



# Stochastic modeling of gene regulatory networks

Hidde de Jong  
IBIS  
INRIA Grenoble – Rhône-Alpes  
Hidde.de-Jong@inria.fr

November 24, 2021

# INRIA Grenoble - Rhône-Alpes and IBIS



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

<http://team.inria.fr/ibis>

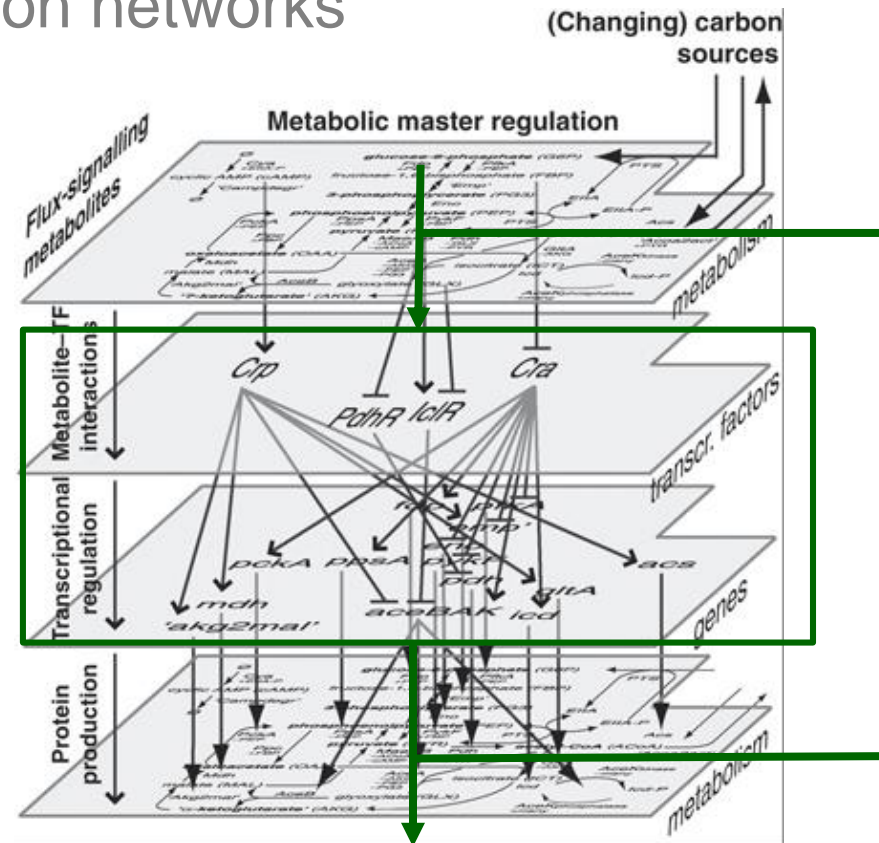


# Overview

- Part 1. Systems biology and kinetic modeling
- Part 2. Metabolic network modeling
- Part 3. Gene regulatory network modeling
  - Quantitative modeling of gene regulatory networks
  - Qualitative modeling of gene regulatory networks
  - **Stochastic modeling of gene regulatory networks**
  - Practical on integrated models of bacterial growth (Matlab)

# Gene regulatory networks

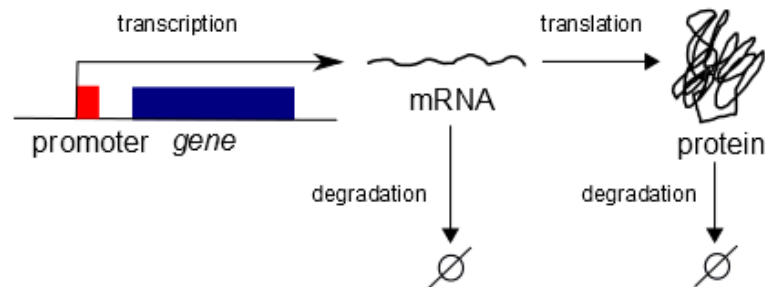
- Focus on **subsystems** that can be studied in isolation due to **modular structure** of reaction networks
  - Time-scale hierarchies
  - Connectivity structure
- **Gene regulatory networks**
  - Genes, proteins, and regulatory interactions
  - Reactions involved in transcription and translation and their regulation
  - Time-scale: min (mRNA) to h (proteins)



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

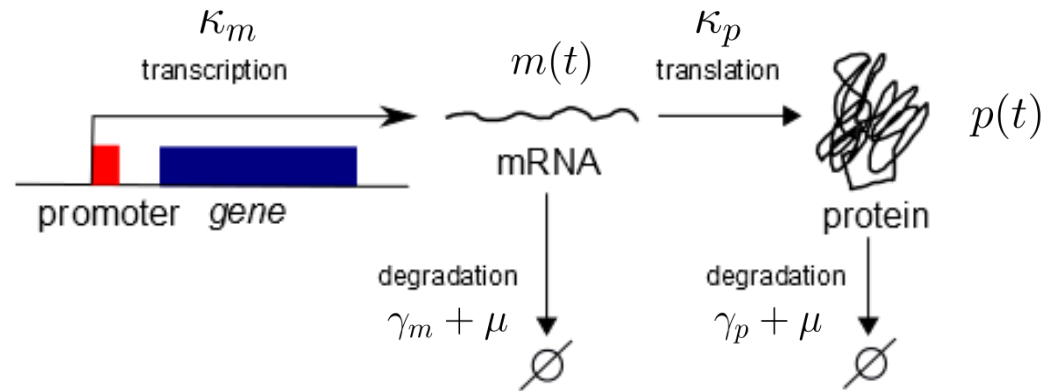


# Modeling of gene regulatory networks

- 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

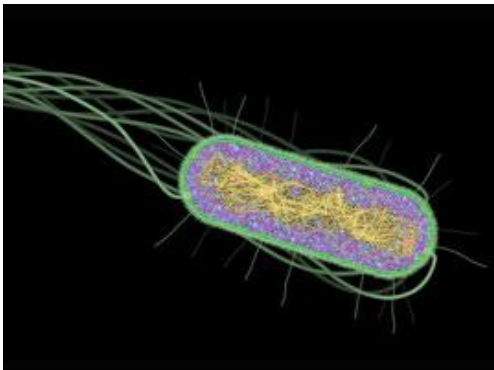
$\gamma_m, \gamma_p > 0$ , degradation rate constants

$\mu \geq 0$ , growth rate

# 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
  - Underlying biochemical reactions are stochastic processes
  - Probability of reaction to occur depends on random encounters of molecules in cell



Goodsell (2010), *The Machinery of Life*, Springer, 2nd ed.

# Stochasticity in gene expression

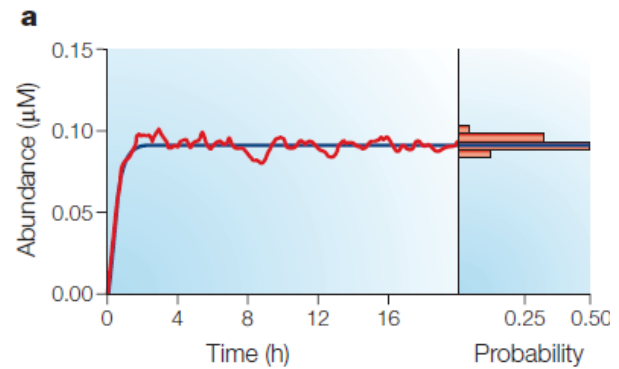
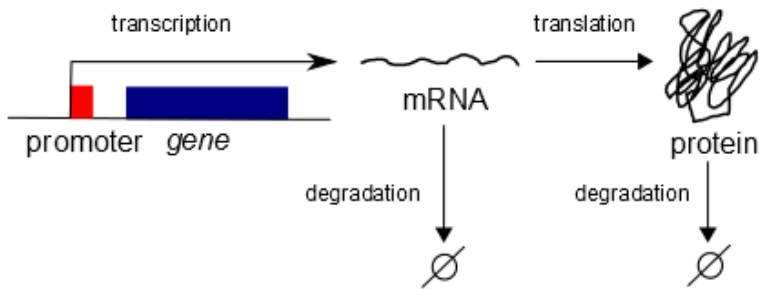
- ODE models make abstraction of underlying biochemical reaction processes involved in gene expression that may not be warranted  
*Kaern et al. (2005), Nat. Rev. Genet., 6(6):451-464*
- Gene expression is **stochastic** instead of **deterministic** process
  - Underlying biochemical reactions are stochastic processes
  - Probability of reaction to occur depends on random encounters of molecules in cell
- **Discrete** number of molecules of reaction species, instead of **continuous** concentrations

Some reactions species involved in gene expression have very low copy numbers (1-10)



# Stochasticity in gene expression

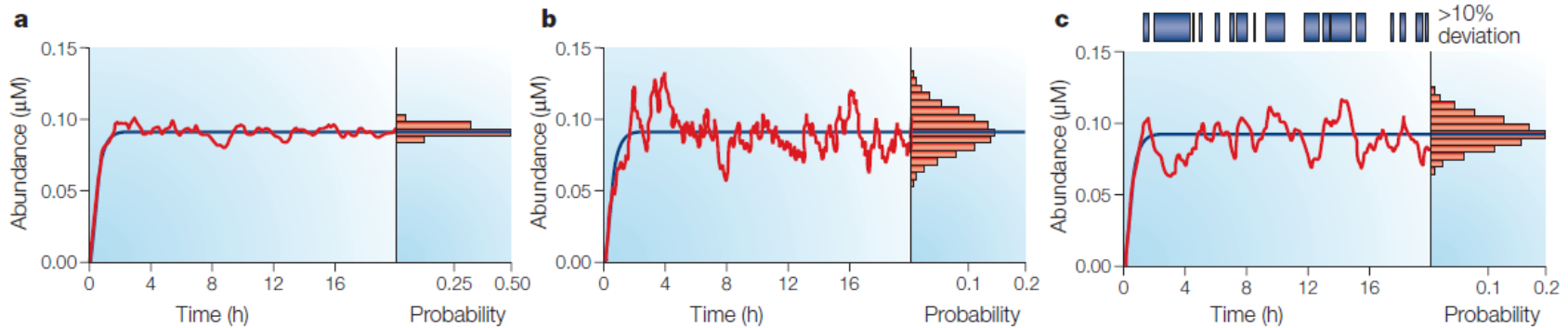
- Stochasticity in gene expression leads to **noise**  
Fluctuations in mRNA and protein concentrations



