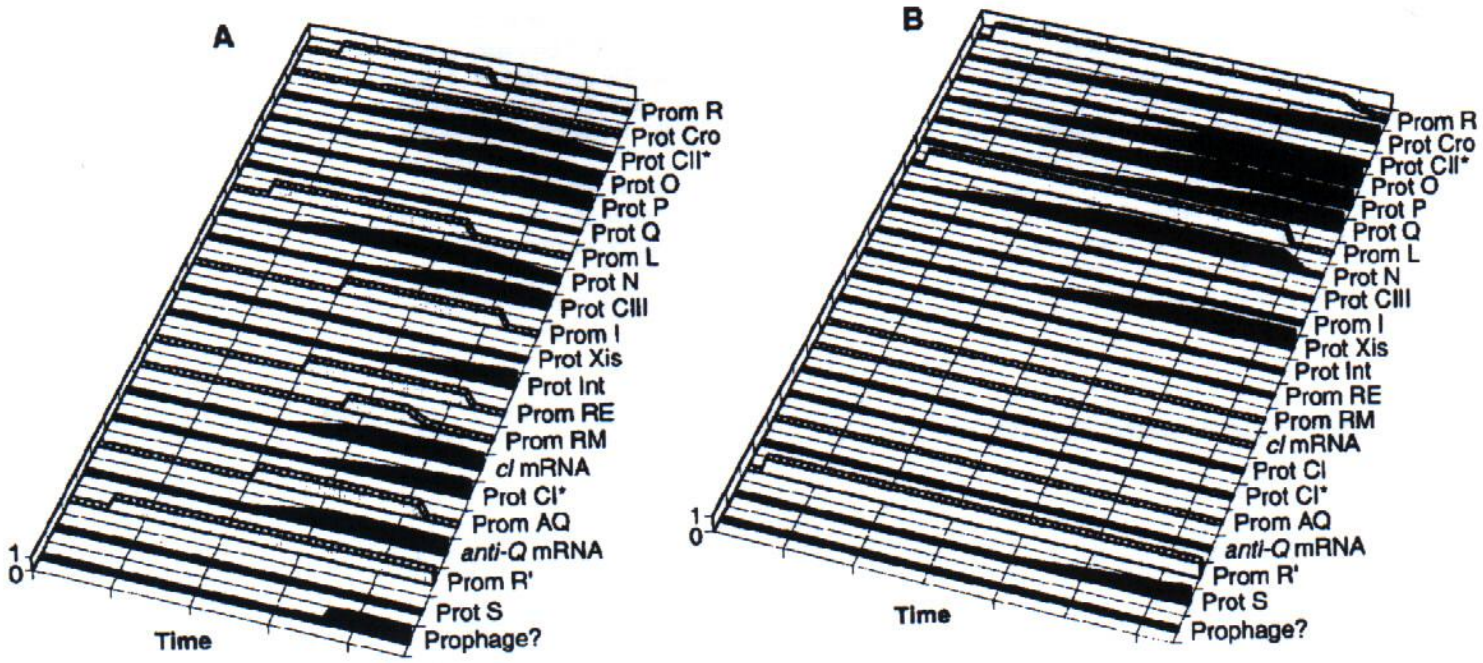
- Recent experimental work downplays importance of mutual inhibition of CI and Cro in lysis-lysogeny decision

Oppenheim *et al.* (2005), *Annu. Rev. Genet.*, 39:409–29



Simulation of phage λ infection

- Numerical simulation of promoter activity and protein concentrations in (a) lysogenic and (b) lytic pathways

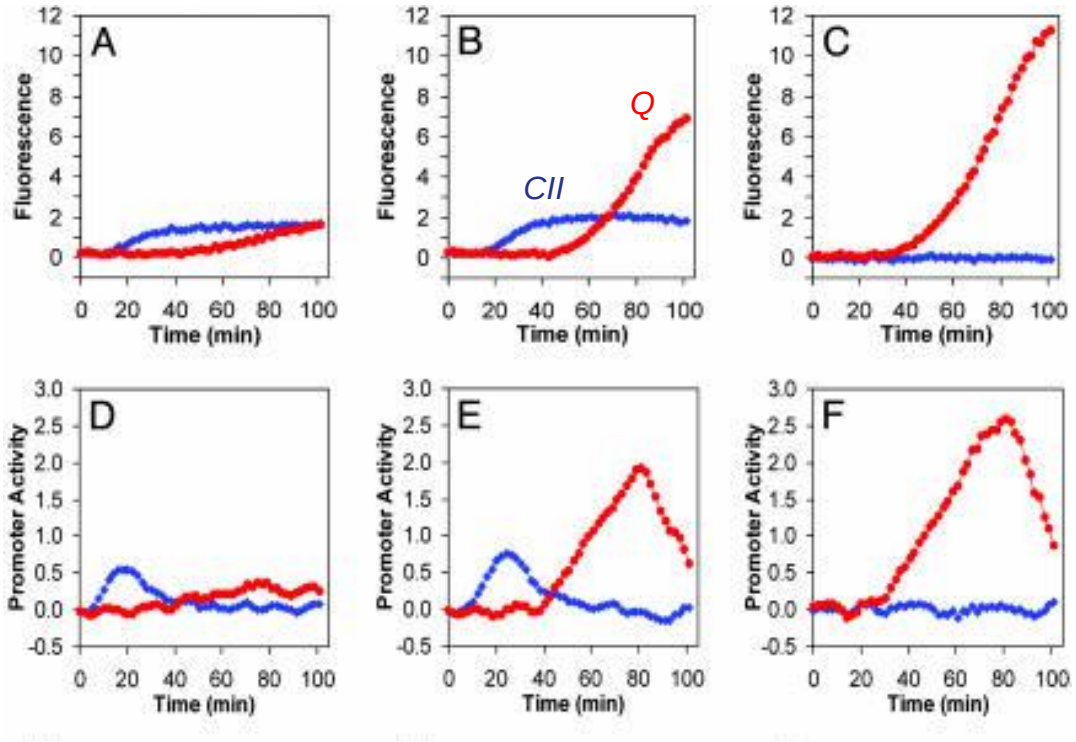


- Cell follows one of two pathways for different initial conditions

Real-time monitoring of phage λ infection

- New measurement techniques allow real-time and *in-vivo* monitoring of the execution of lytic and lysogenic pathways

  Use of fluorescent reporter genes in combination with automated plate readers



Kobiler *et al.* (2005), *Proc. Natl. Acad. Sci. USA*, 102(12): 4470-5

Other examples of bistability

- Many other examples of bistability exist in bacteria
  - Lactose utilization in *E. coli*
  - Persister cells and antibiotic resistance in *E. coli*
  - Genetic competence in *B. subtilis*
  - …

Dubnau and Losick (2006), *Mol. Microbiol.*, 61 (3):564–72

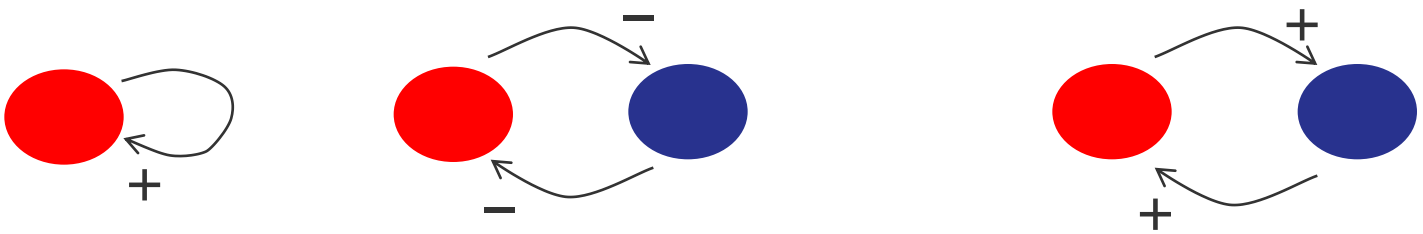
- Can we find general **design principles**, relating network structure to bistability and other properties of network dynamics?

Alon (2007), *An Introduction to Systems Biology,* Chapmann&Hall/CRC

Necessary condition for bistability

- **Necessary condition** for bistability, or multistability, is the occurrence of **positive feedback** loops in the regulatory network

Thomas and d'Ari (1990), *Biological Feedback,* CRC Press



- Increasingly general mathematical proofs of necessary condition for bistability, or multistability, in regulatory networks

  Regulatory interactions (activation/inhibition) lead to non-zero signs (+/-) in Jacobian matrix

  Soulé(2003), *ComPlexUs,* 1:123-133

- **Condition is not sufficient**, as the actual occurrence of bistability depends on parameter values

Other ODE models

- Circadian clock in mammals

  Leloup and Goldbeter (2003), *Proc. Natl. Acad. Sci. USA*, 100(12):7051-7056

- Cell cycle in yeast

  Chen *et al.* (2004), *Mol. Biol. Cell*, 15(8):3841-3862

- Carbon starvation in bacteria

  Bettenbrock (2005), *J. Biol. Chem.*, 281(5):2578-2584

- Signal transduction cascades and developmental decisions

  Ferrell and Machleder (1998), *Science*, 280(5365):852-853

- Pattern formation in fruit fly embryon

  Jaeger *et al.* (2004), *Nature*, 430(6997):368-371

Evaluation of differential equations

- **Pro**: general formalism for which powerful analysis and simulation techniques exist
- **Pro**: well-developed theoretical framework for application to genetic regulatory networks
- **Contra**: numerical techniques are often not appropriate due to lack of quantitative data on model parameters
- **Contra**: assumptions of continuous and deterministic change of concentrations may not be valid on molecular level

Lack of quantitative information: strategies

- Three main strategies to deal with lack of quantitative data:
  - **Parameter sensitivity and robustness**
  - Parameter estimation from time-series data
  - Model reduction

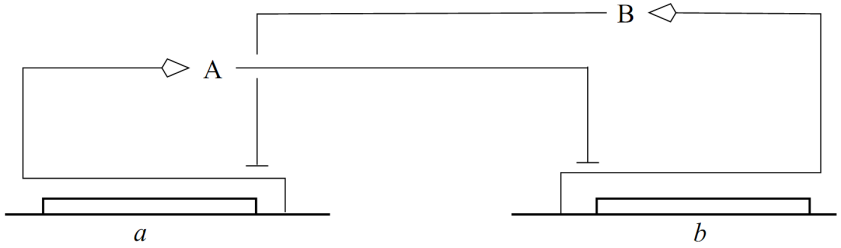
De Jong and Ropers (2006), *Brief. Bioinform.*, 7(4):354-363

