Introduction to Modeling of Genetic and Metabolic Networks

Hidde de Jong

INRIA Grenoble - Rhône-Alpes, 655 Avenue de l'Europe,
Montbonnot, 38334 St. Ismier Cedex, France
http://ibis.inrialpes.fr/

INSA 5BIM

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

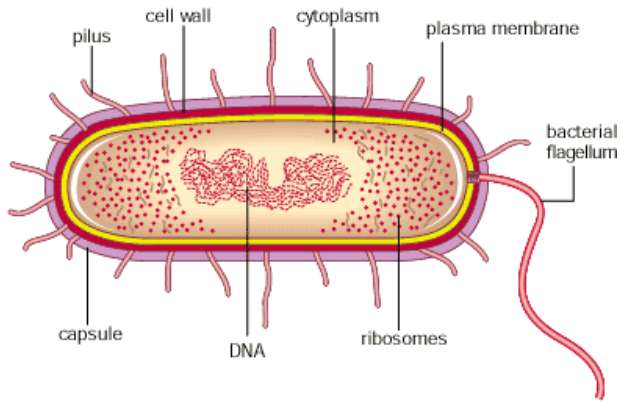
Outline

1. Biochemical reaction networks in bacteria
2. Mathematical modeling of biochemical reaction networks
3. Identification and validation using experimental data
4. Objectives and program of course

Bacteria

Bacteria form much of the world's biomass

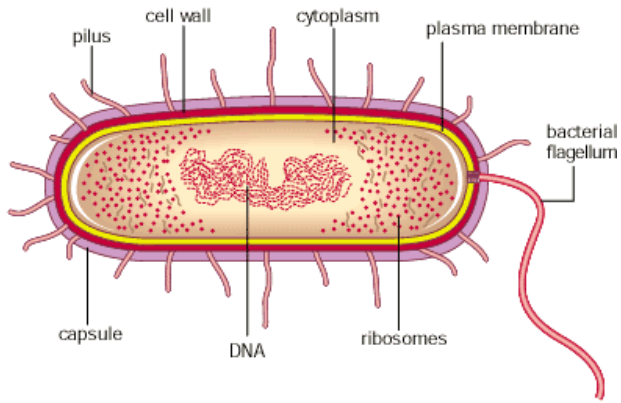
- $40 \cdot 10^6$ bacterial cells in 1 g of soil and 10^6 bacterial cells in 1 mL of fresh water
- 10 times as many bacterial cells as human cells in human body
- Wide range of shapes (spheres, rods, spirals, ...), typically 0.5 - 5.0 μm in length



Bacteria

Bacteria form much of the world's biomass

- $40 \cdot 10^6$ bacterial cells in 1 g of soil and 10^6 bacterial cells in 1 mL of fresh water
- 10 times as many bacterial cells as human cells in human body
- Wide range of shapes (spheres, rods, spirals, ...), typically 0.5 - 5.0 μm in length



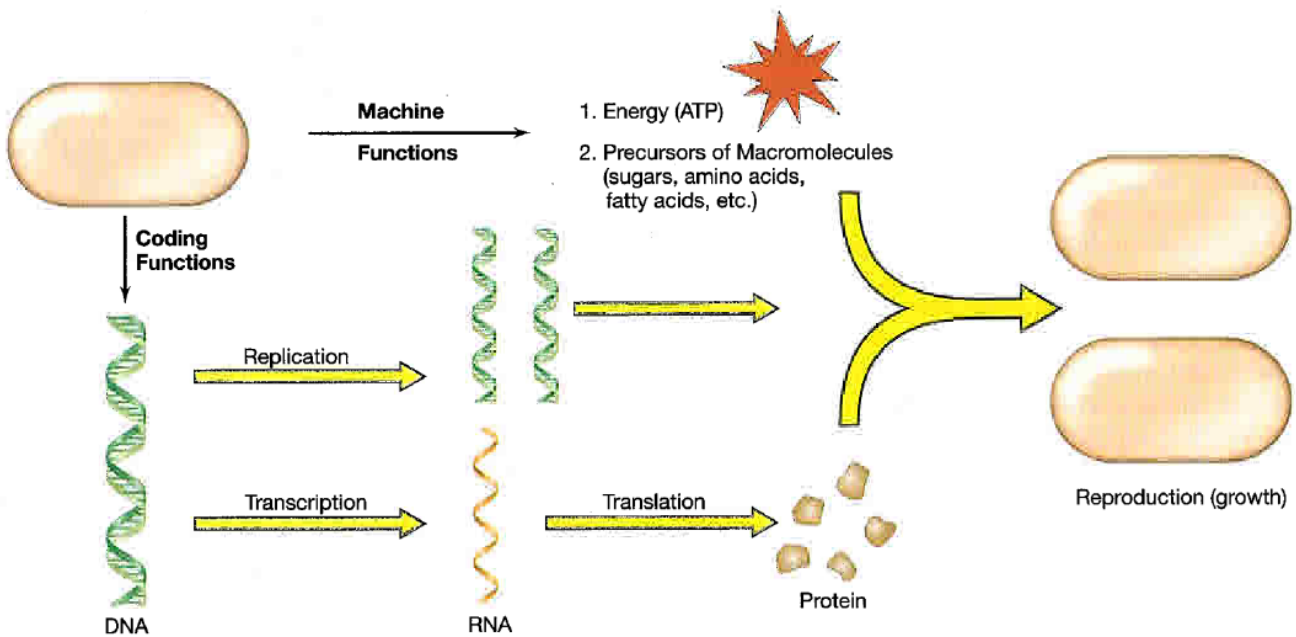
Impact of bacteria on humans

- Causative agents of **infective diseases**: cholera, syphilis, anthrax, leprosy, ...
- **Beneficial** bacteria: gut flora, probiotics, ...
- Bacteria in **technology and industry**: food industry, waste treatment, biotechnology, ...

Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation



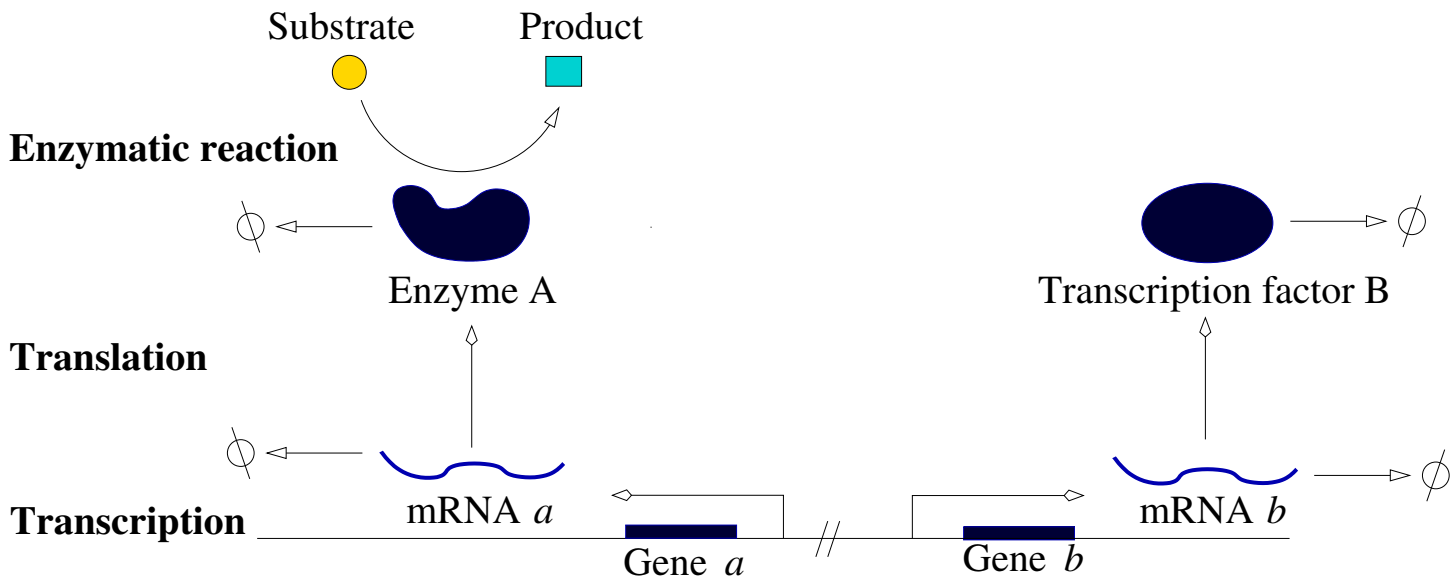
Madigan *et al.* (2003), *Brock Biology of Microorganisms*, Prentice Hall, 10th ed.

Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation

Major biochemical mechanisms of regulation

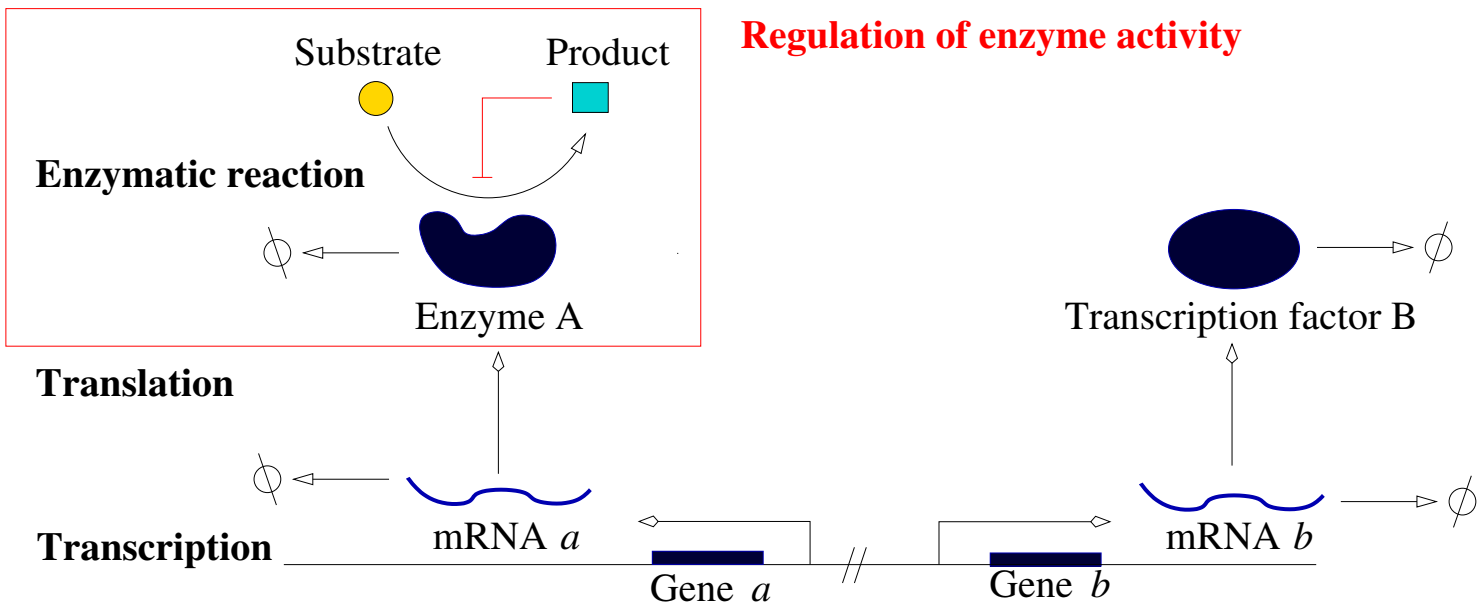


Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation

Major biochemical mechanisms of regulation

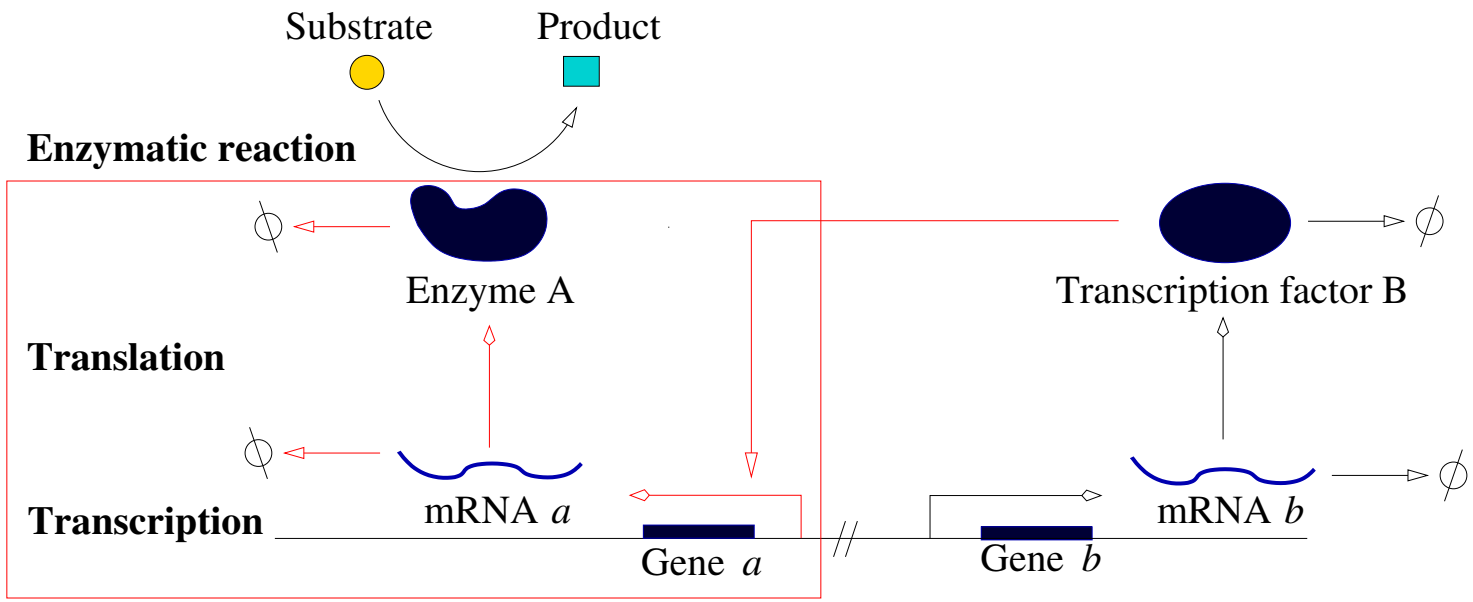


Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation

Major biochemical mechanisms of regulation



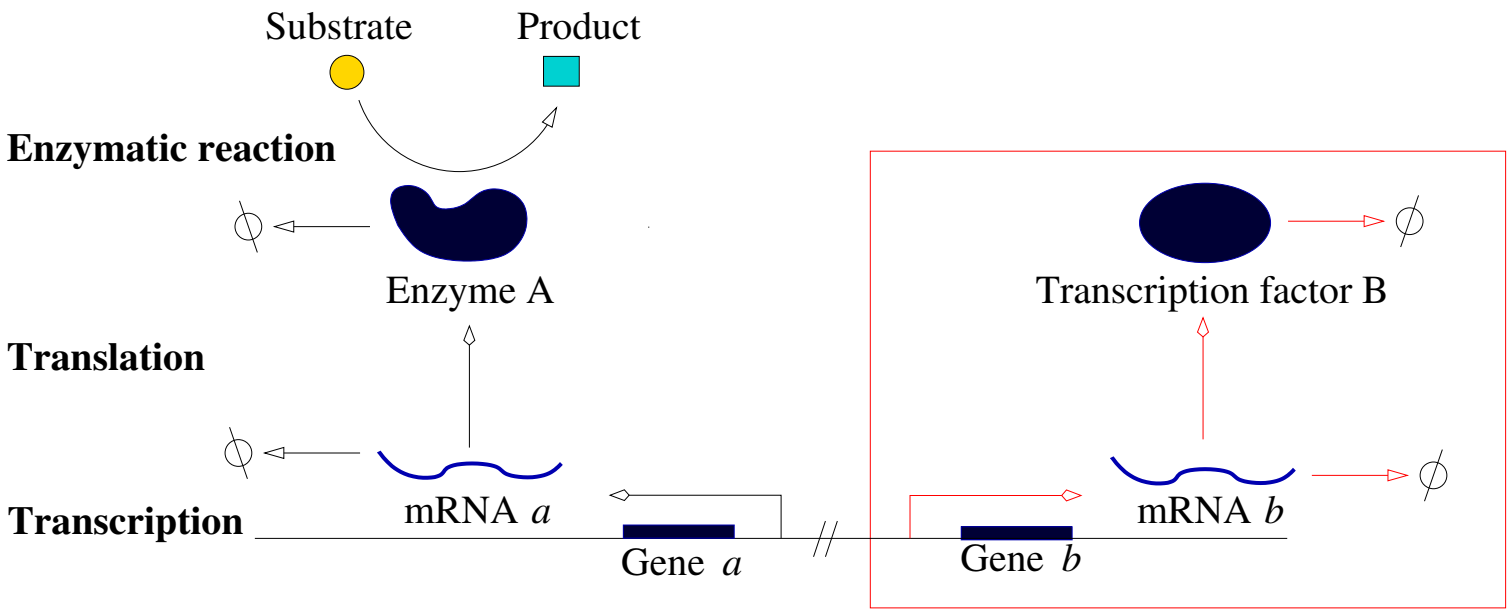
Regulation of enzyme synthesis/degradation

Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation

Major biochemical mechanisms of regulation



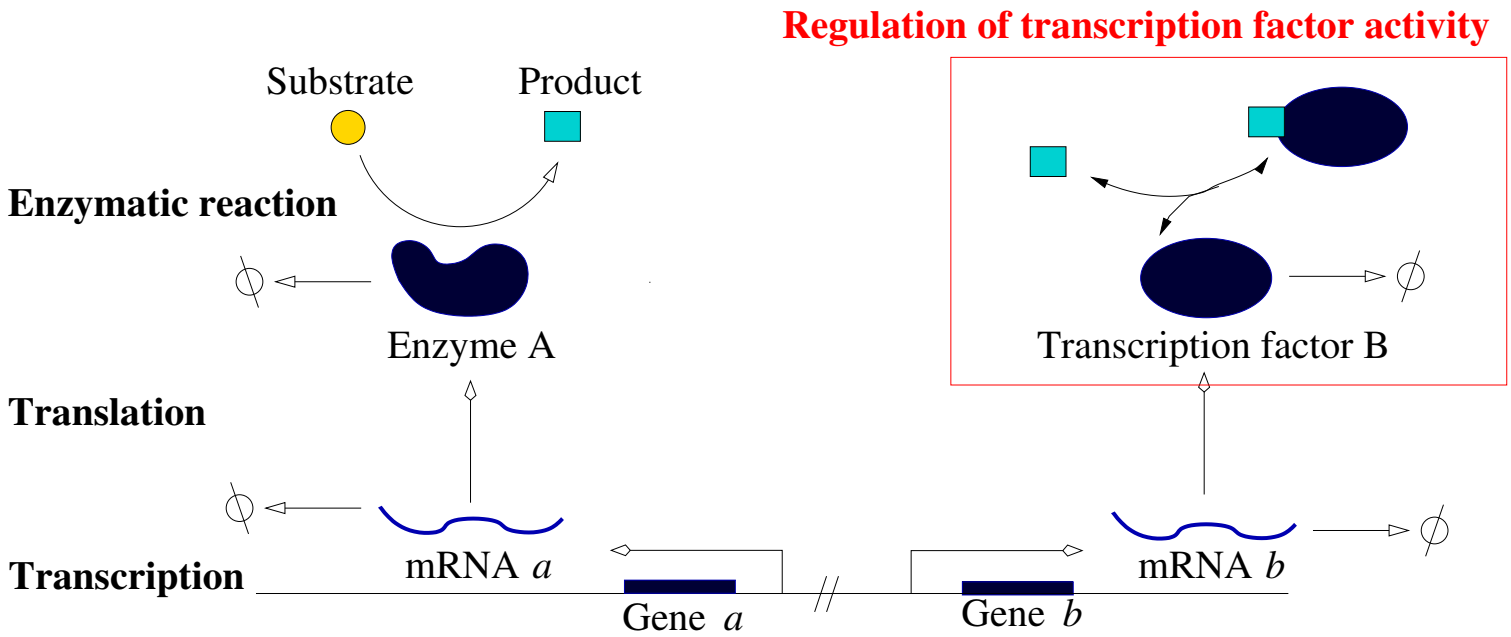
Regulation of transcription factor synthesis/degradation

Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation

Major biochemical mechanisms of regulation



Bacterial adaptation to a variety of environments

Adaptation is achieved by regulation of cellular functions

- **machine functions**: metabolism, involving production of energy and precursors of macromolecules
- **coding functions**: replication, transcription, translation

Major biochemical mechanisms of regulation

Biochemical species and reactions form **biochemical reaction networks**

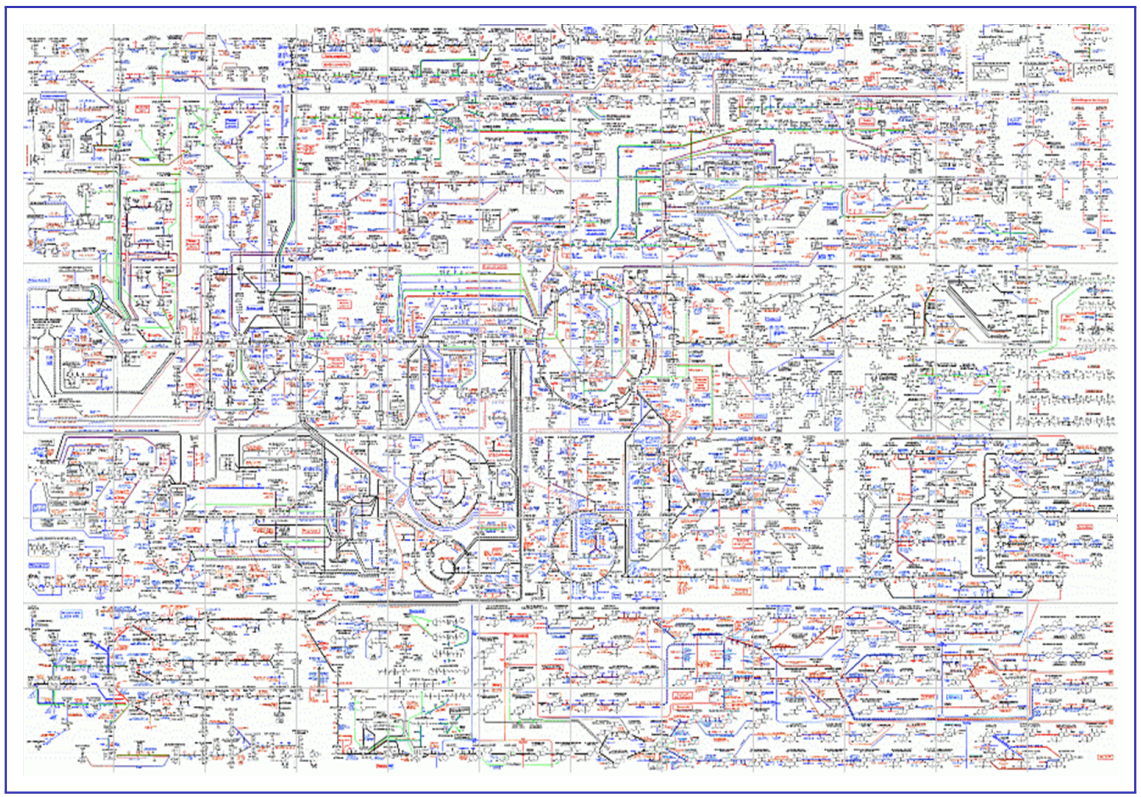
Complexity of biochemical reaction networks

Most networks of interest are large and complex

For instance, *E. coli* has:

- ∼4500 genes, with 330 genes coding for 170 transcription factors
- 194 metabolic pathways, involving 900 enzymes and ∼1000 biochemical reactions

Karp *et al.* (2007), *Nucleic Acids Res.*, 35(22): 7577-90

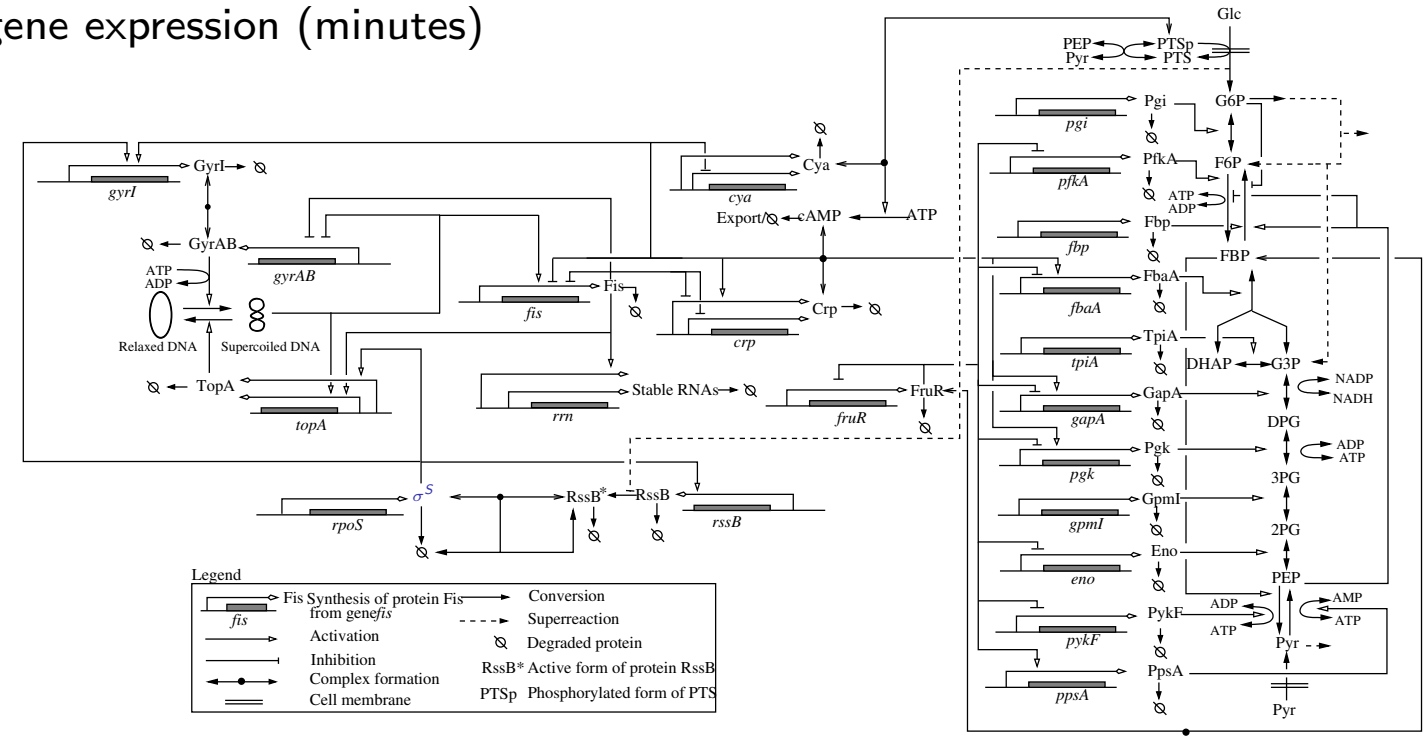


Heterogeneity of biochemical reaction networks

Most networks involve variety of biochemical reaction mechanisms, operating on different time-scales