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

# Stochasticity in gene expression

- Stochasticity in gene expression leads to **noise**  
Fluctuations in mRNA and protein concentrations
- Noise amplified by small number of molecules



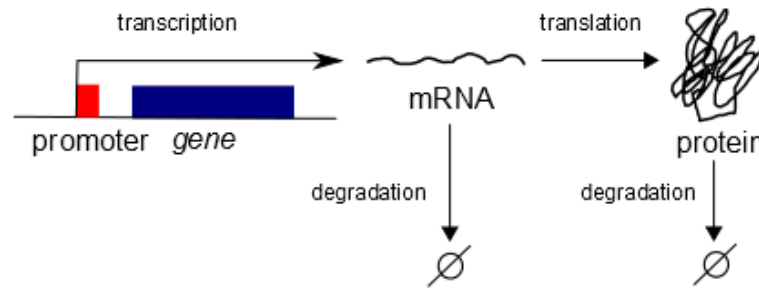
Decrease in number of mRNA and protein molecules wrt **a**

Decrease in number of mRNA molecules wrt **a**, for same number of protein molecules

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

# Stochasticity in gene expression

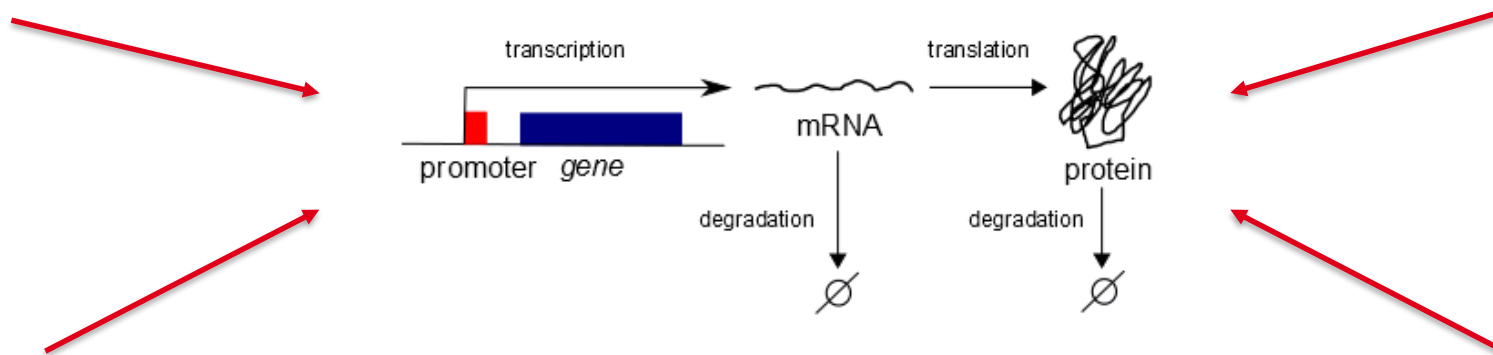
- Stochasticity in gene expression leads to **noise**  
Fluctuations in mRNA and protein concentrations
- Noise amplified by small number of molecules



- Different types of noise:
  - **Intrinsic noise:** fluctuations due to stochasticity of processes involved in gene expression (transcription, translation, ...)

# Stochasticity in gene expression

- Stochasticity in gene expression leads to **noise**  
Fluctuations in mRNA and protein concentrations
- Noise amplified by small number of molecules

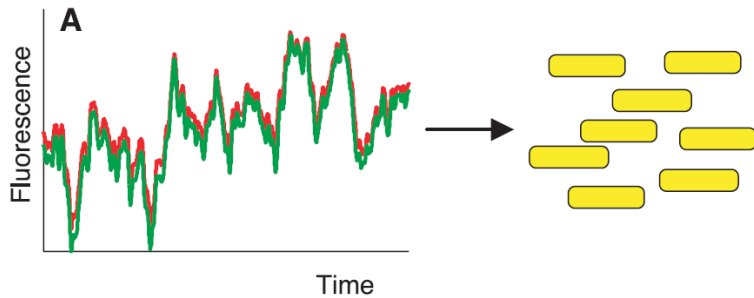


- Different types of noise:
  - **Intrinsic noise:** fluctuations due to stochasticity of processes involved in gene expression (transcription, translation, ...)
  - **Extrinsic noise:** fluctuations due to variability in external factors (temperature, ribosome availability, ...). Impact on rate constants.

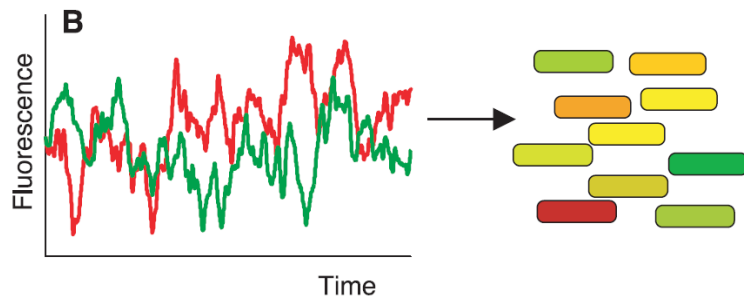
# Stochasticity in gene expression

- Experimental discrimination between intrinsic and extrinsic noise

Expression in a single cell with two different reporter genes (*gfp* and *cfp*) controlled by same promoter



**No intrinsic noise**, so relative amount of both proteins is constant over time and across individual cells in population



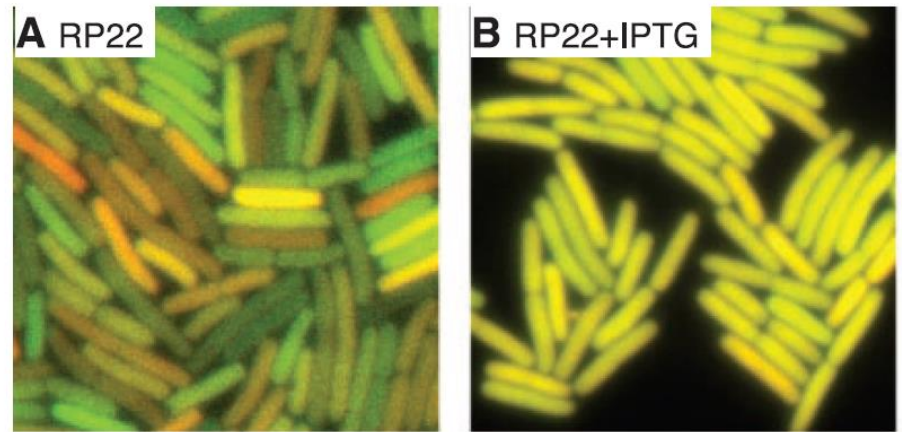
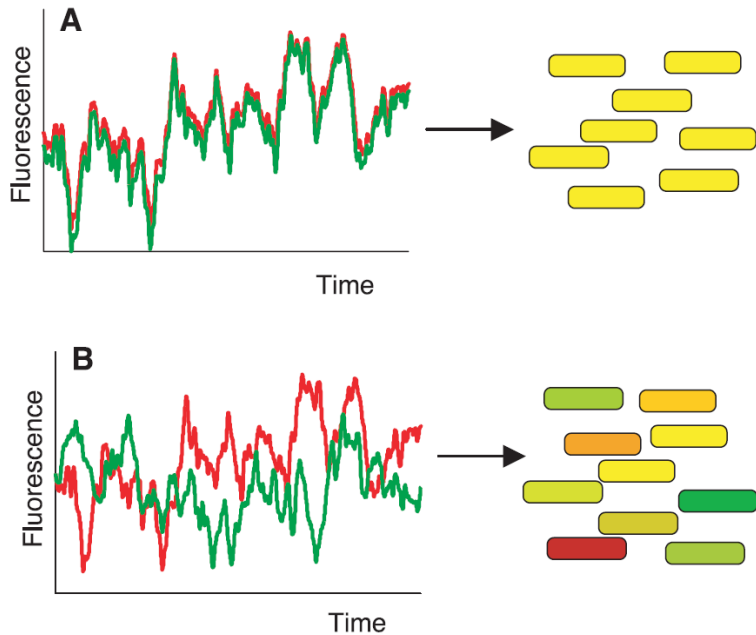
**Intrinsic noise**, so relative amount of both proteins varies over time and across individual cells in population

Elowitz *et al.* (2002), *Science*, 297(5584):1183-6

# Stochasticity in gene expression

- Experimental discrimination between intrinsic and extrinsic noise

Expression in a single cell with two different reporter genes (*gfp* and *cfp*) controlled by same promoter