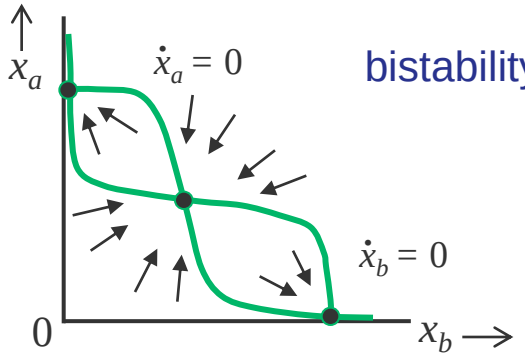
Lack of quantitative data: robustness

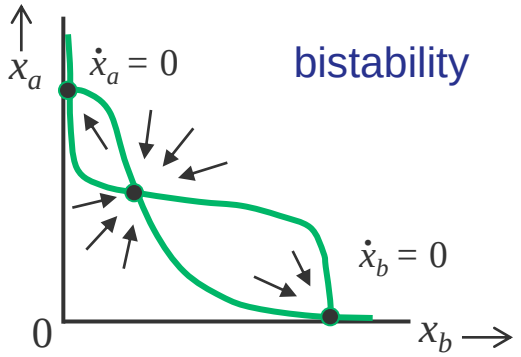
- Important dynamic properties are expected to be **robust** over large ranges of parameter values

  Important dynamic properties should be insensitive to moderate variations in parameter values

Stelling *et al.* (2004), *Cell*, 118(6):675-685



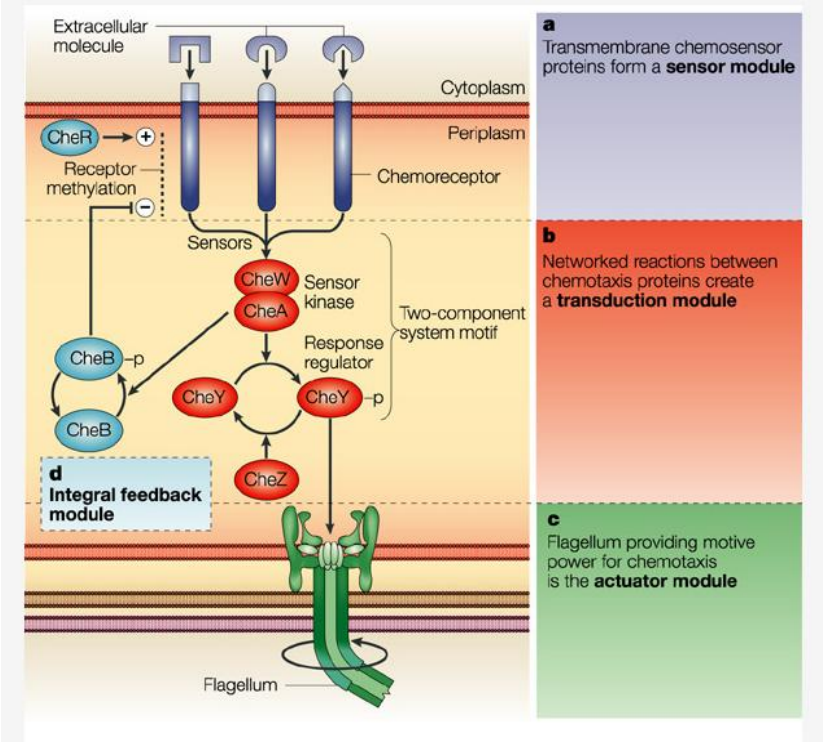


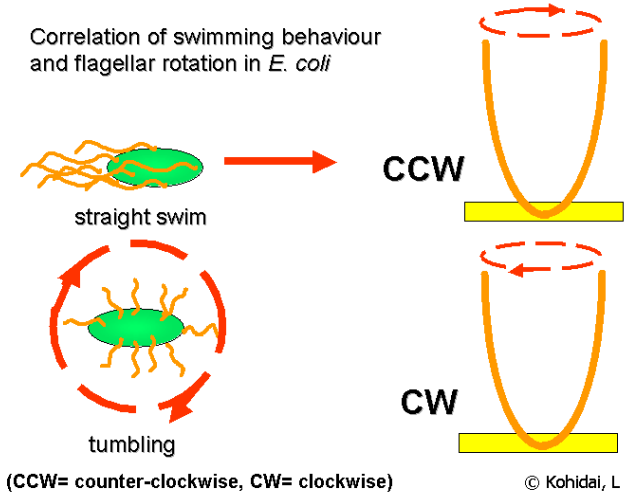


Robustness in *E. coli* chemotaxis

- **Chemotaxis** in bacteria is ability to sense gradient of chemical ligands in environment

  Adjustment of tumbling frequency of molecular motor

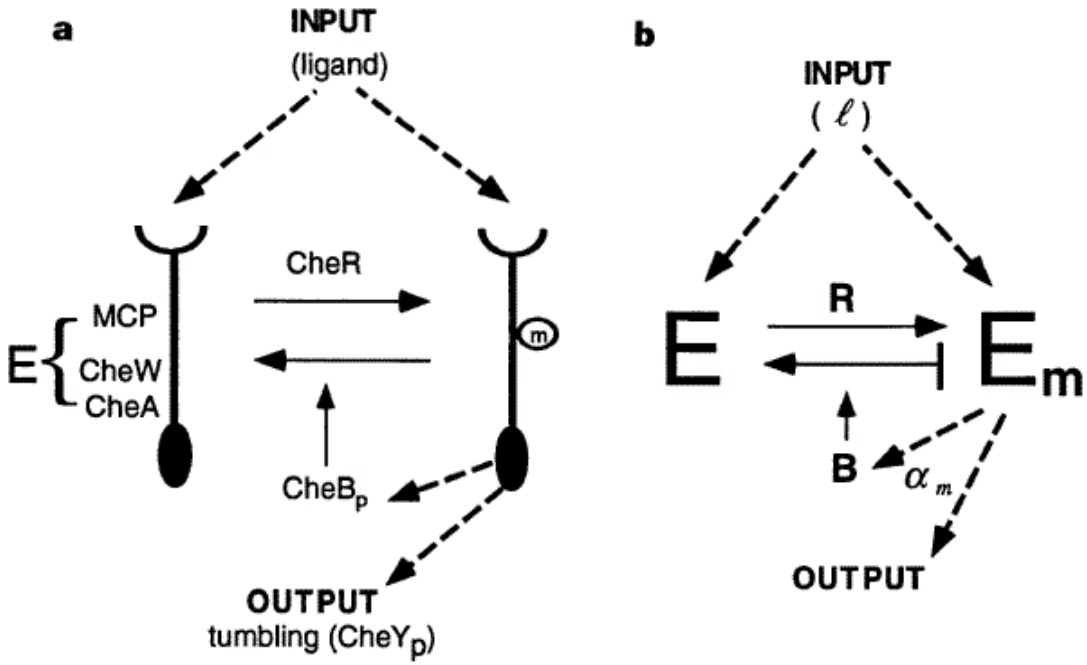




McAdams *et al.* (2004), *Nat. Rev. Genet.*, 5:169-178

Robustness in *E. coli* chemotaxis

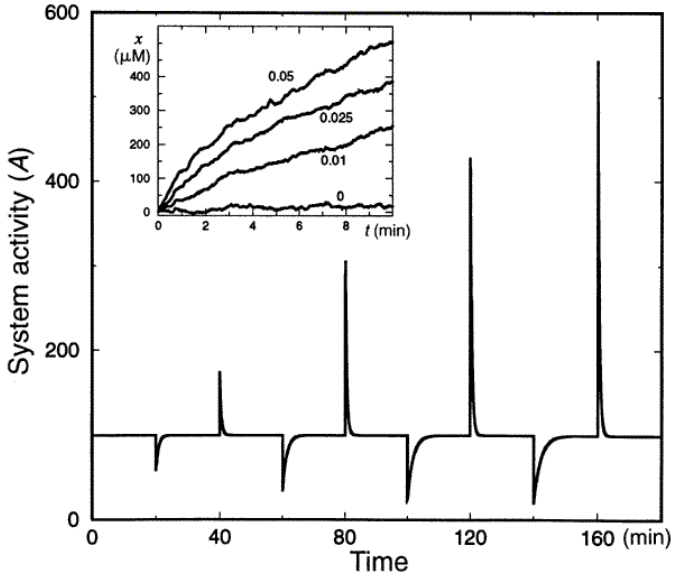
- Differential equation model of signal transduction network underlying bacterial chemotaxis



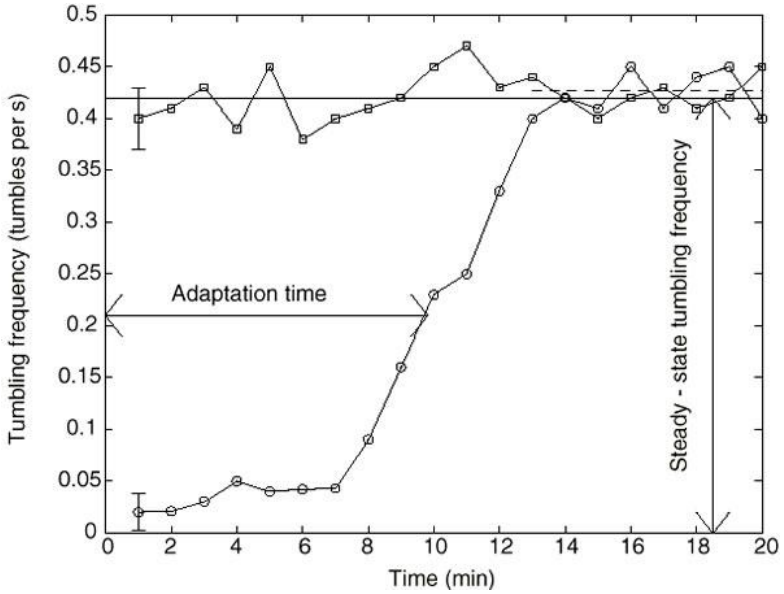
Barkai and Leibler (1997), *Nature*, 387(6636):913-917

Robustness in *E. coli* chemotaxis

- **Adaptation** property is insensitivity of steady-state tumbling frequency to ligand concentration
- Robustness of adaptation property over wide range of parameter values (model and experiments)