For instance, *E. coli* carbon assimilation involves:

- signal transduction (milliseconds)
- enzymatic reactions (seconds)
- gene expression (minutes)

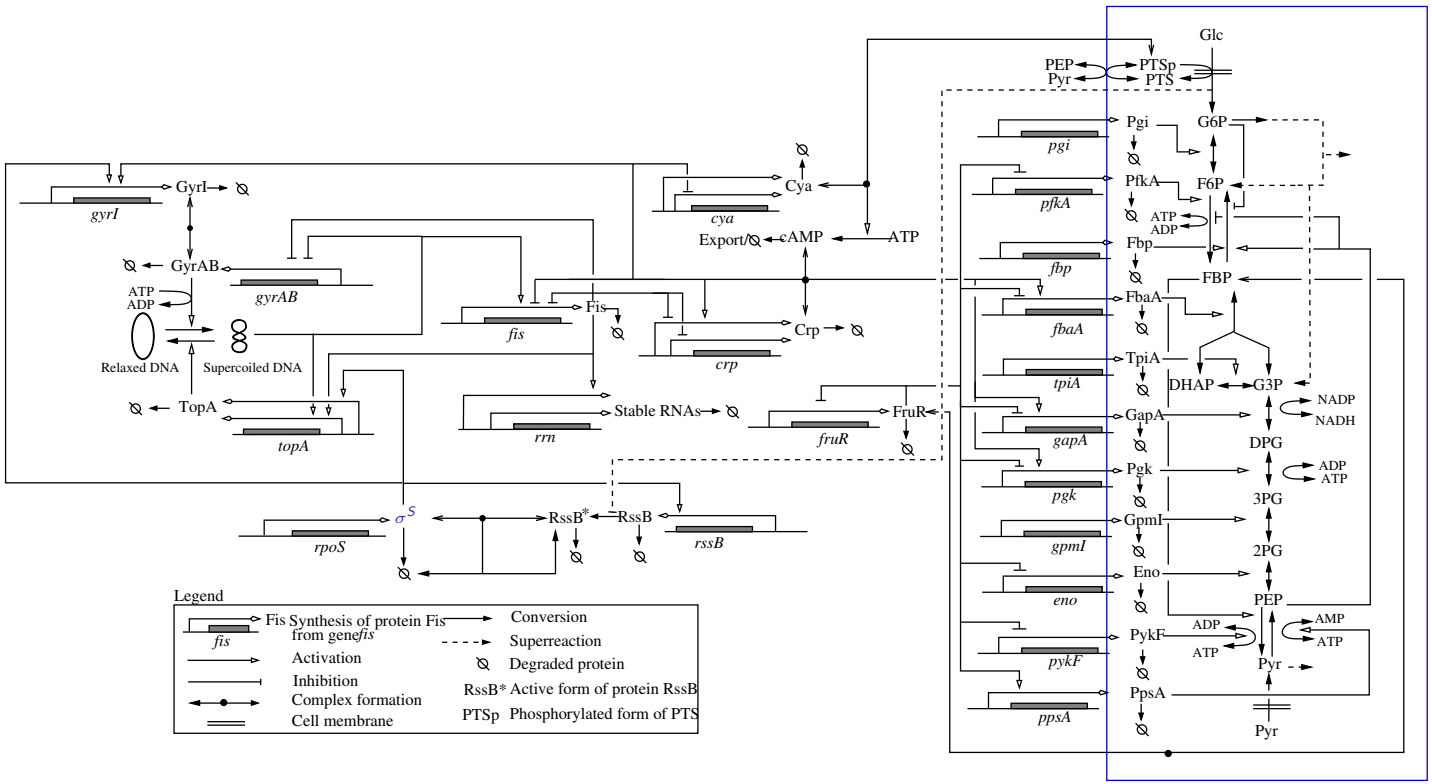


Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

Types of biochemical reaction networks:

Different types of networks distinguished by focusing on particular types of interactions and time-scales:

- **metabolic networks**: metabolites and enzymatic reactions

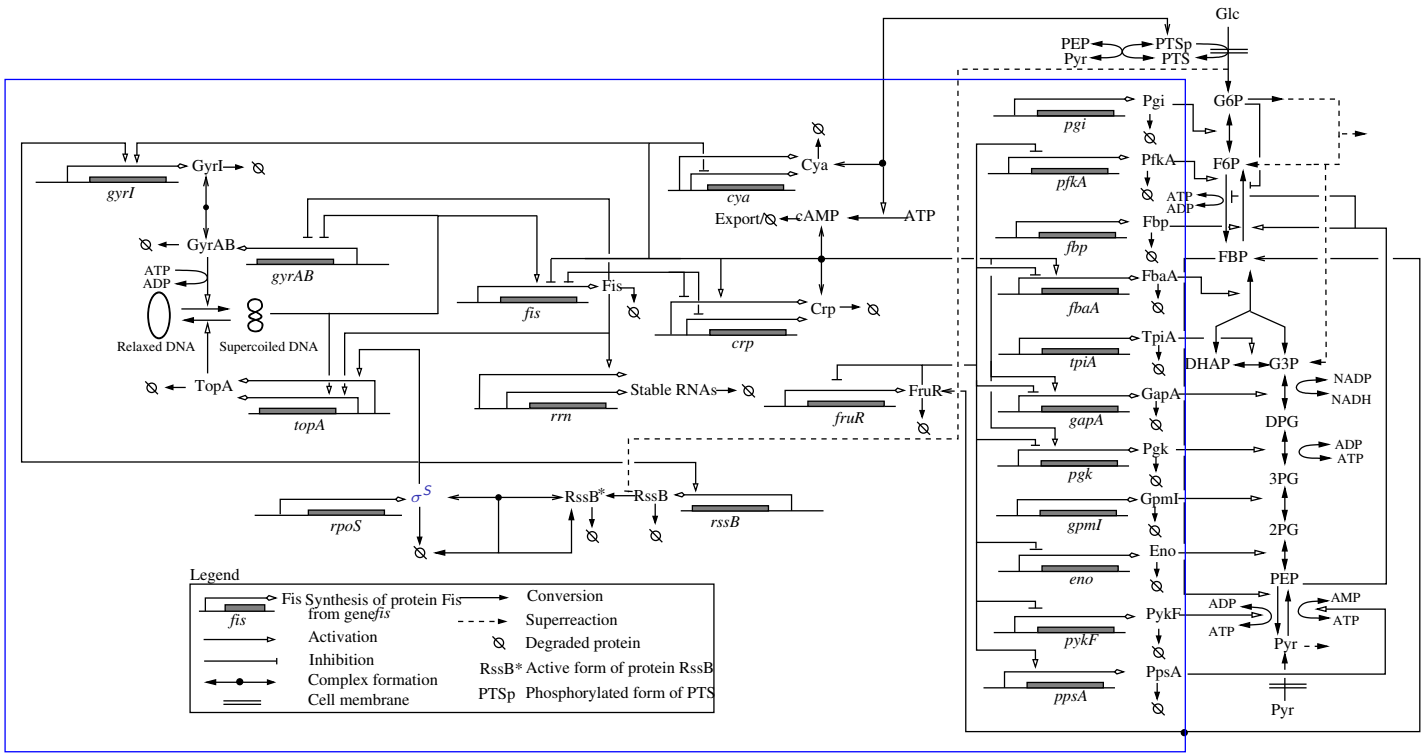


Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

Types of biochemical reaction networks:

Different types of networks distinguished by focusing on particular types of interactions and time-scales:

- **metabolic networks**: metabolites and enzymatic reactions
- **gene regulatory networks**: genes, RNAs, proteins, and direct and indirect regulation of gene expression



Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

Analysis of network functioning: from structure to dynamics

Wealth of knowledge on network structure in many bacteria

- Scientific databases and repositories
- Primary experimental literature

Comprehension of network functioning requires structure of network to be related to dynamics

Analysis of network functioning: from structure to dynamics

Wealth of knowledge on network structure in many bacteria

- Scientific databases and repositories
- Primary experimental literature

Comprehension of network functioning requires structure of network to be related to dynamics

Mathematical modeling and computer simulation indispensable for dynamic analysis of biochemical reaction networks

Analysis of network functioning has a central place in emerging field of **systems biology**

Kitano (2002), *Science*, 295(5560):564

Historical note

Systems biology, and more particularly the mathematical modeling and computer simulation of biochemical reaction networks, have a long history

Westerhoff and Palsson, *Nat. Biotechnol.*,22(10):1249-52

Simulation of metabolic pathways (glycolysis)

Garfinkel *et al.* (1970), *Ann. Rev. Biochem.*, 39:473-98

Historical note

Systems biology, and more particularly the mathematical modeling and computer simulation of biochemical reaction networks, have a long history

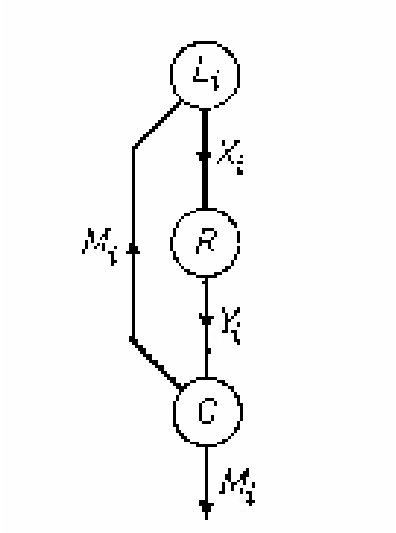
Westerhoff and Palsson, *Nat. Biotechnol.*,22(10):1249-52

Simulation of metabolic pathways (glycolysis)

Garfinkel *et al.* (1970), *Ann. Rev. Biochem.*, 39:473-98

Modeling of gene regulatory networks

Goodwin (1963), *Temporal Organization in Cells*, Academic Press



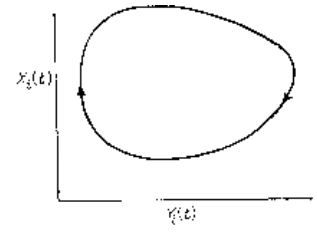
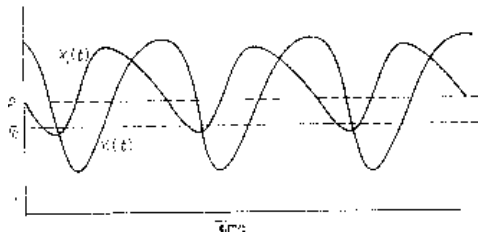


FIGURE 3.



Mathematical modeling of biochemical reaction networks

Well-established framework for modeling of biochemical reaction networks using ordinary differential equation (ODE) models

Heinrich and Schuster, *The Regulation of Cellular Systems*, Chapman & Hall, 1996

General form of ODE models of biochemical reaction networks

$$\dot{x} = N \cdot v(x)$$

- x: vector of concentrations of biochemical species
- N: stoichiometric matrix
- v: rate vector describing synthesis and degradation of proteins, metabolites, and biochemical complexes

ODE model of enzymatic reactions

- Long tradition in **kinetic theory**: precise description of catalytic mechanisms
  Segel (1993), *Enzyme kinetics*, Wiley & Sons
- Detailed description of enzyme kinetics (and complex formation): **mass-action law**

$$
\begin{aligned}
v_E &= (k_{-1} + k_2)\, x_{ES} - k_1\, x_E \\
v_{ES} &= k_1\, x_E - (k_{-1} + k_2)\, x_{ES} \\
v_S &= k_{-1}\, x_{ES} + k_1\, x_S \\
v_P &= k_2\, x_{ES}
\end{aligned}
$$

ODE model of enzymatic reactions

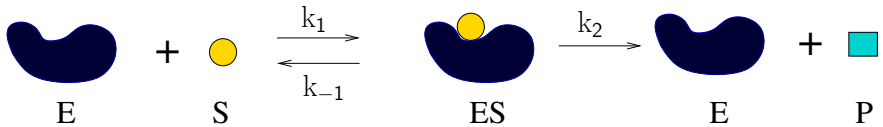
- Long tradition in **kinetic theory**: precise description of catalytic mechanisms
  Segel (1993), *Enzyme kinetics*, Wiley & Sons
- Detailed description of enzyme kinetics (and complex formation): **mass-action law**

$$v_E = (k_{-1} + k_2)\, x_{ES} - k_1\, x_E$$
$$v_{ES} = k_1\, x_E - (k_{-1} + k_2)\, x_{ES}$$
$$v_S = k_{-1}\, x_{ES} + k_1\, x_S$$
$$v_P = k_2\, x_{ES}$$