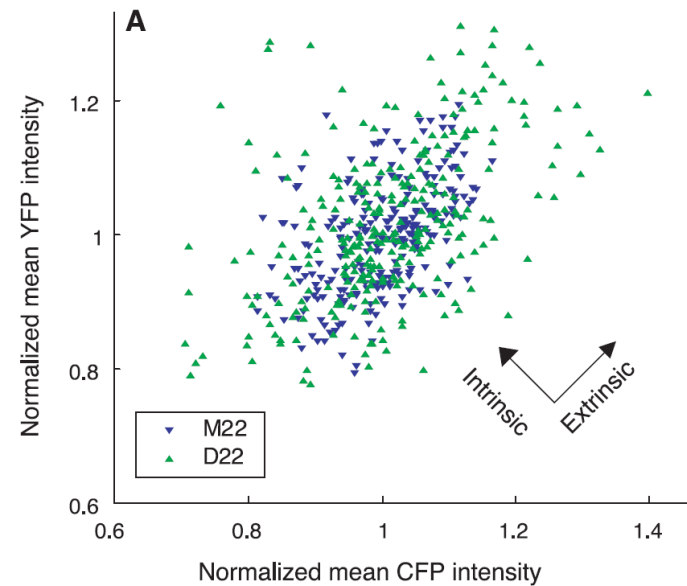
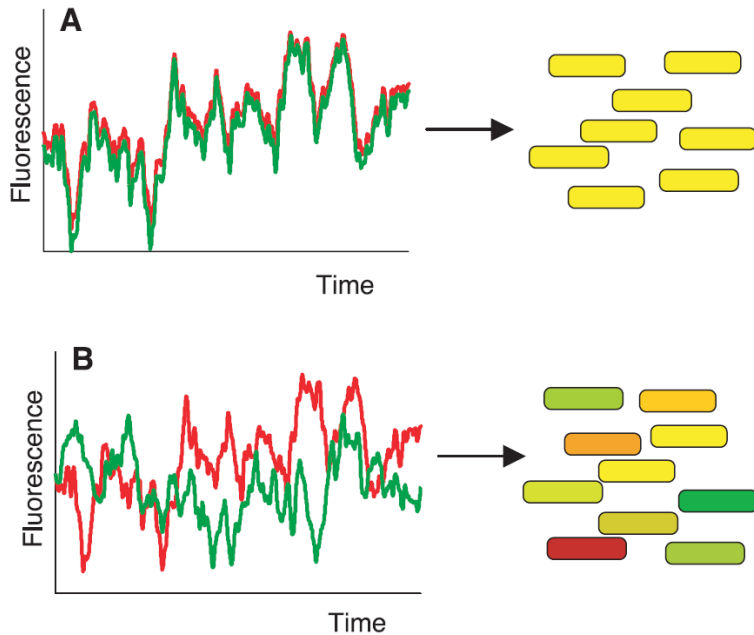


Elowitz *et al.* (2002), *Science*, 297(5584):1183-6

# Stochasticity in gene expression

- Experimental discrimination between intrinsic and extrinsic noise

Expression in a single cell with two different reporter genes (*gfp* and *cfp*) controlled by same promoter



Elowitz *et al.* (2002), *Science*, 297(5584):1183-6

# Stochasticity in gene expression

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

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

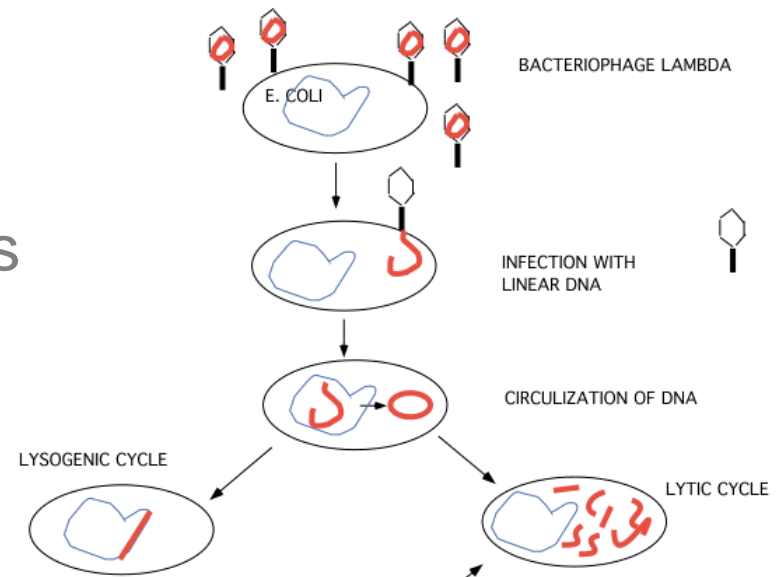
Raj and van Oudenaarden (2008), *Cell*, 135(2):216-26

- 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

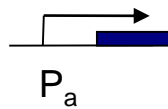
- ODE models are not suitable for studying origin and effects of noise





# Stochastic models

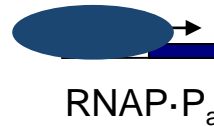
- Stochastic models of gene regulation are more appropriate
- Number of molecules of each species  $i$  at time-point  $t$  represented by discrete variable  $X_i(t) \in \mathbb{N}$



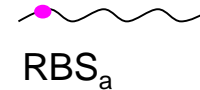
$X_1$



$X_2$



$X_3$

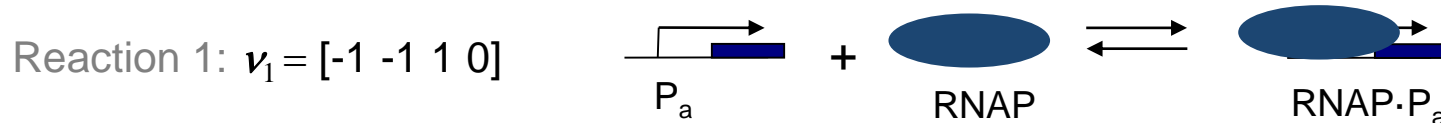


$X_4$

# Stochastic models

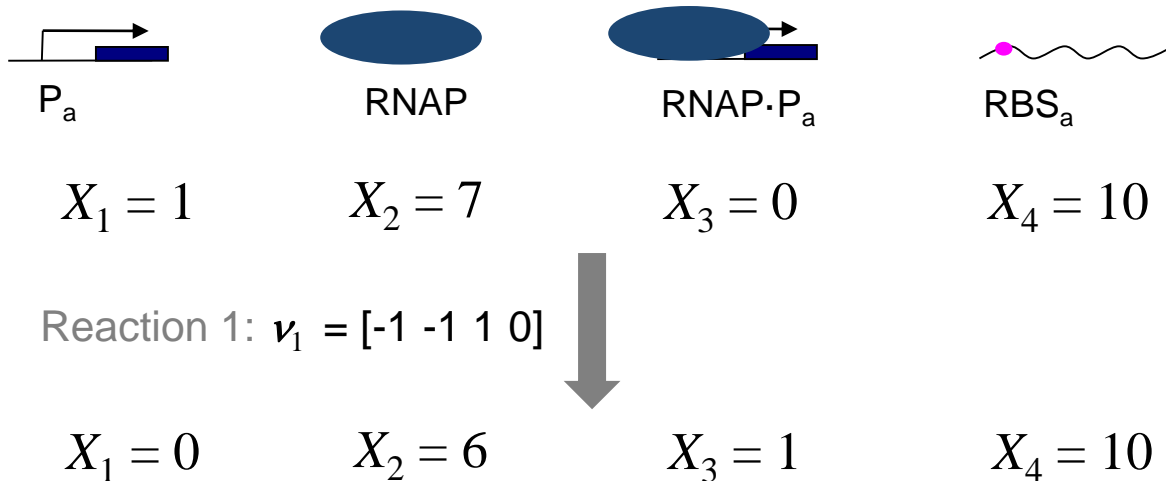
- **Stochastic models** of gene regulation are more appropriate
- Number of molecules of each species  $i$  at time-point  $t$  represented by discrete variable  $X_i(t) \in \mathbb{N}$
- Reactions between molecular species lead to change in state of system from  $\mathbf{X}(t)$  to  $\mathbf{X}(t+\Delta t)$  over time-interval  $\Delta t$ , where  $\mathbf{X} = [X_1, \dots, X_n]'$

Change of state by reaction  $k$  described by vector  $\mathbf{v}_k$



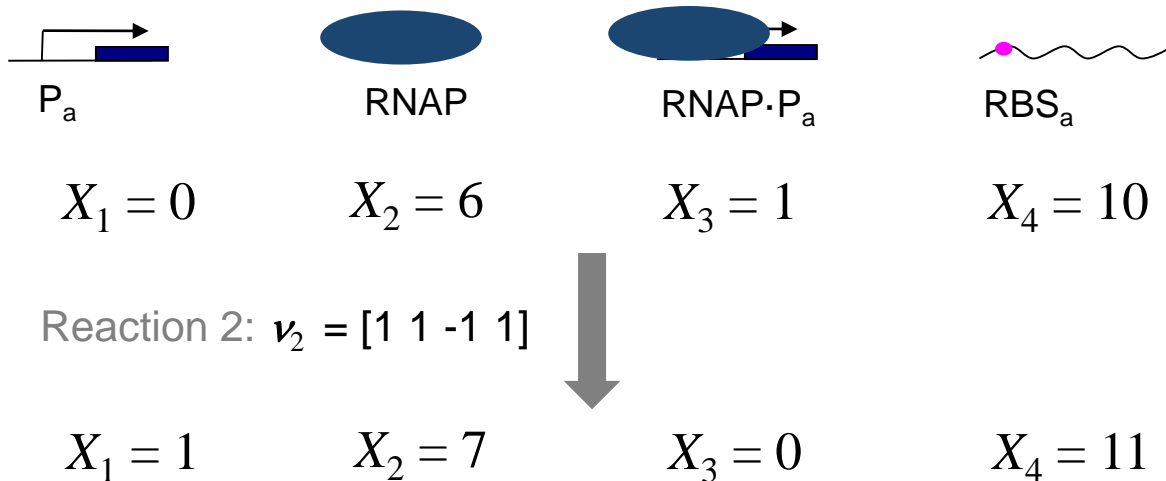
# Stochastic models

- **Stochastic models** of gene regulation are more appropriate
- 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  $\mathbf{X}(t)$  to  $\mathbf{X}(t+\Delta t)$  over time-interval  $\Delta t$ , where  $\mathbf{X} = [X_1, \dots, X_n]'$