Barkai and Leibler (1997), *Nature*, 387:913-917



Alon *et al.* (1999), *Nature*, 397:168-171

Lack of quantitative information: strategies

- Three main strategies to deal with lack of quantitative data:
  - Parameter sensitivity and robustness
  - **Parameter estimation from time-series data**
  - Model reduction

De Jong and Ropers (2006), *Brief. Bioinform.*, 7(4):354-363

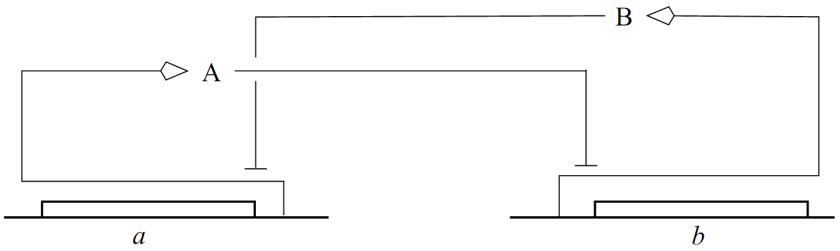
Lack of quantitative data: estimation

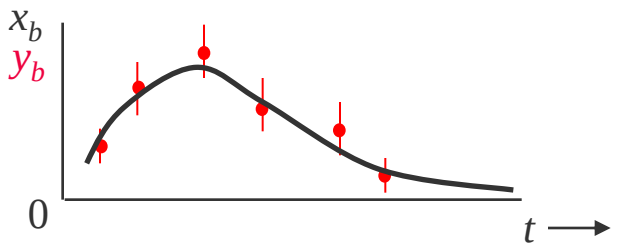
- **Estimate** parameter values from experimental time-series data

  Systems identification in control and engineering

  Ljung (1999), *System Identification: Theory for the User*

- Given model structure, search parameter values for which model predictions best fit experimental data





- Minimization of objective function, for instance sum of squared errors:

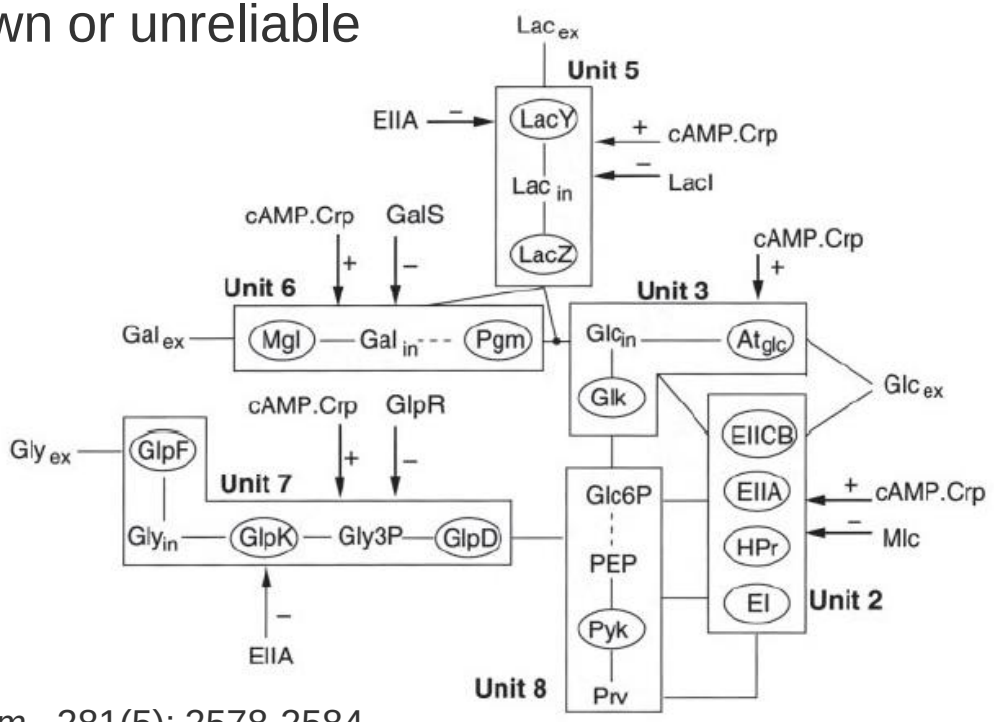
$$\sum_t (x(t,\theta) - y(t))^2$$

  Possibility to add constraint or penalty terms to restrict parameter space

Estimation of parameter values

- Nonlinear differential equation model of uptake of carbon sources (glucose, lactose, glycerol, …) by *E. coli*

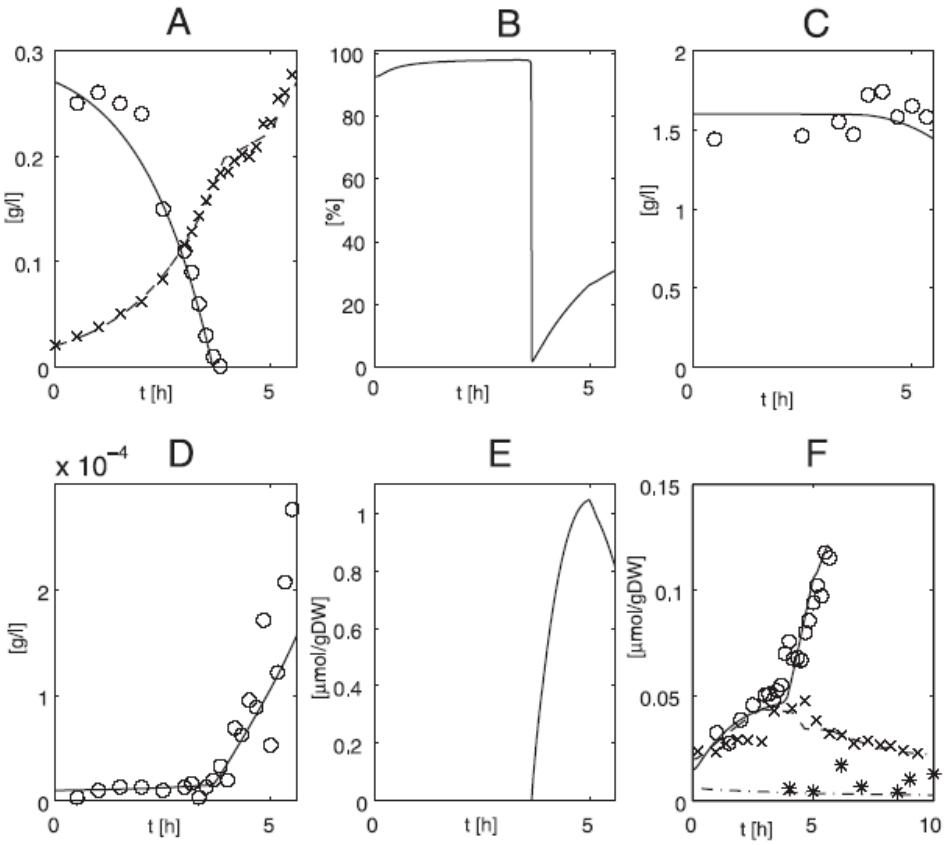
  Several dozens of equations and more than a hundred parameters, many of them unknown or unreliable



Bettenbrock *et al.* (2005), *J. Biol. Chem.*, 281(5): 2578-2584

Estimation of parameter values

- Estimation of parameter values from time-series measurements of metabolite concentrations on wild-type and mutant strains



Bettenbrock *et al.* (2005), *J. Biol. Chem.*, 281(5): 2578-2584

Limitations of system identification

- No algorithms that guarantee globally optimal solution for parameter estimation in nonlinear models

  Evolutionary algorithms, simulated annealing, genetic algorithms, …

- Model identifiability demands experimental data of sufficient quantity and quality

  Common problems: noise, sampling density, unobserved variables, …

  Van Riel (2006), *Brief. Bioinform.*, 7(4):364-374

- However, models of cellular regulatory networks may be non-identifiable by principle, and …

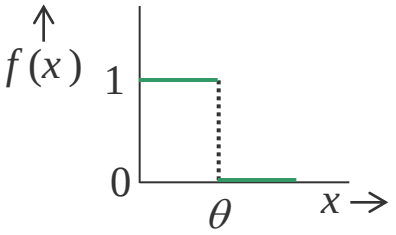
  … even partially identifiable models may yield interesting results

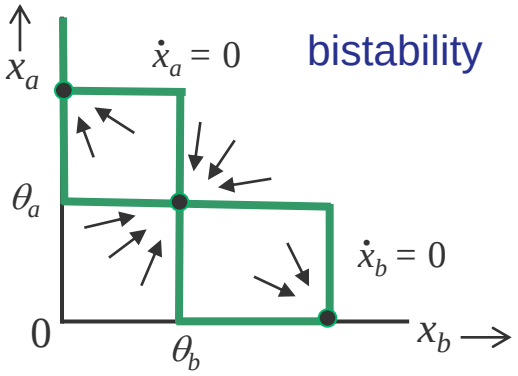
Lack of quantitative data: reduction

- Use **model reduction** to obtain simpler models that can be analyzed with less information on parameter values

  Piecewise-linear instead of nonlinear models

$$\dot{x}_a = \kappa_a f(x_b) - \gamma_a x_a$$
$$\dot{x}_b = \kappa_b f(x_a) - \gamma_b x_b$$





Glass and Kauffman (1973), *J. Theor. Biol.*, 39(1):103-29

de Jong *et al.* (2004), *Bull. Math. Biol.*, 66(2):301-340

- Other example of model reduction: quasi-steady state assumption

Heinrich and Schuster (1996), *The Regulation of Cellular Systems*

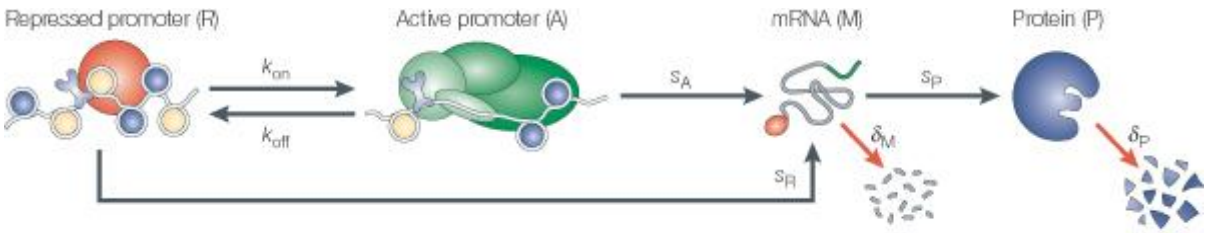
Stochasticity in gene expression

- ODE models make abstraction of underlying biochemical reaction processes involved in gene expression that may not be warranted

  Kaern *et al.* (2005), *Nat. Rev. Genet.*, 6(6):451-464

- Gene expression is **stochastic** instead of **deterministic** process

  Stochasticity gives rise to fluctuations in gene products (**noise**)



- **Discrete** number of molecules of reaction species, instead of **continuous** concentrations

  Noise amplified by low number of molecules of each species

Stochasticity in gene expression

- Major question is how cells both tolerate and exploit noise.