- **Michaelis-Menten kinetics**, based on two approximations
  - Quasi-steady-state approximation: the concentration of ES changes more slowly than those of S and P: $v_{ES} \simeq 0$
  - The total enzyme concentration (x_E^0) does not change over time: $x_E^0 = x_E + x_{ES} \simeq const.$

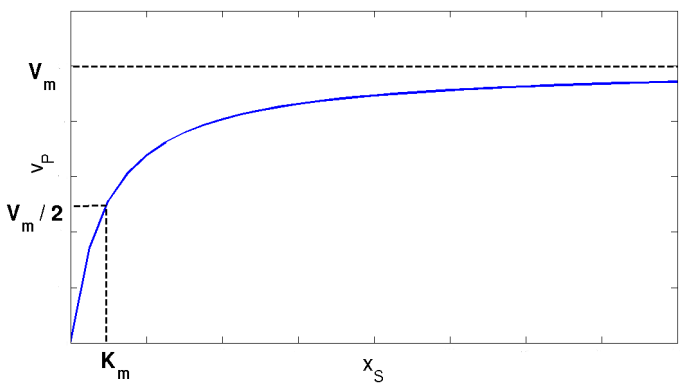
ODE model of enzymatic reactions

- Long tradition in **kinetic theory**: precise description of catalytic mechanisms
  Segel (1993), *Enzyme kinetics*, Wiley & Sons
- Detailed description of enzyme kinetics (and complex formation): **mass-action law**



$$v_E = (k_{-1} + k_2)\, x_{ES} - k_1\, x_E$$
$$v_{ES} = k_1\, x_E - (k_{-1} + k_2)\, x_{ES}$$
$$v_S = k_{-1}\, x_{ES} + k_1\, x_S$$
$$v_P = k_2\, x_{ES}$$

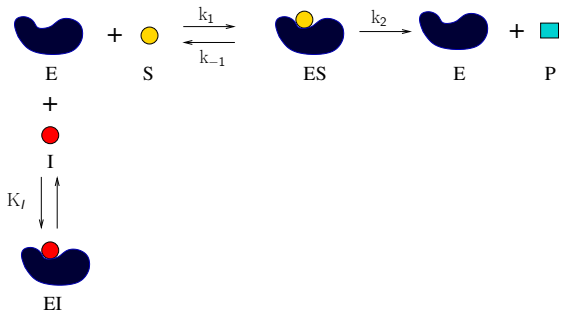
- **Michaelis-Menten kinetics**, based on two approximations



$$v_P = V_m \frac{x_S}{K_m + x_S} \text{ with:}$$
$$V_m = k_2\, x_E^0$$
$$K_m = \frac{k_{-1} + k_2}{k_1}$$

ODE models of enzymatic reactions, taking into account regulation of enzymatic activity

- **Michaelis-Menten kinetics with competitive inhibition**

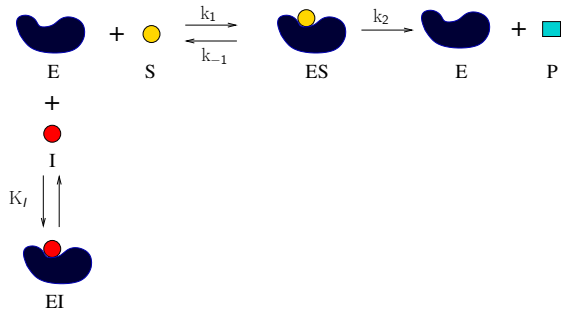


$$v_P = V_m \frac{x_S}{K_m\,(1 + \frac{x_I}{K_I}) + x_S}$$

Mathematical modeling of biochemical reaction networks

ODE models of enzymatic reactions, taking into account regulation of enzymatic activity

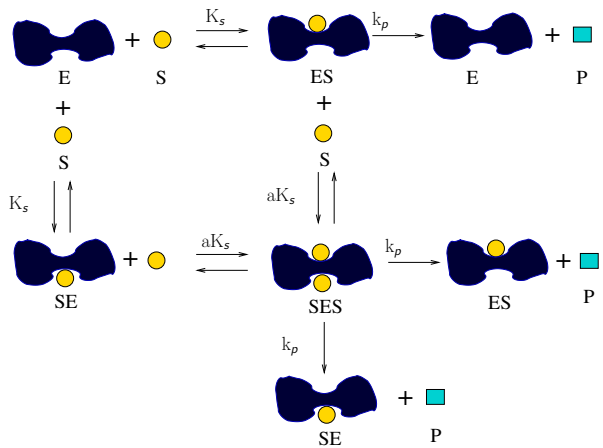
- **Michaelis-Menten kinetics with competitive inhibition**



$$v_P = V_m \frac{x_S}{K_m\,(1 + \frac{x_I}{K_I}) + x_S}$$

ODE models of enzymatic reactions with allostery

- **Cooperative binding: the sigmoidal Hill equation**



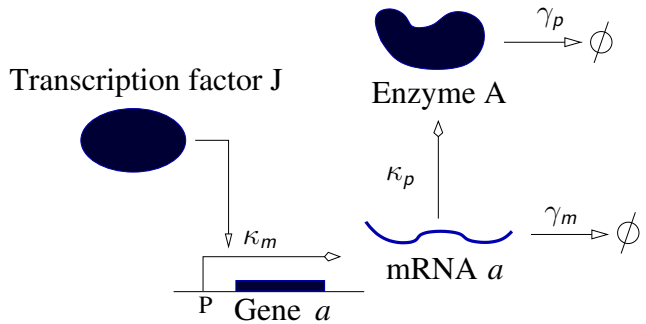
a: interaction factor

n: number of binding sites

$$v_P = V_m \frac{x_S^n}{a\,K_s^n + x_S^n}$$

ODE model of gene expression, taking into account regulation on transcriptional level

x_j: transcription factor concentration



$$v_{tr} = \kappa_m f(x_j)$$
$$v_{tl} = \kappa_p x_m$$
$$v_{dm} = \gamma_m x_m$$
$$v_{dp} = \gamma_p x_p$$