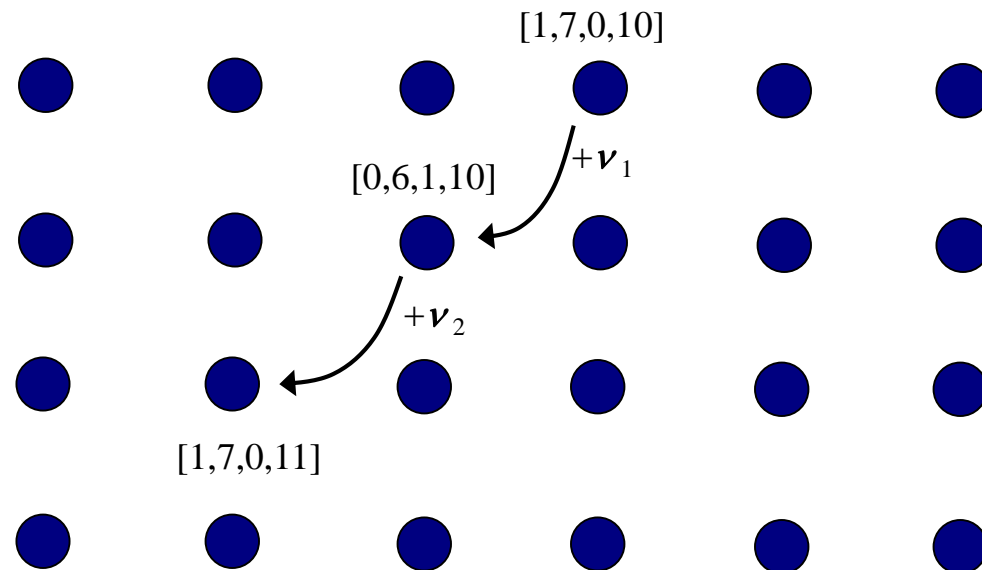
# Stochastic models

- **Stochastic models** of gene regulation are more appropriate
- 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  $\mathbf{X}(t)$  to  $\mathbf{X}(t+\Delta t)$  over time-interval  $\Delta t$ , where  $\mathbf{X} = [X_1, \dots, X_n]'$



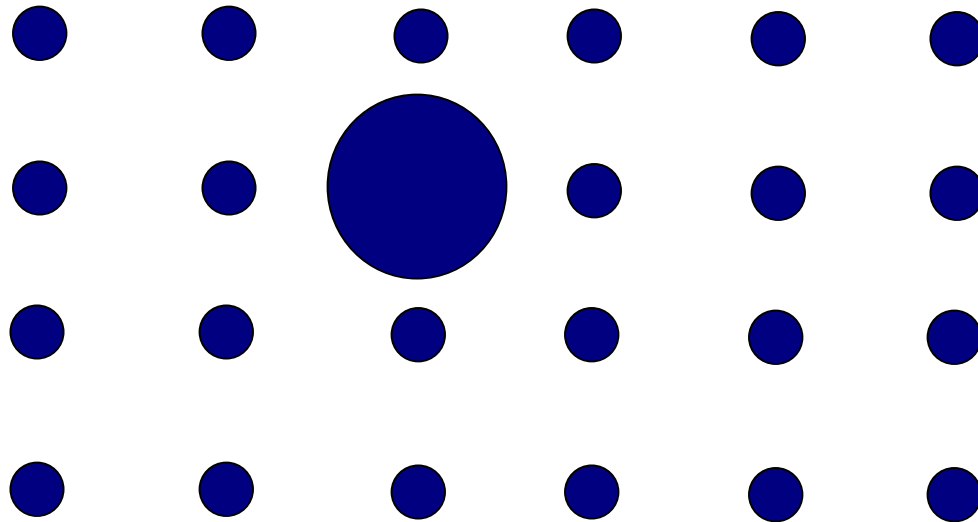
# Stochastic models

- Possible states are given by possible value combinations for variables:  $\mathbf{X} = \mathbf{V}$ , with  $\mathbf{V} = [V_1, \dots, V_n]'$
- Transitions between states are given by possible reactions  $k$



# Stochastic models

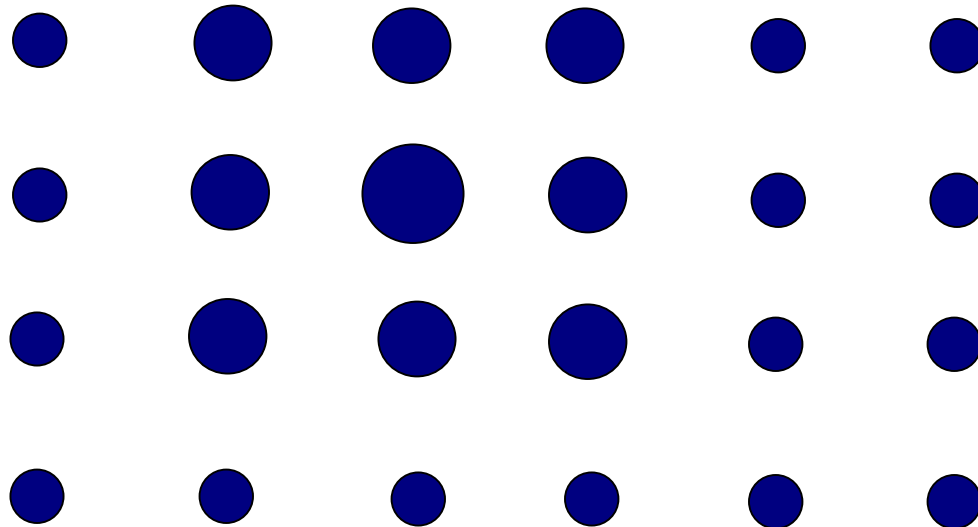
- Probability distribution  $p[\mathbf{X}(t)=\mathbf{V}]$  describes probability that at time-point  $t$  there are  $\mathbf{V} = [V_1, \dots, V_n]'$  molecules



Time  $t_0$

# Stochastic models

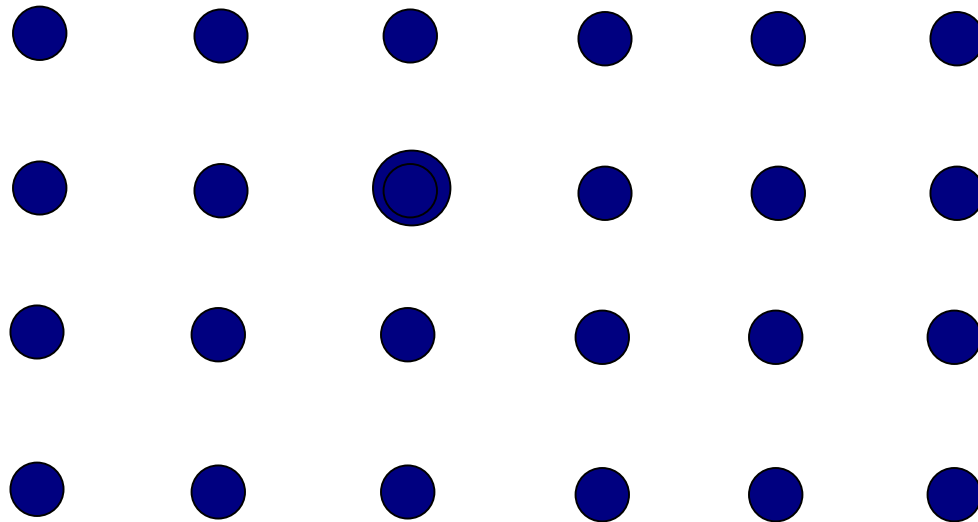
- Probability distribution  $p[\mathbf{X}(t)=\mathbf{V}]$  describes probability that at time-point  $t$  there are  $\mathbf{V} = [V_1, \dots, V_n]'$  molecules



Time  $t_1$

# Stochastic models

- Probability distribution  $p[\mathbf{X}(t)=\mathbf{V}]$  describes probability that at time-point  $t$  there are  $\mathbf{V} = [V_1, \dots, V_n]'$  molecules



Time  $t_2$



# Stochastic master equation

- Evolution of probability distribution  $p[\mathbf{X}(t)=\mathbf{V}]$  given by

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

- $m$  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  $\mathbf{X}(t)=\mathbf{V}$
- $\beta_k \Delta t$  is the probability that reaction  $k$  will bring the system from  $\mathbf{X}(t)=\mathbf{V}-\mathbf{v}_k$  to  $\mathbf{X}(t+\Delta t)=\mathbf{V}$  in  $[t, t + \Delta t]$

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

# Stochastic master equation

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

$$dp[X(t)=V] / dt = \sum_{j=1}^m p[X(t)=V-\nu_j] \beta_j - p[X(t)=V] \alpha_j$$

- Probabilities  $\alpha_j, \beta_j$  are defined in terms of kinetic constants of reactions and number of reactant molecules
- Unimolecular reaction  $j$ :  $S_1 \rightarrow \text{product(s)}$

$$\alpha_j = k_j X_1 (X_1-1)/2$$

- Bimolecular reaction  $j$ :  $S_1 + S_2 \rightarrow \text{product(s)}$

$$\alpha_j = k_j X_1 X_2 / \Omega \qquad \Omega : \text{cell volume}$$

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

# Stochastic master equation

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

$$dp[X(t)=V] / dt = \sum_{j=1}^m p[X(t)=V-\nu_j] \beta_j - p[X(t)=V] \alpha_j$$

- Probabilities  $\alpha_j, \beta_j$  are defined in terms of kinetic constants of reactions and number of reactant molecules
- Analytical solution of master equation is not possible in most situations of practical interest