  Rao *et al.* (2002), *Nature*, 420(6912):231-237

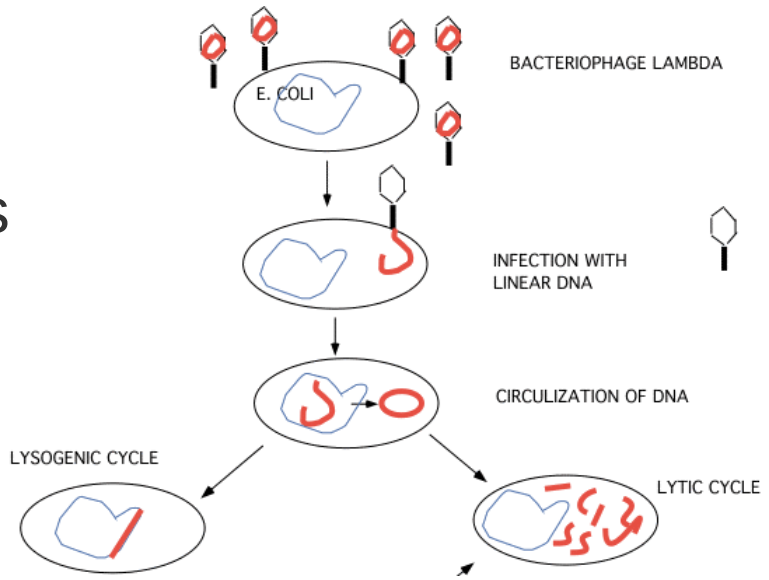
  Raj and van Oudenaarden (2008), *Cell*, 135(2):216-26

- Most cellular processes are **robust** to noise, despite stochasticity of underlying system of biochemical reactions

- Sometimes, intracellular noise drives **population heterogeneity** that may be beneficial for a species

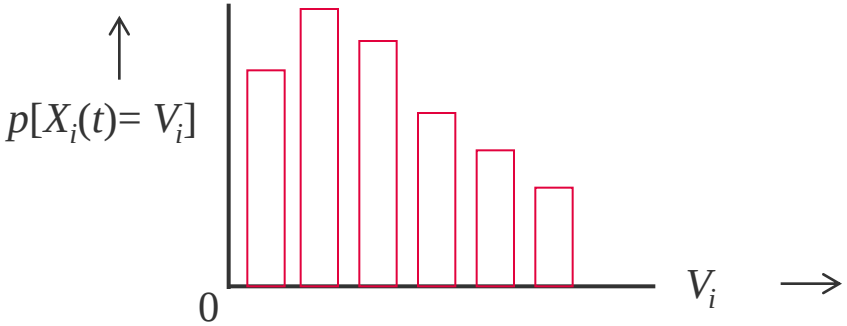
  After infection, only fraction of cells lyse

- ODE models are not suitable for studying origin and effects of noise



Stochastic models

- **Stochastic models** of gene regulation are more realistic
- Number of molecules of each species i at time-point t represented by discrete variable $X_i(t) \in \mathbb{N}$
- Reactions between molecular species lead to change in state of system from $\boldsymbol{X}(t)$ to $\boldsymbol{X}(t+\Delta t)$ over time-interval Δt, where $\boldsymbol{X} = [X_1, \ldots, X_n]'$
- Probability distribution $p[X_i(t)=V_i]$ describes probability that at time-point t there are V_i molecules of species i



Rao *et al.* (2002), *Nature*, 420(6912): 231-237

Stochastic master equation

- Equation describes evolution of state $\boldsymbol{X}$ of regulatory system

$$p[\boldsymbol{X}(t+\Delta t)=\boldsymbol{V}] = p[\boldsymbol{X}(t)=\boldsymbol{V}]\left(1-\sum_{j=1}^{m}\alpha_j\,\Delta t\right) + \sum_{k=1}^{m} p[\boldsymbol{X}(t)=\boldsymbol{V}-\boldsymbol{\nu}_k]\,\beta_k\,\Delta t$$

- m is the number of reactions that can occur in the system
- $\alpha_j\,\Delta t$ is the probability that reaction j will occur in $[t, t+\Delta t]$ given that $\boldsymbol{X}(t)=\boldsymbol{V}$
- $\beta_k\,\Delta t$ is the probability that reaction k will bring the system from $\boldsymbol{X}(t)=\boldsymbol{V}-\boldsymbol{\nu}_k$ to $\boldsymbol{X}(t+\Delta t)=\boldsymbol{V}$ in $[t, t+\Delta t]$

Van Kampen (1997), *Stochastic Processes in Physics and Chemistry*

Stochastic master equation

- For $\Delta t \to 0$ we obtain **stochastic master equation**

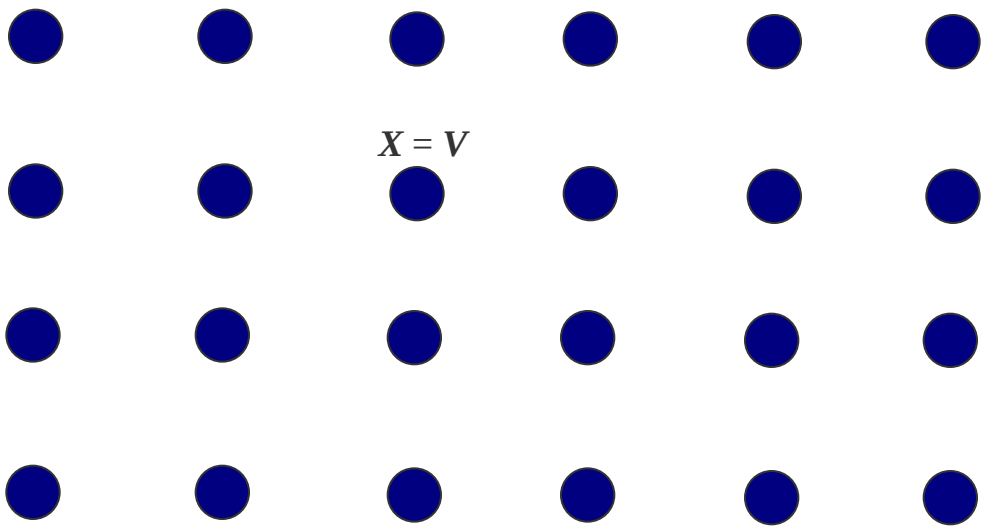
$$\partial p[\boldsymbol{X}(t)=\boldsymbol{V}] / \partial t = \sum_{j=1}^{m} p[\boldsymbol{X}(t)=\boldsymbol{V}-\boldsymbol{\nu}_j]\ \beta_j - p[\boldsymbol{X}(t)=\boldsymbol{V}]\ \alpha_j$$

Van Kampen (1997), *Stochastic Processes in Physics and Chemistry*

- Probabilities α_j, β_j are defined in terms of kinetic constants of reactions
- Analytical solution of master equation is not possible in most situations of practical interest

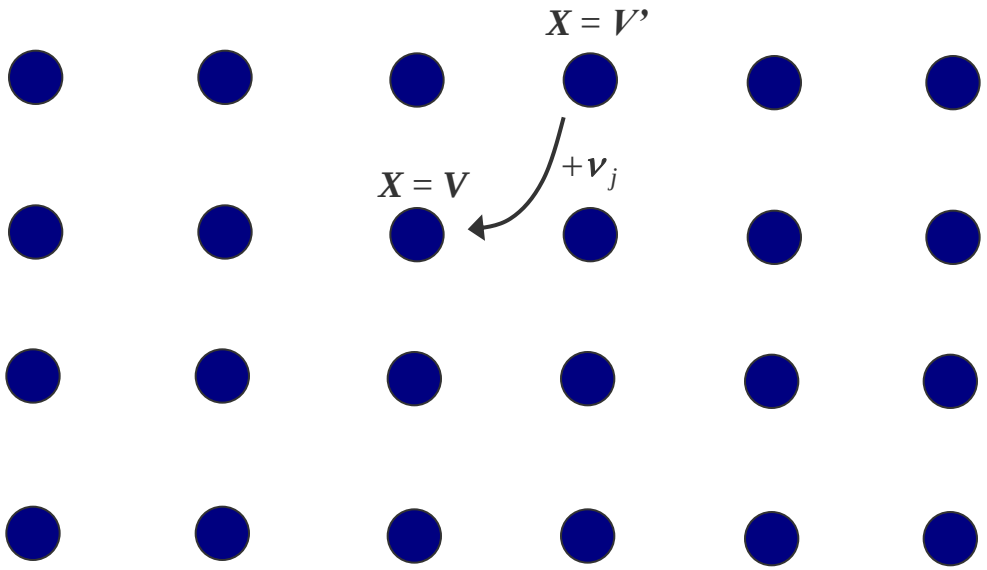
Interpretation of stochastic master equation

- Each state of reaction system corresponds to state of Markov chain with value $\boldsymbol{V}$ for species vector $\boldsymbol{X}$