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

# 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  $\mathbf{X}(0) = \mathbf{V}_0$

- Stochastic simulation samples joint probability density function

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

$\tau$  = time until occurrence of next reaction

$j$  = index of next reaction

- **Interpretation:**  $p[\tau, j/\mathbf{X}(t) = \mathbf{V}]d\tau$  is probability, given  $\mathbf{X}(t) = \mathbf{V}$ , that next reaction will occur in  $[t+\tau, t+\tau+d\tau]$  and is reaction  $j$

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** generates sequences of reactions that change state of system, starting from initial state  $X(0) = V_0$ 
  - Stochastic simulation samples joint probability density function
$$p[\tau, j/X(t) = V]$$

$\tau$  = time until occurrence of next reaction

$j$  = index of next reaction
  - Probability density function defined in terms of  $\alpha_j, \beta_k$  (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** generates sequences of reactions that change state of system, starting from initial state  $X(0) = V_0$
- Stochastic simulation based on sampling of  $p[\tau, j/X(t) = V]$  generates sequences in exact accordance with stochastic master equations
- Repeating stochastic simulation many times (Monte-Carlo procedure) yields approximation of probability distribution  $p(X(t) = V)$

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

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

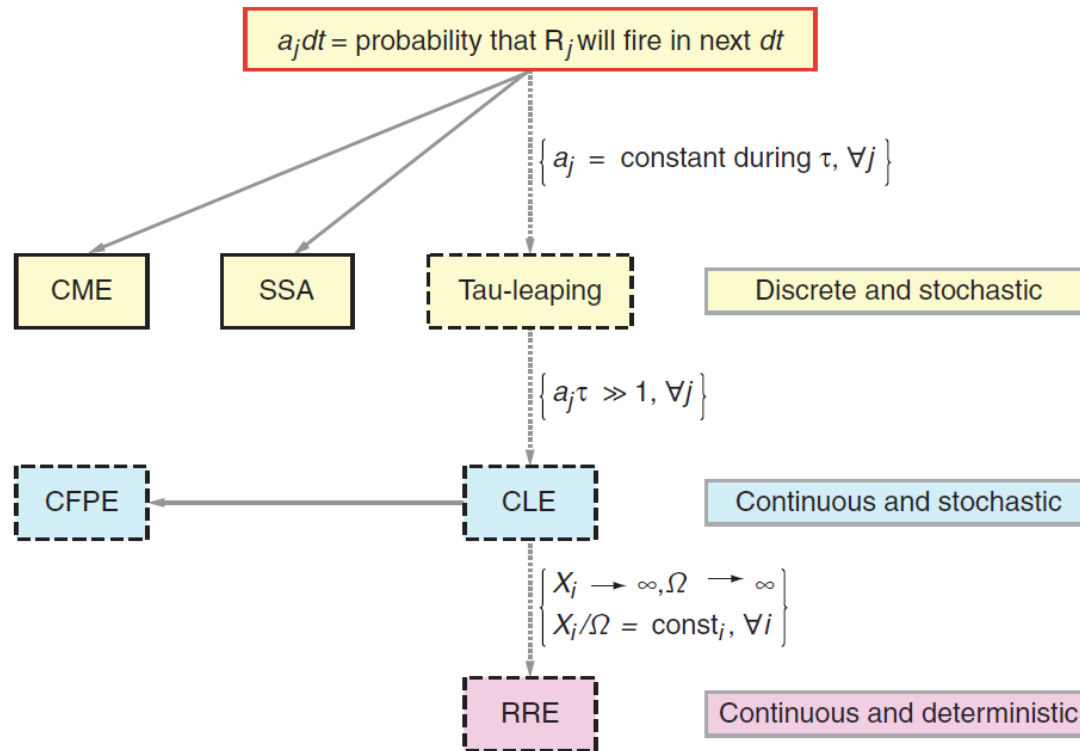
# Stochastic simulation

- Analytical solution of master equations is not possible in most situations of practical interest
- **Stochastic simulation** generates sequences of reactions that change state of system, starting from initial state  $X(0) = V_0$
- Various approximations of basic stochastic simulation algorithm, trading exactness for simulation speed:
  - Tau-leaping approaches: choose  $\tau$  such that  $\alpha_j, \beta_j$  remain approximately constant over time interval (encapsulate several reactions in one step)
  - Quasi-steady-state approximations (distinguish between slow and fast reactions)
  - ...

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

# Stochastic simulation

- Relation of stochastic simulation models with other modeling approaches

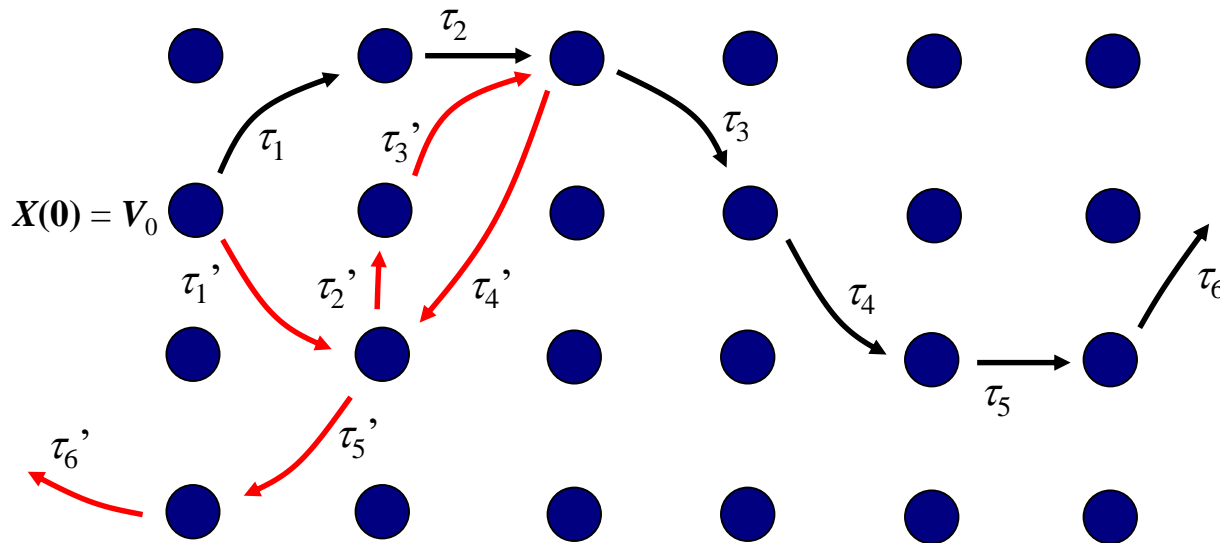


Gillespie (2007), *Annu. Rev. Phys. Chem.*, 58:35-55



# Stochastic simulation

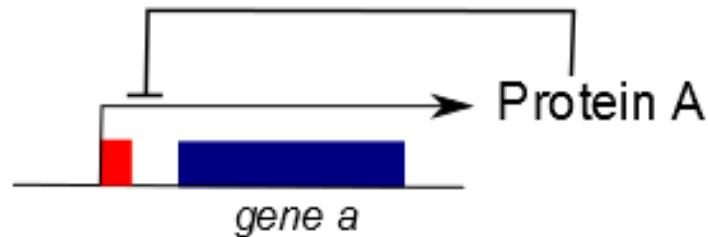
- **Stochastic simulation** generates sequences of reactions that change state of system, starting from initial state  $X(0) = V_0$



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

# Auto-inhibition network

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

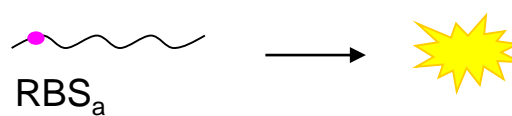
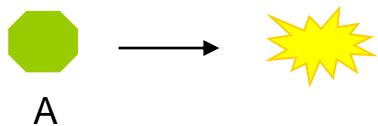
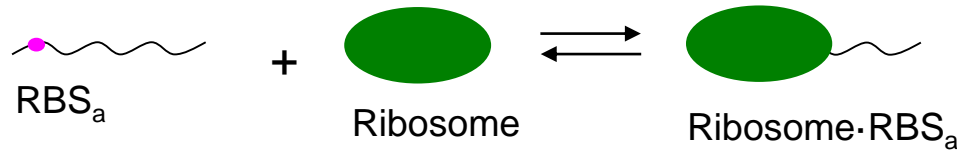
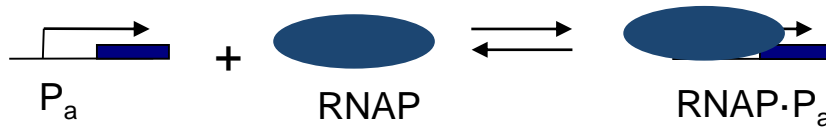


- Auto-inhibition is example of **negative feedback**, and frequently occurs in bacterial regulatory networks

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

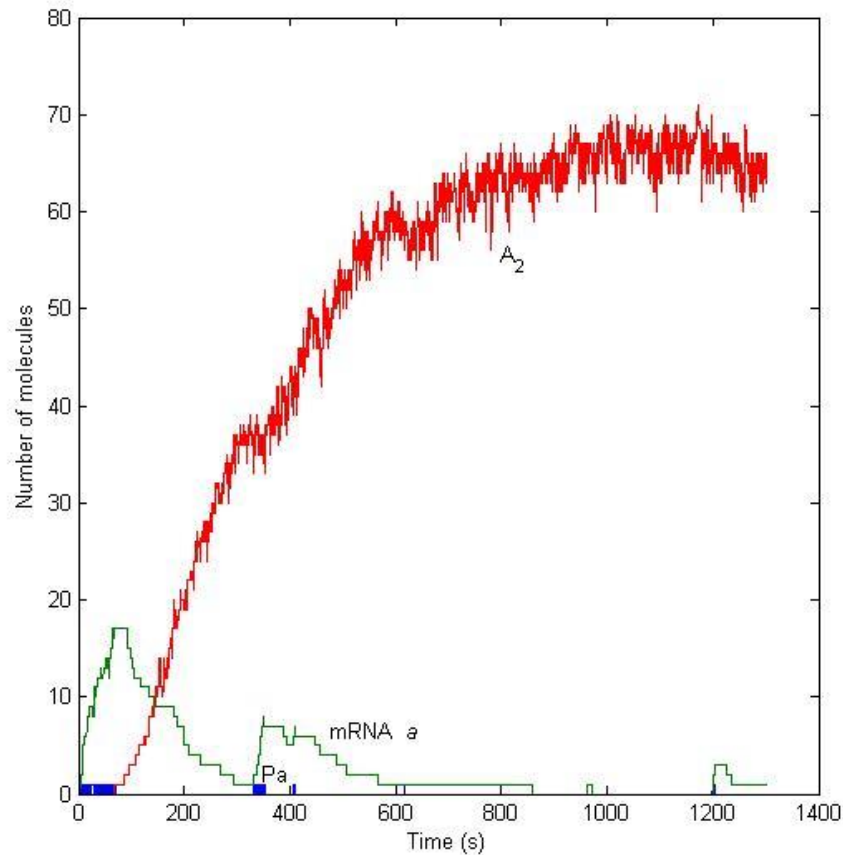
- Development of stochastic model requires list of species, reactions, and kinetic constants

# Reactions and species



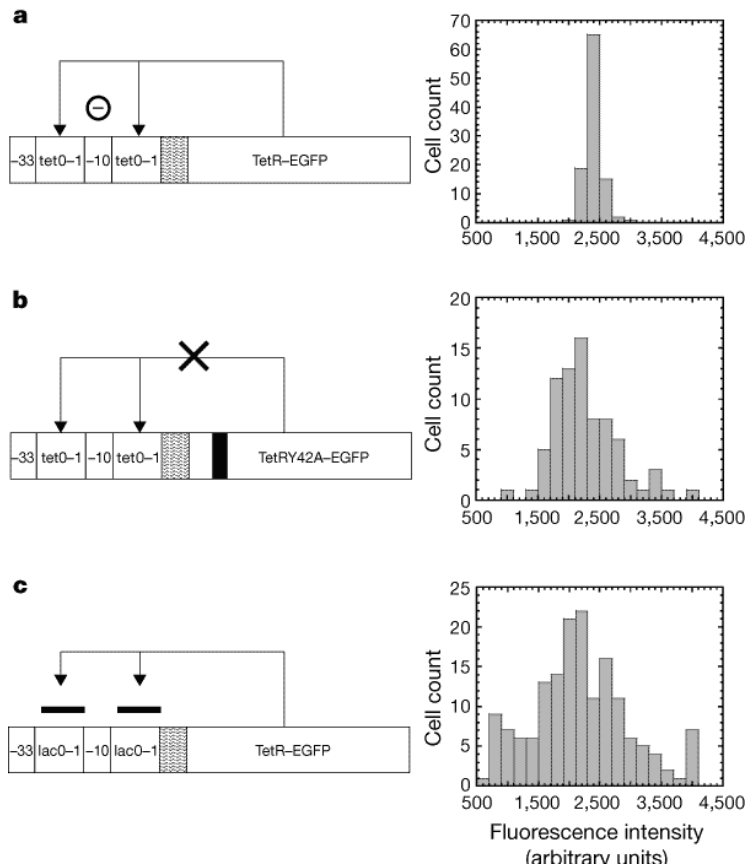
# Stochastic simulation of auto-inhibition

- Occurrence of fluctuations and bursts in gene expression



# Auto-inhibition and noise reduction

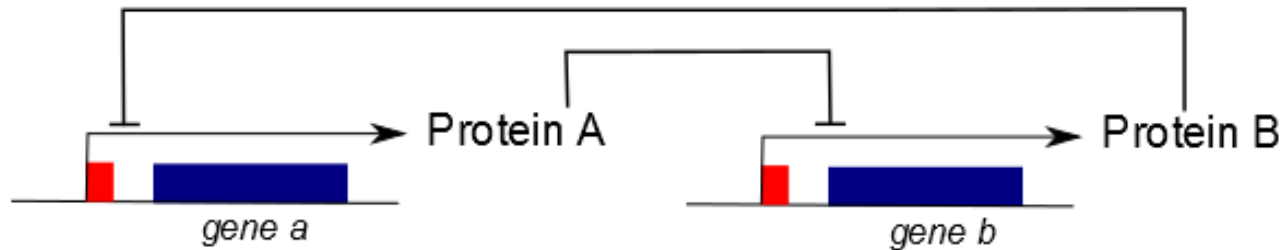
- Auto-inhibition reduces fluctuations in gene expression level



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

# 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*, CRC Press

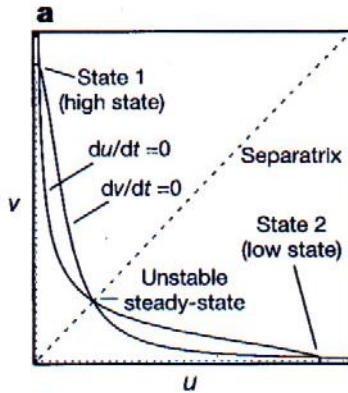
- Construction of cross inhibition network *in vivo*: **toggle switch**

Gardner *et al.* (2000), *Nature*, 403(6786): 339-42

# Dynamics of toggle switch

- ODE model predicts bistability of toggle switch

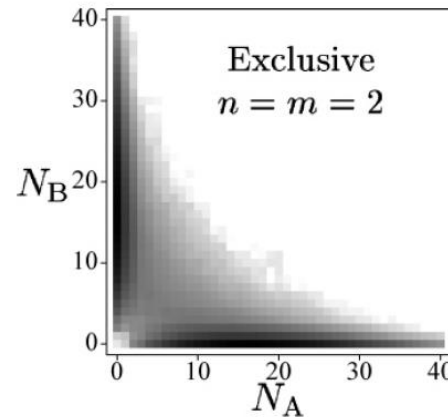
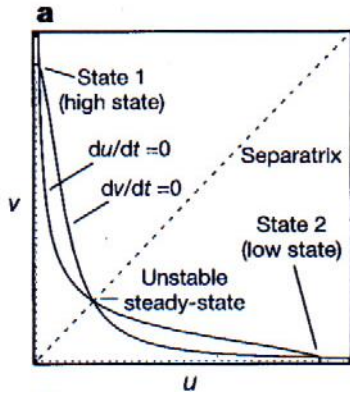
Gardner *et al.* (2000), *Nature*, 403(6786): 339-42



- Question: what will be predicted long-term dynamics in stochastic model of toggle switch?

# Dynamics of toggle switch

- ODE model predicts bistability of toggle switch



- Stochastic model predicts bimodal state (two attractors)

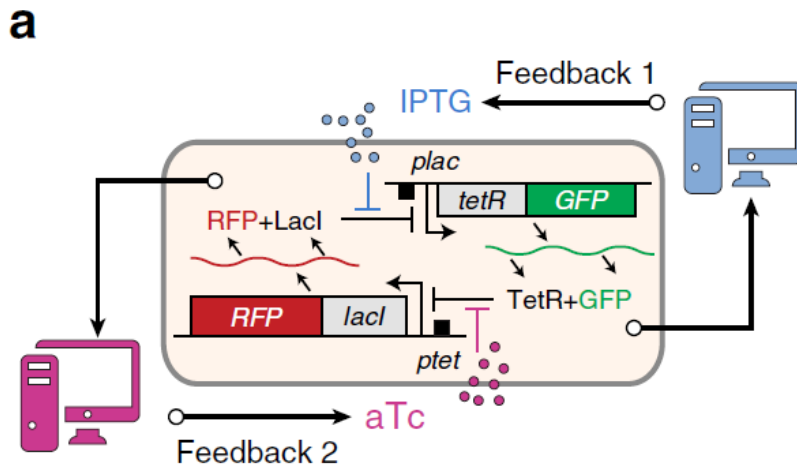
Warner and ter Wolde (2005), *J. Phys. Chem. B*, 109(4):6812-23

- Depending on noise characteristics, system can **spontaneously switch** from one attractor to another