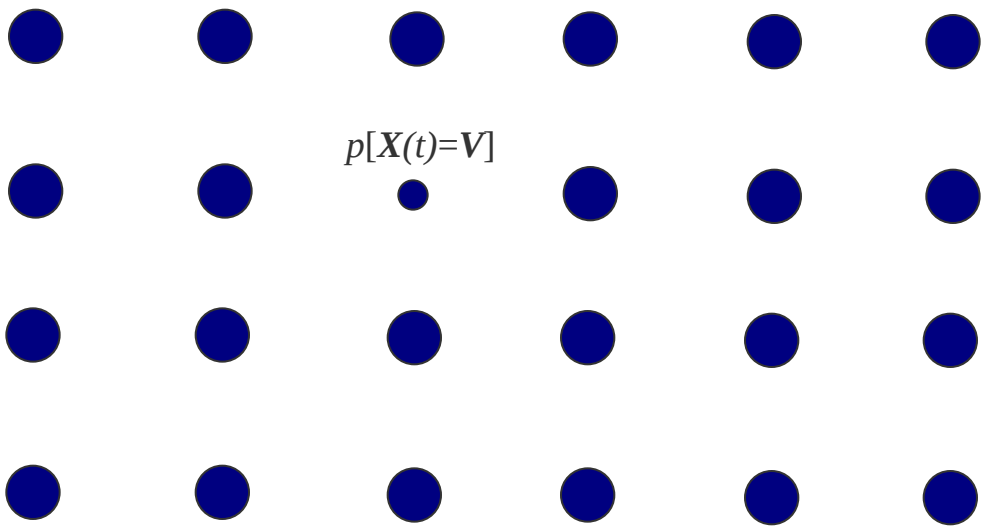
Interpretation of stochastic master equation

- Each reaction j corresponds to change of state in Markov chain, with state update $\boldsymbol{V} = \boldsymbol{V'} + \boldsymbol{\nu}_j$



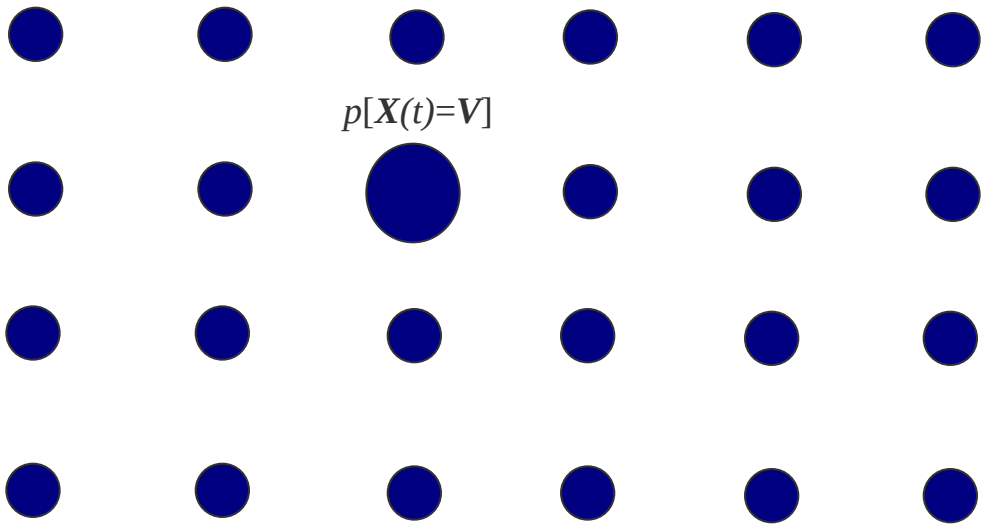
Interpretation of stochastic master equation

- $p[\boldsymbol{X}(t)=\boldsymbol{V}]$ describes probability of state $\boldsymbol{X}=\boldsymbol{V}$ at time t in Markov chain



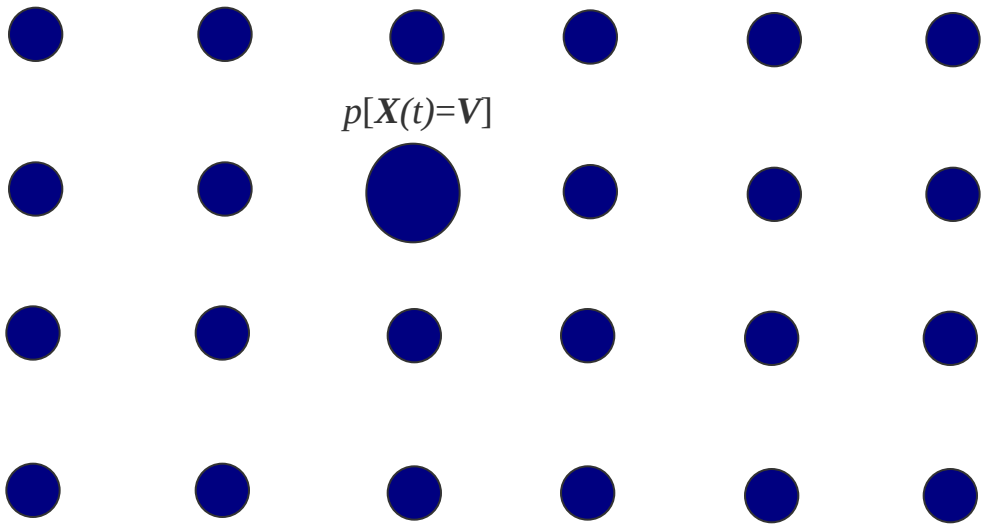
Interpretation of stochastic master equation

- $p[\boldsymbol{X}(t)=\boldsymbol{V}]$ describes probability of state $\boldsymbol{X}=\boldsymbol{V}$ at time t in Markov chain



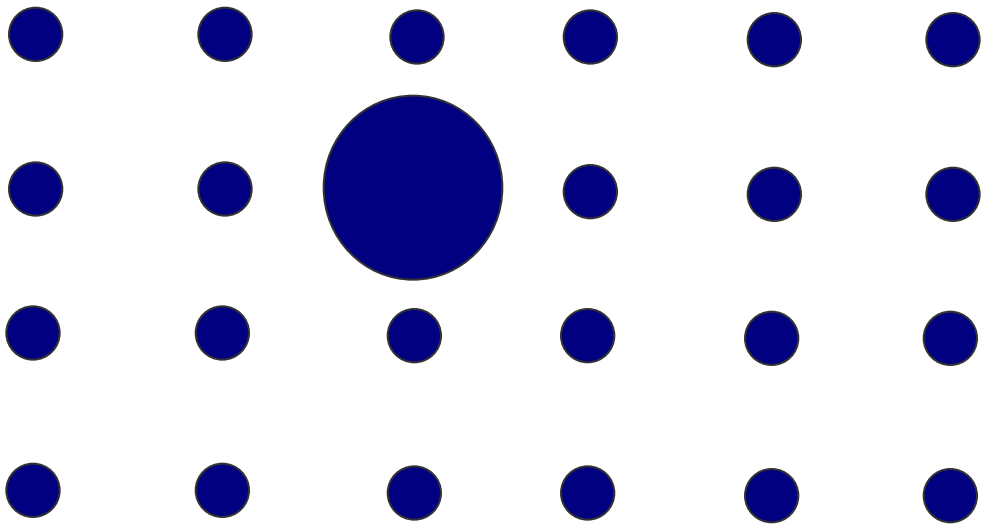
Interpretation of stochastic master equation

- $p[\boldsymbol{X}(t)=\boldsymbol{V}]$ describes probability of state $\boldsymbol{X}=\boldsymbol{V}$ at time t in Markov chain



Interpretation of stochastic master equation

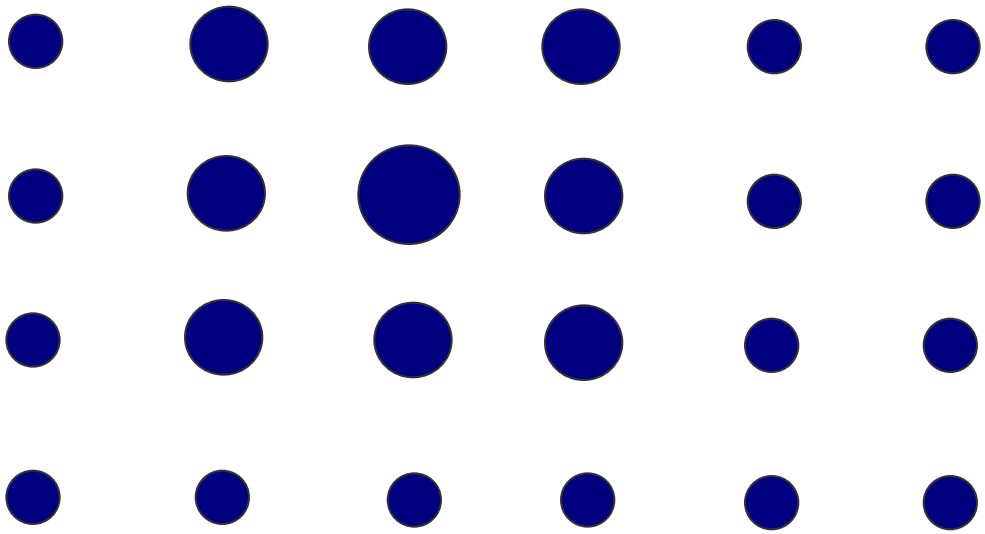
- Stochastic master equations for all states V together describe dynamics of system over time



Time t_0

Interpretation of stochastic master equation

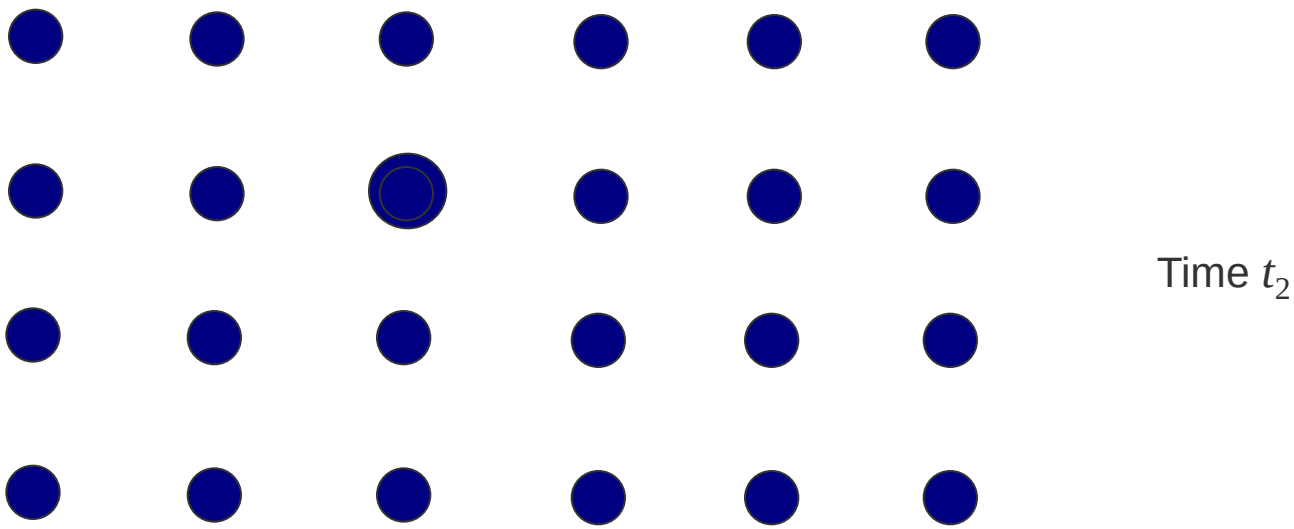
- Stochastic master equations for all states together describe dynamics of system over time