# Control of toggle switch

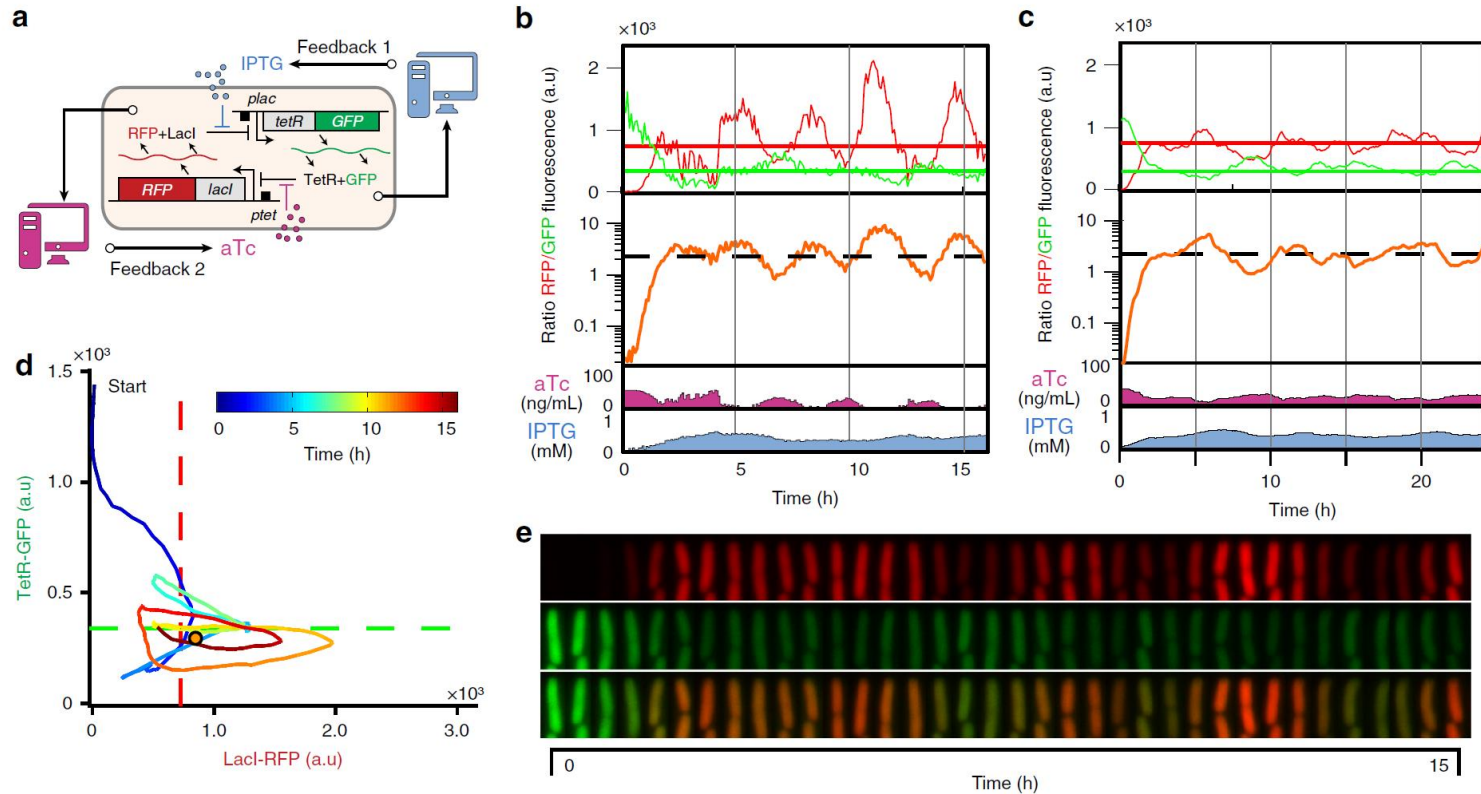
- Is it possible to stabilize toggle switch around unstable steady state in ODE model?



Lugagne *et al.* (2017), *Nat. Commun.*, 8:1671

# Control of toggle switch

- Is it possible to stabilize toggle switch around unstable steady state in ODE model?



Lugagne *et al.* (2017), *Nat. Commun.*, 8:1671

# Control of toggle switch

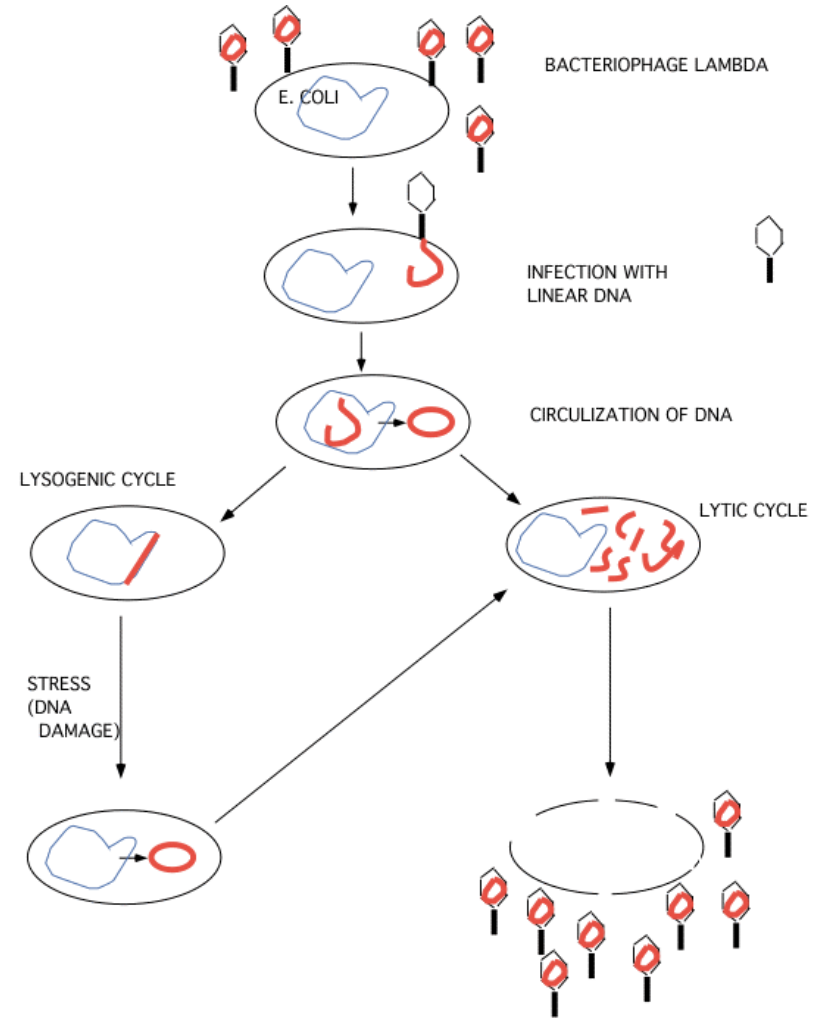
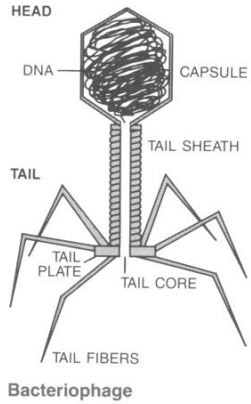
- Is it possible to stabilize toggle switch around unstable steady state in ODE model?
- Applications of control theory in synthetic biology: **cybergenetics**

<https://bsse.ethz.ch/ctsb/research/cybergenetics.html>

# Bacteriophage $\lambda$ infection of *E. coli*

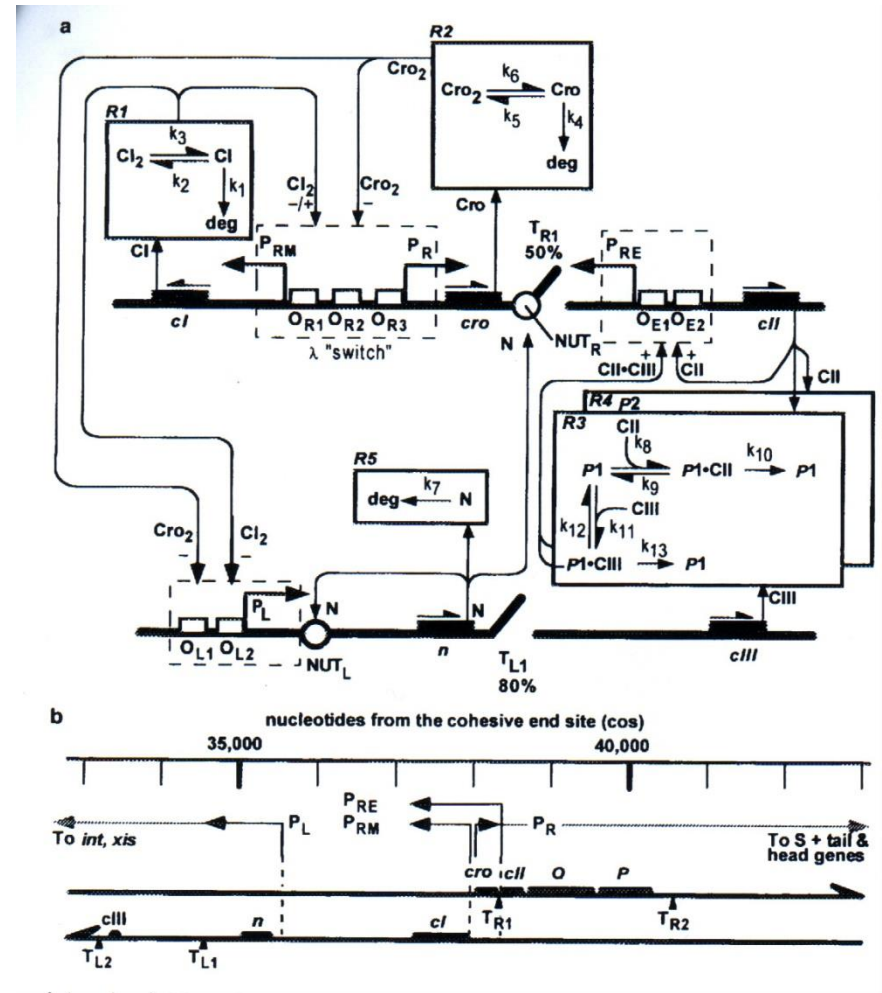
- Response of *E. coli* to phage  $\lambda$  infection involves decision between alternative developmental pathways: **lytic cycle** and **lysogeny**