Time t_1

Interpretation of stochastic master equation

- Stochastic master equations for all states together describe dynamics of system over time



Stochastic simulation

- Analytical solution of master equations is not possible in most situations of practical interest
- **Stochastic simulation** predicts sequences of reactions that change state of system, starting from initial state $\boldsymbol{X}(0) = \boldsymbol{V_0}$

Stochastic simulation samples joint probability density function

$p[\tau, j|\boldsymbol{X}(t) = \boldsymbol{V}]$

$\tau =$ time interval until occurrence of next reaction

$j =$ index of next reaction

Probability density function defined in terms of α_j, β_k (reaction constants)

Gillespie (2002), *J. Phys. Chem.*, 81(25): 2340-61

Gillespie (2007), *Annu. Rev. Phys. Chem.*, 58:35-55

Stochastic simulation

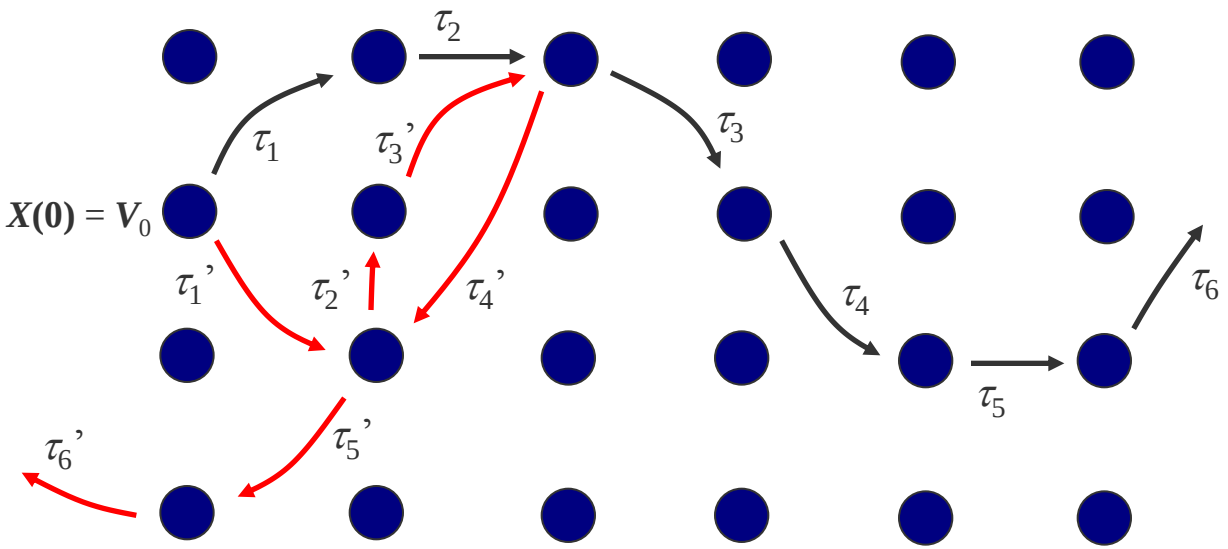
- Analytical solution of master equations is not possible in most situations of practical interest
- **Stochastic simulation** predicts sequences of reactions that change state of system, starting from initial state $\boldsymbol{X}(0) = \boldsymbol{V_0}$
- Repeating stochastic simulation many times yields approximation of probability distribution $p(\boldsymbol{X}\,(t\,)=\boldsymbol{V})$, and thus solution of stochastic master equation

Gillespie (2002), *J. Phys. Chem.*, 81(25): 2340-61

Gillespie (2007), *Annu. Rev. Phys. Chem.*, 58:35-55

Stochastic simulation

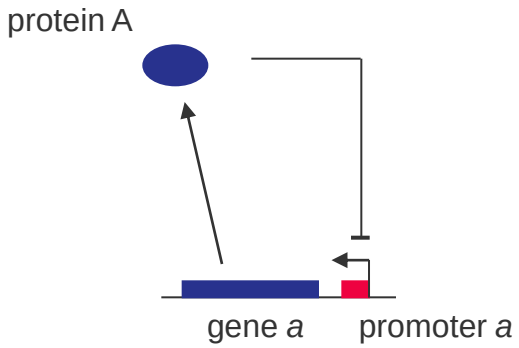
- Stochastic simulation generates sequences of reactions and time intervals between reactions, starting from initial state $\boldsymbol{X}(0)$



- Stochastic simulation may lead to different dynamical behaviors starting from identical initial conditions

Auto-inhibition network

- **Auto-inhibition** network consists of a single gene, coding for transcription regulator inhibiting expression of its own gene

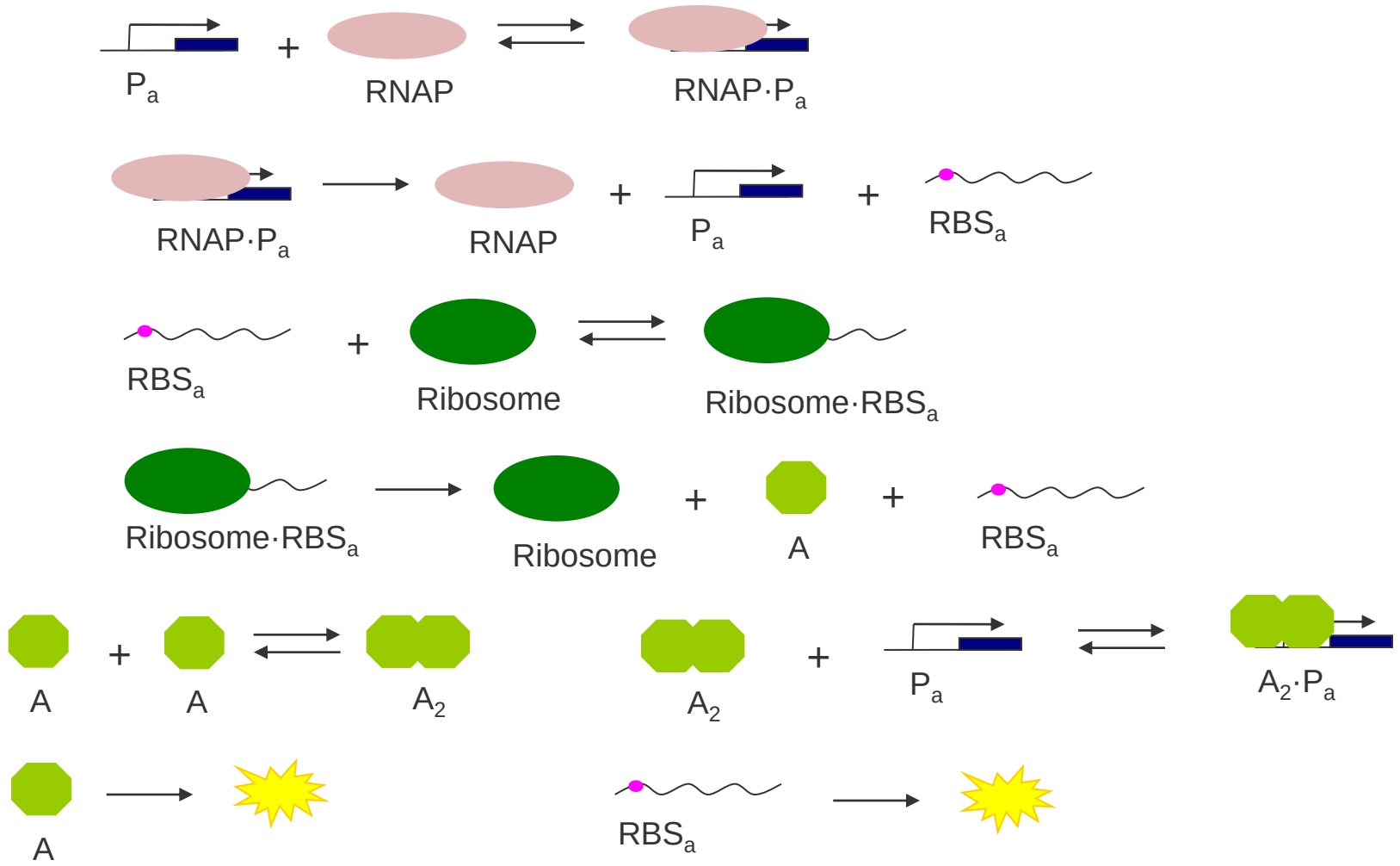


- Auto-inhibition is example of **negative feedback**, and frequently occurs in bacterial regulatory networks

Thieffry *et al.* (1998), *BioEssays*, 20(5):433-440

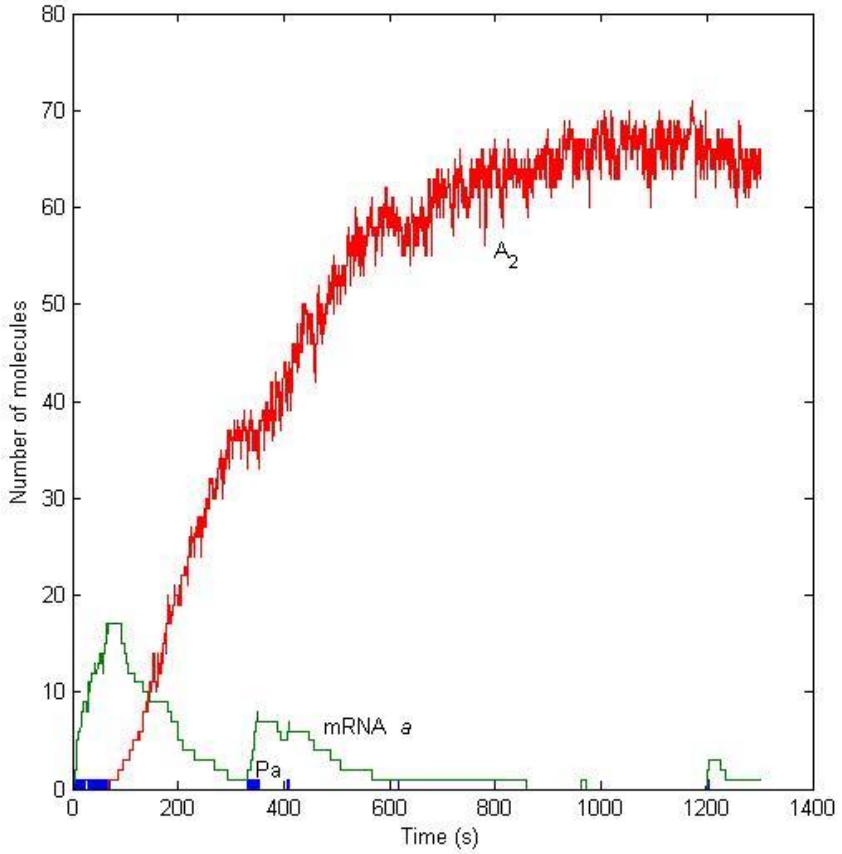
- Development of stochastic model requires list of species, reactions, and kinetic constants

Reactions and species



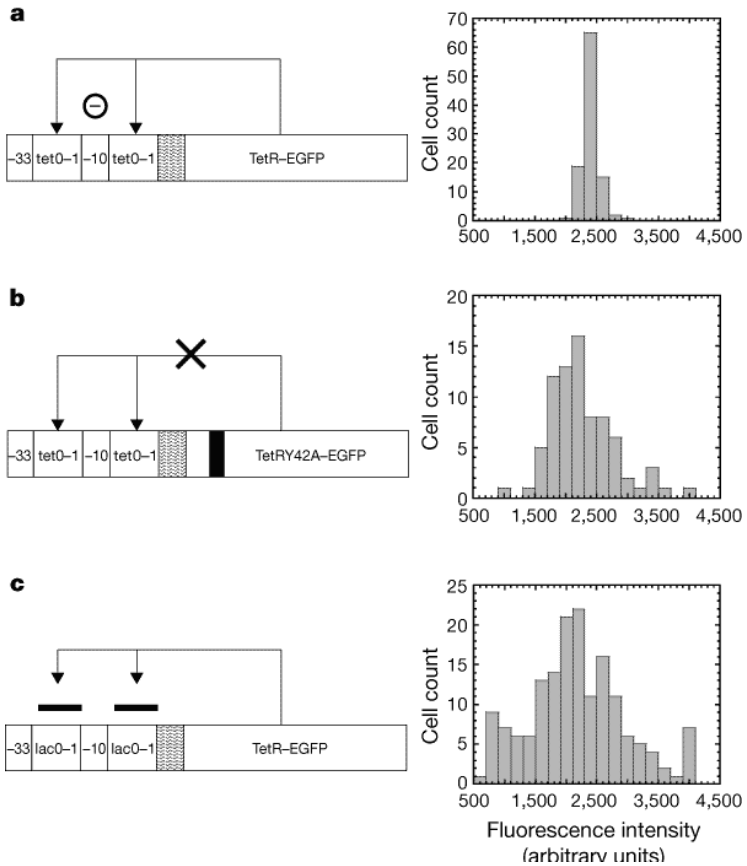
Stochastic simulation of auto-inhibition

- Occurrence of fluctuations and bursts in gene expression



Auto-inhibition and noise reduction

- Auto-inhibition reduces fluctuations in gene expression level

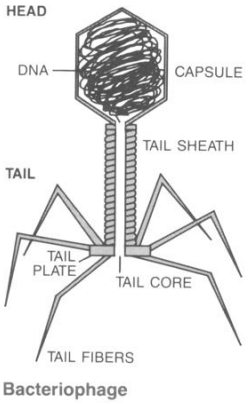


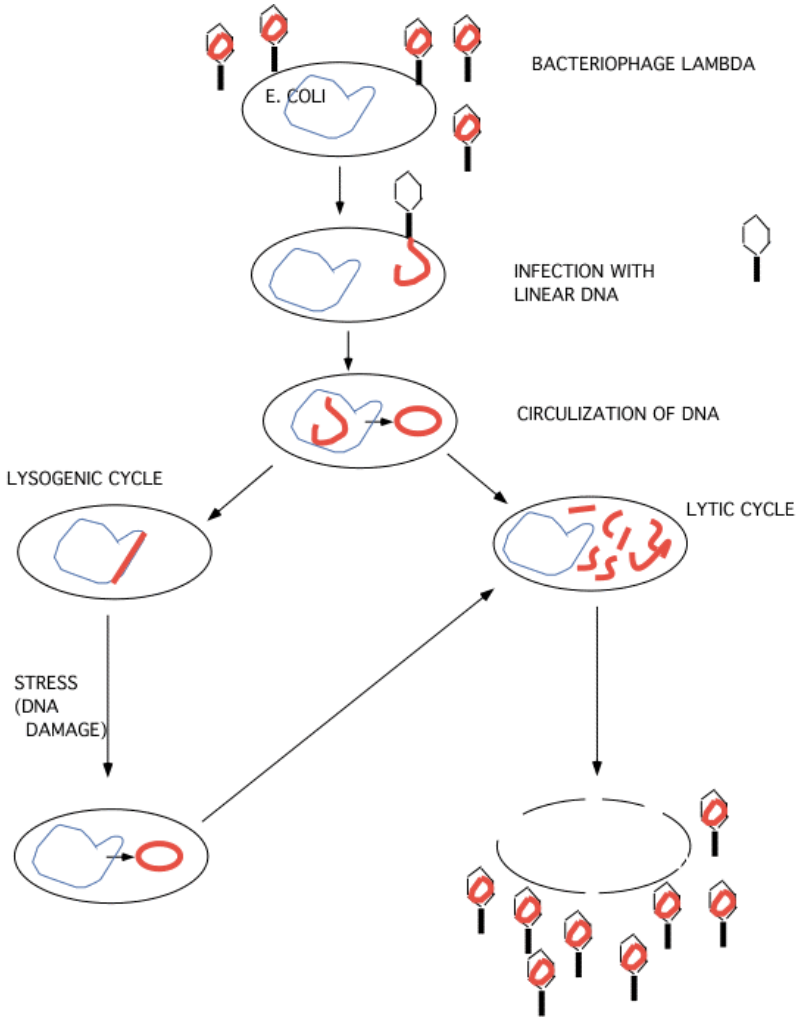
Becskei and Serrano (2000), *Nature*, 405(6785):590-591

Bacteriophage λ infection of *E. coli*

- Response of *E. coli* to phage λ infection involves decision between alternative developmental pathways: **lytic cycle** and **lysogeny**

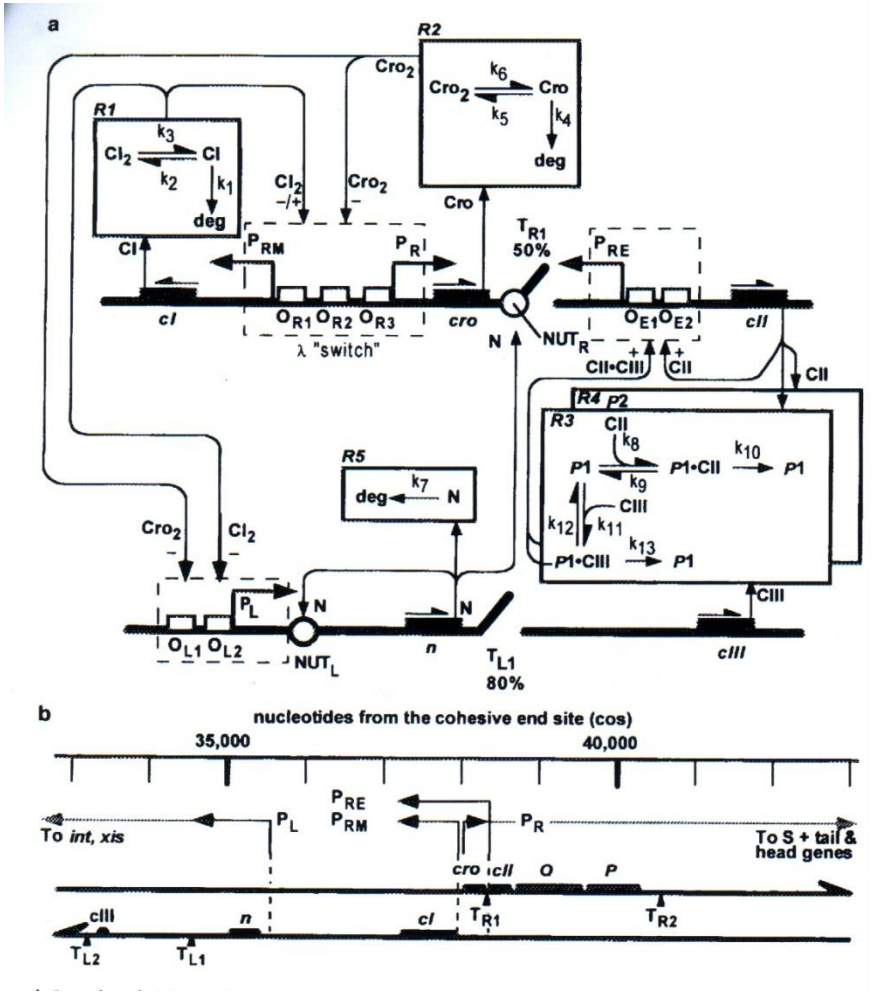
Ptashne (1997), *A Genetic Switch: Phage λ and Higher Organisms*