Ptashne (1997), *A Genetic Switch: Phage  $\lambda$  and Higher Organisms*, Cell Press



# Stochastic analysis of phage $\lambda$ infection

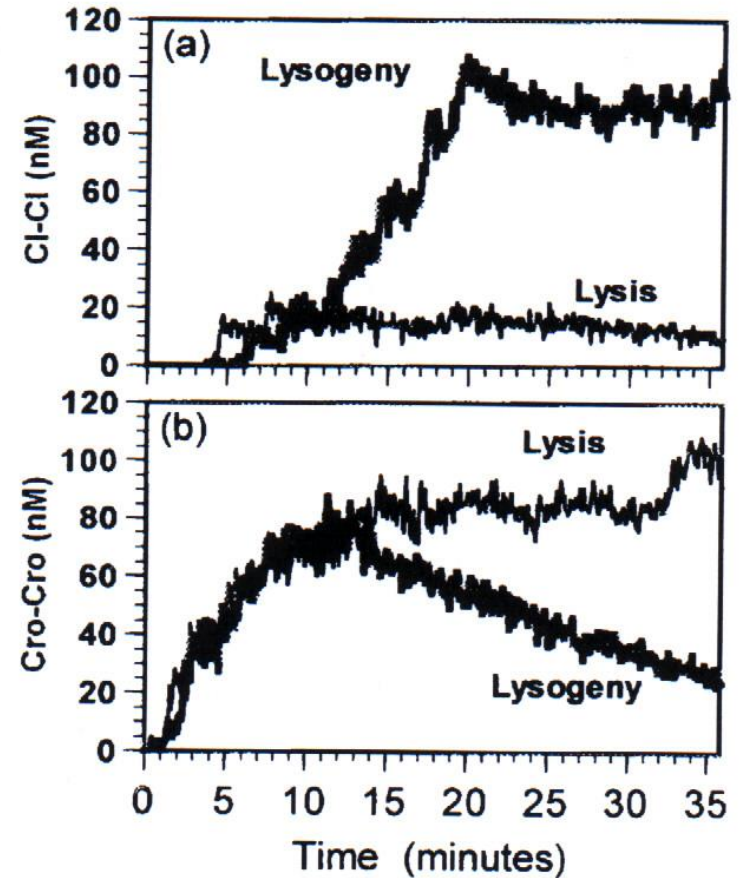
- Stochastic model of  $\lambda$  lysis-lysogeny decision network



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

# Stochastic analysis of phage $\lambda$ 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

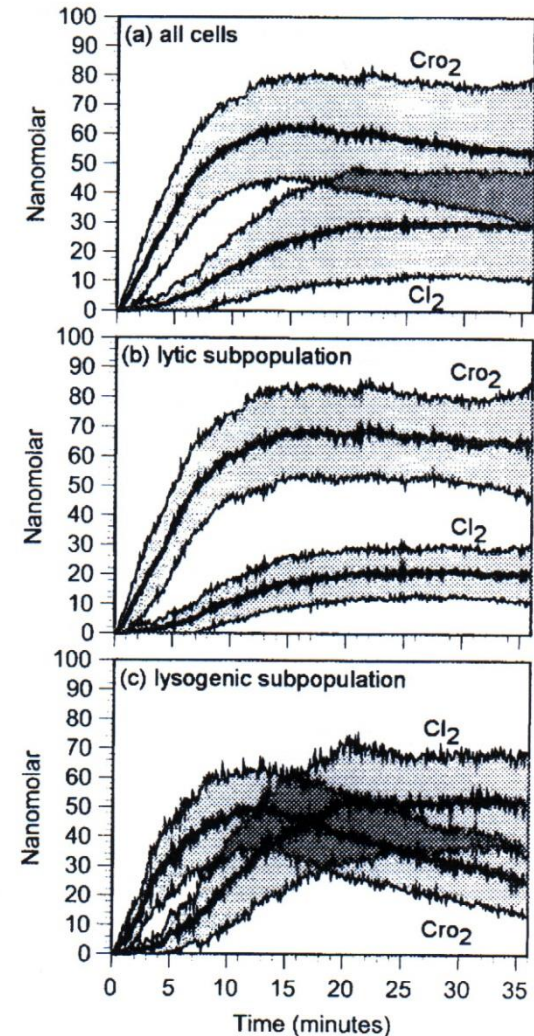
- Deterministic models can be seen as predicting **average behavior** of cell population

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

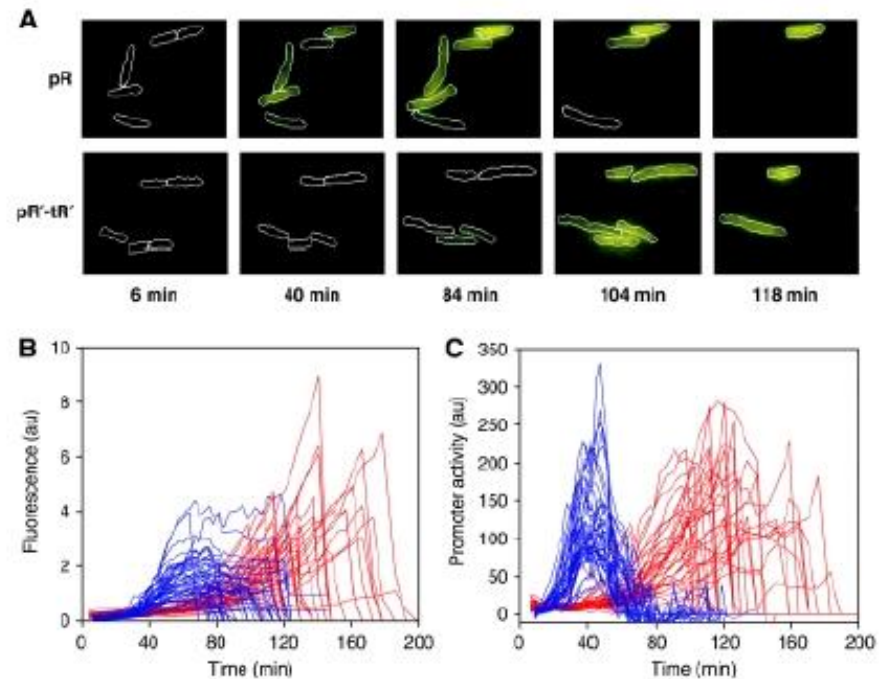
- However, under some conditions deterministic models yield good approximation



# Measurements of phage $\lambda$ 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



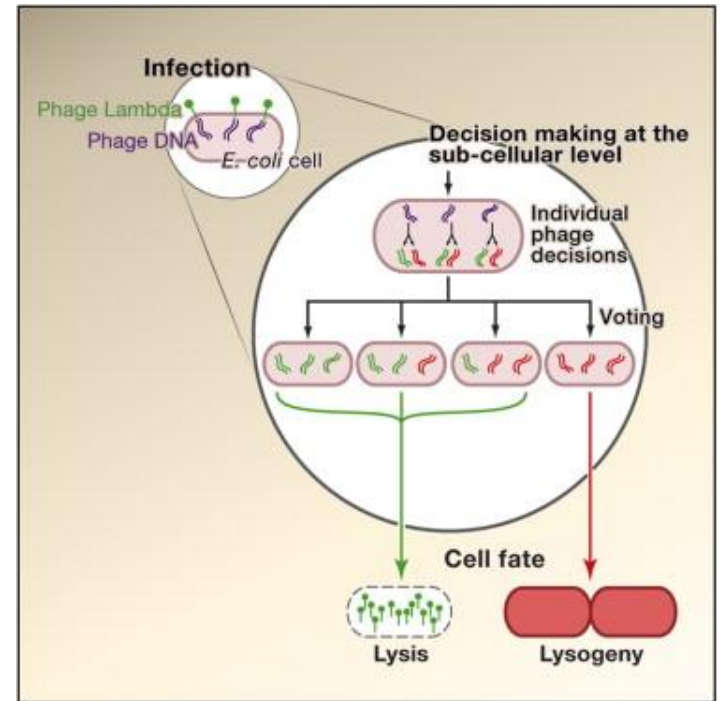
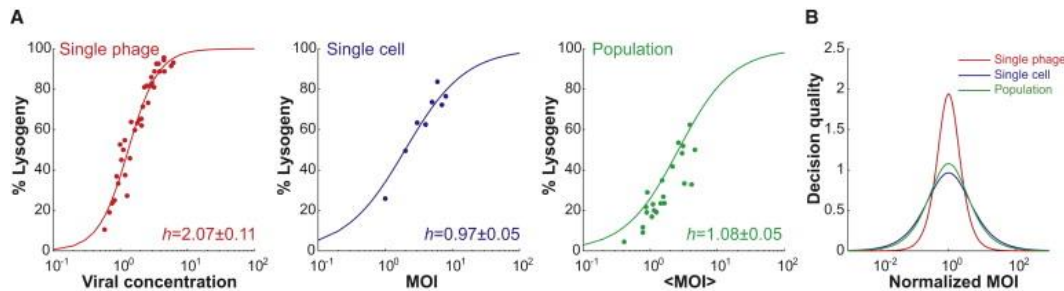
Amir *et al.* (2007), *Mol. Syst. Biol.*, 3:71



# 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

# Conclusions

- Stochastic models provide more realistic picture of gene expression
- Difficulty of stochastic models is that required information on regulatory mechanisms on molecular level usually not available

Reaction schemas and kinetic constants, necessary for generating values of parameters  $\tau$  and  $\rho$ , are not or incompletely known

- Another difficulty is that stochastic simulation is computationally expensive

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

**Merci !**

*inria*  
informatics mathematics

[www.inrialpes.fr/ibis](http://www.inrialpes.fr/ibis)