Stochastic analysis of phage λ infection

- Stochastic model of λ lysis-lysogeny decision network

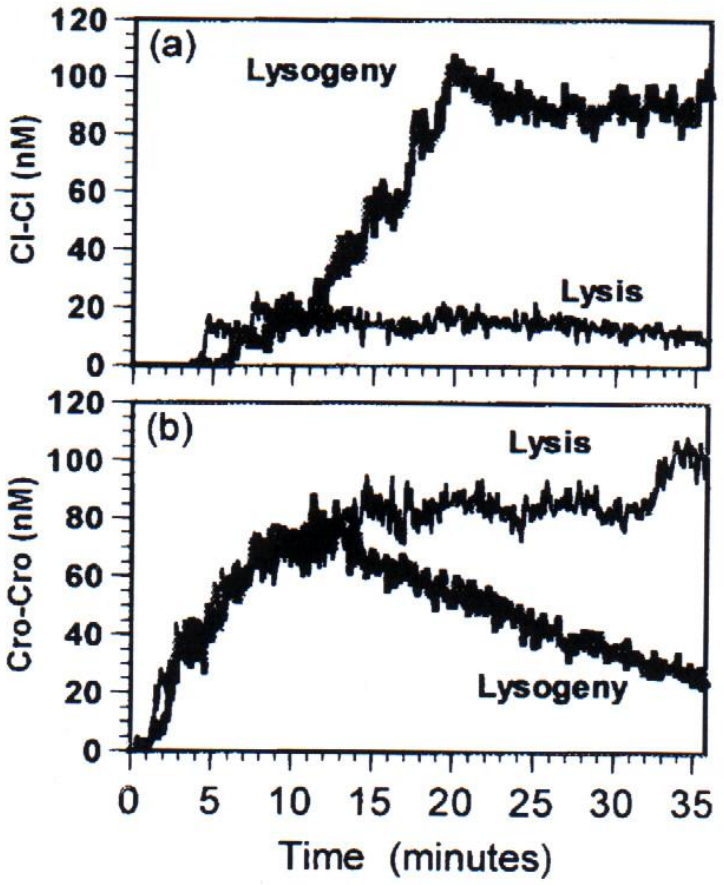


Arkin *et al.* (1998), *Genetics*, 149(4): 1633-1648

Stochastic analysis of phage λ infection

- Time evolution of Cro and CI dimer concentrations

- Due to stochastic fluctuations, under identical conditions cells follow one or other pathway (with some probability)



Arkin *et al.* (1998), *Genetics*, 149(4): 1633-1648

Comparison with deterministic approach

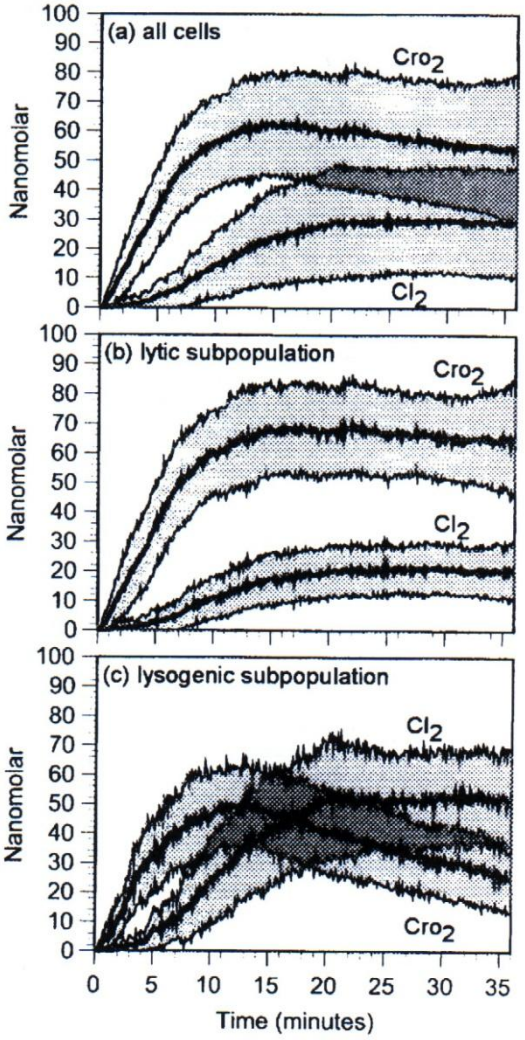
- Deterministic models can be seen as predicting **average behavior** of cell population

  Gillespie. (2000), *J. Chem. Phys.*, 113(1): 297-306

- Analysis of average behavior may obscure that one part of population chooses one pathway rather than another

  Arkin *et al.* (1998), *Genetics*, 149(4): 1633-1648

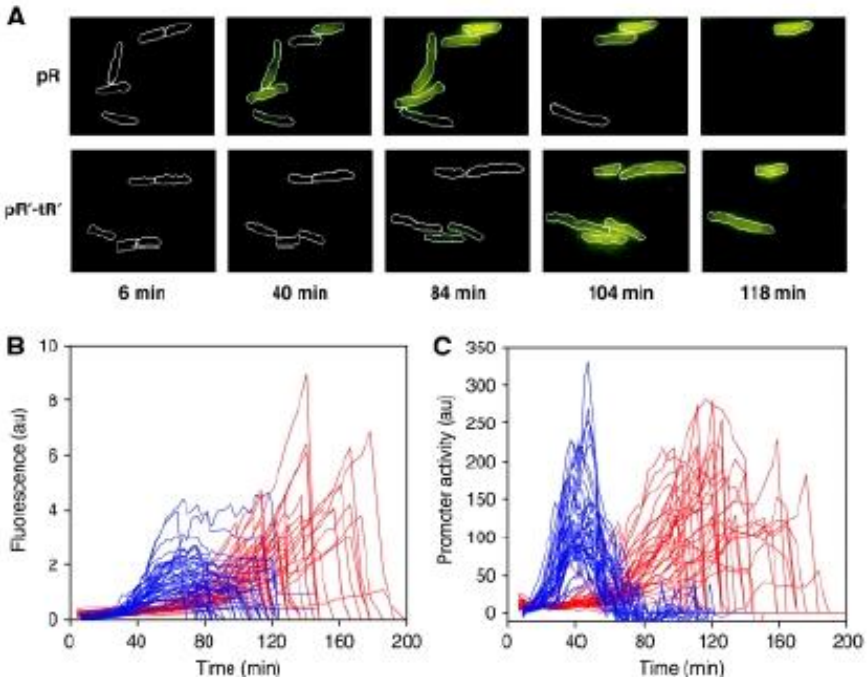
- However, under some conditions deterministic models yield good approximation



Measurements of phage λ infection

- New measurement techniques allow real-time and *in-vivo* monitoring of the execution of lytic and lysogenic pathways in individual cells

  Use of reporter genes in combination with fluorescence microscopy

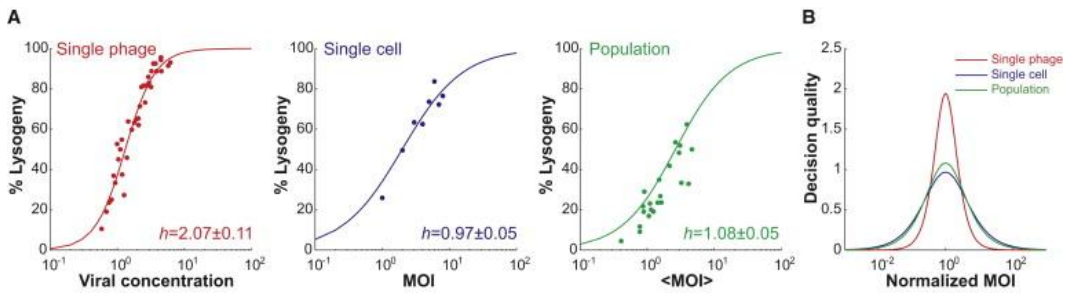


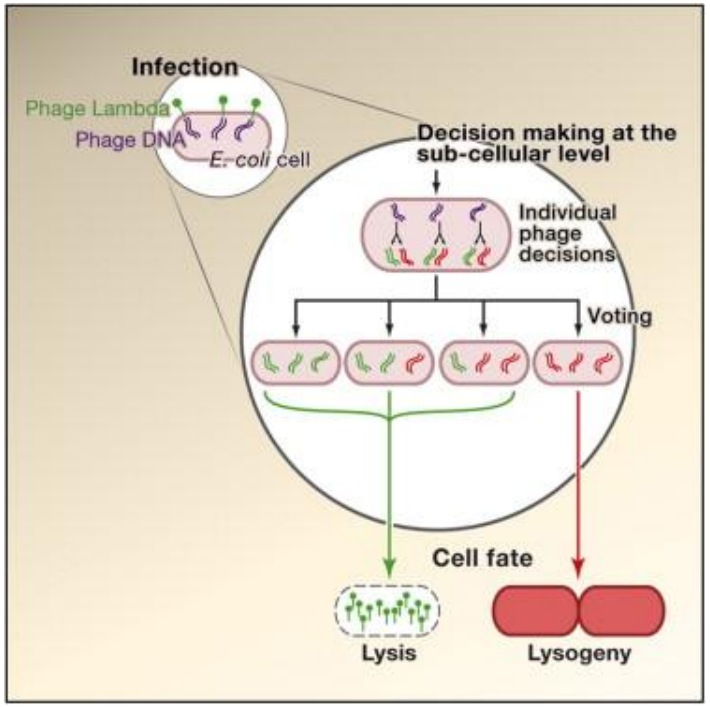
Amir *et al.* (2007), *Mol. Syst. Biol.*, 3:71

Stochasticity and hidden variables

- Is observed population heterogeneity entirely due to stochastic dynamics of biochemical reactions?
- **Hidden variables** that deterministically set outcome of what seems noisy decision process

  Deterministic voting of stochastic decision in single phages





Zeng *et al.* (2010), *Cell*, 141(4):682-91

Other stochastic models

- Effect of noise on carbon assimilation in *E. coli*

  Puchalka and Kierzek (2004), *Biophys. J.*, 86(3):1357-1372

- Regulation of expression of virulence factor in pathogenic *E. coli*

  Jarboe *et al.* (2004), *Biotechnol. Bioengin.*, 88(2):189-203

Evaluation of stochastic equations

- **Pro**: more realistic models of gene regulation
- **Contra**: required information on regulatory mechanisms on molecular level usually not available

  Reaction schemas and kinetic constants, necessary for generating values of parameters τ and ρ, are not or incompletely known

- **Contra**: stochastic simulation is computationally expensive

  Large networks cannot currently be handled, but a host of extensions and approximations have been developed

Conclusions

- Mathematical methods and **computer tools** for modeling and simulation necessary to understand genetic regulatory processes
- Variety of approaches available, representing genetic regulatory systems on different **levels of abstraction**
- Choice of approach depends on **biological problem** and on **available information**:
  - knowledge on reaction mechanisms
  - quantitative data on model parameters and gene expression levels
- Lots of **applications** on bacteria and higher organisms