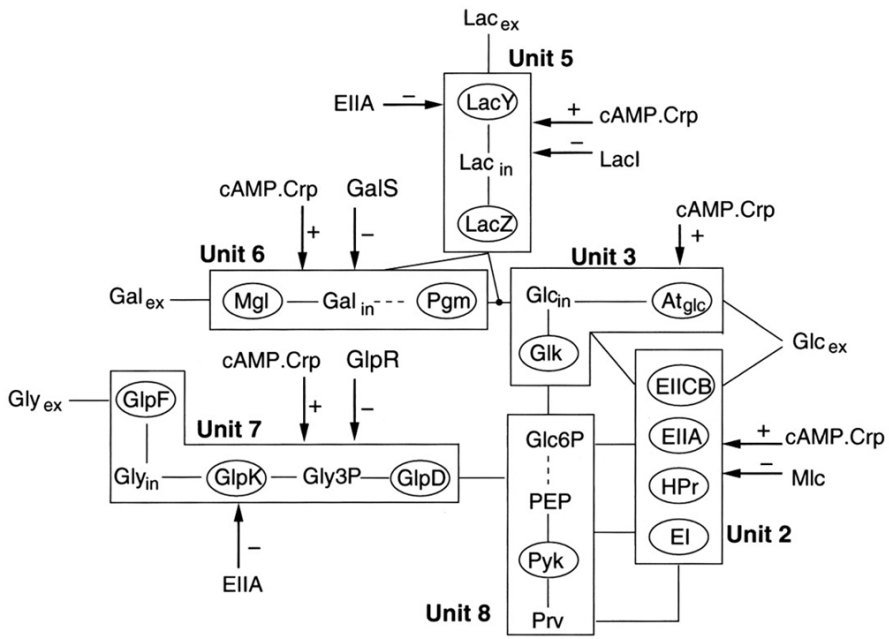
Regulation function $f(x_j)$ often has sigmoidal form, accounting for cooperative nature of regulation

A detailed model of carbohydrate metabolism in *E. coli*

Diauxic growth of the bacterium *Escherichia coli*

- When several carbon sources are available, *E. coli* bacteria choose the **nutrient sustaining fastest growth**, i.e. glucose is preferred over lactose
- Glucose depletion is followed by a **growth arrest**, when the bacteria modify their pattern of gene expression so as to produce the enzymes necessary for the uptake and metabolism of lactose.

Diauxic growth controlled by a complex biochemical regulatory network



Bettenbrock *et al.* (2006), *J. Biol. Chem.*, 281:2578-84

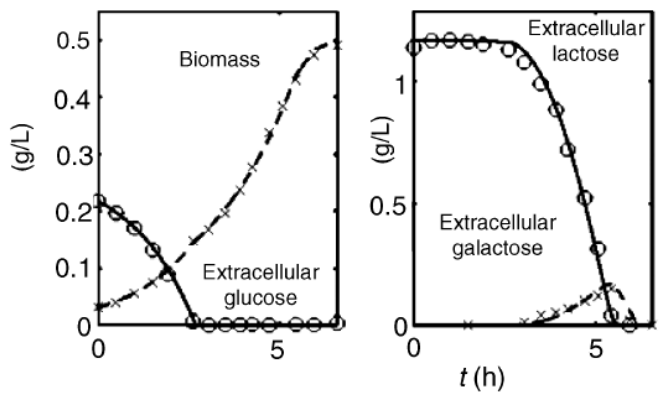
A detailed model of carbohydrate metabolism in *E. coli*

Detailed description of diauxic growth: kinetic model with 50 ODEs and 14 algebraic equations

- Parameter values reported in the literature could not all be included in the model: obtained in different experimental conditions and with different strains
- Bettenbrock *et al.* carried out their own experiments: **measurement of metabolite concentrations over time**
- **Parameter estimation** from experimental data

Confrontation of model predictions with experimental data

- Diauxic growth on glucose and lactose



Bettenbrock *et al.* (2006), *J. Biol. Chem.*, 281:2578-84

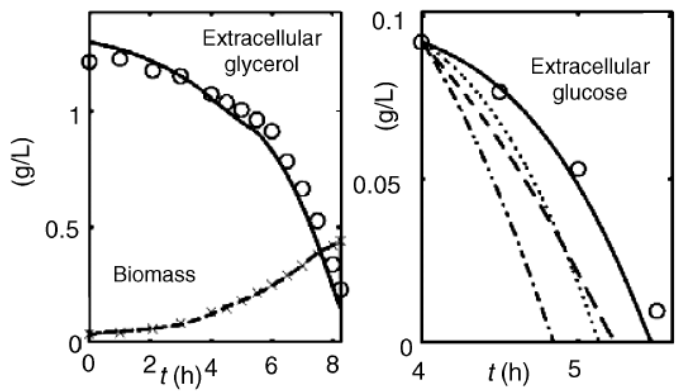
A detailed model of carbohydrate metabolism in *E. coli*

Detailed description of diauxic growth: kinetic model with 50 ODEs and 14 algebraic equations

- Parameter values reported in the literature could not all be included in the model: obtained in different experimental conditions and with different strains
- Bettenbrock *et al.* carried out their own experiments: **measurement of metabolite concentrations over time**
- **Parameter estimation** from experimental data

Confrontation of model predictions with experimental data

- Diauxic growth on glucose and lactose
- Disturbed batch experiment with application of a pulse of glucose on bacteria growing on glycerol



Bettenbrock *et al.* (2006), *J. Biol. Chem.*, 281:2578-84

Issues in mathematical modeling

Mathematical models are used for explanation, prediction, and control

Modeler confronted with several practical problems

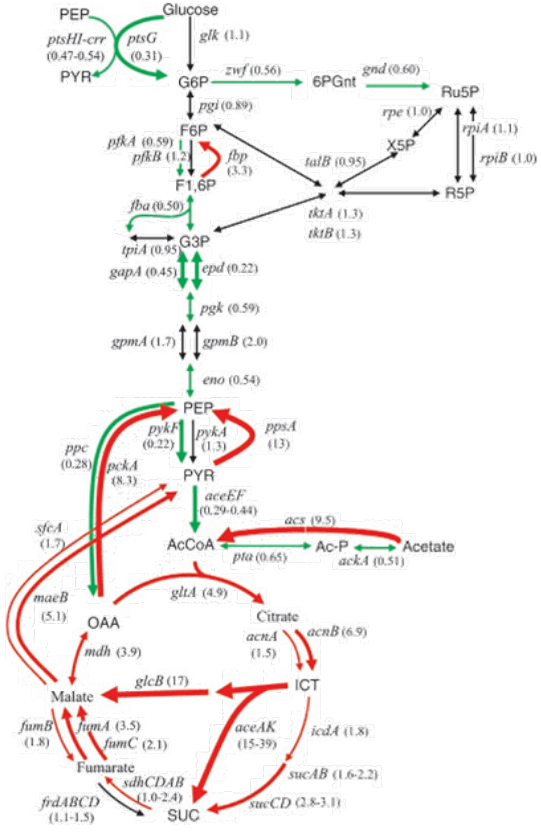
- Models of actual networks are large systems of nonlinear ODEs
- Parameter values are generally unknown and difficult to measure directly
- Reaction mechanisms often unknown
- Experimental measurements of variables are scarce, noisy, and indirect

This raises issues in model reduction and approximation, parameter estimation, network inference, data analysis, ...

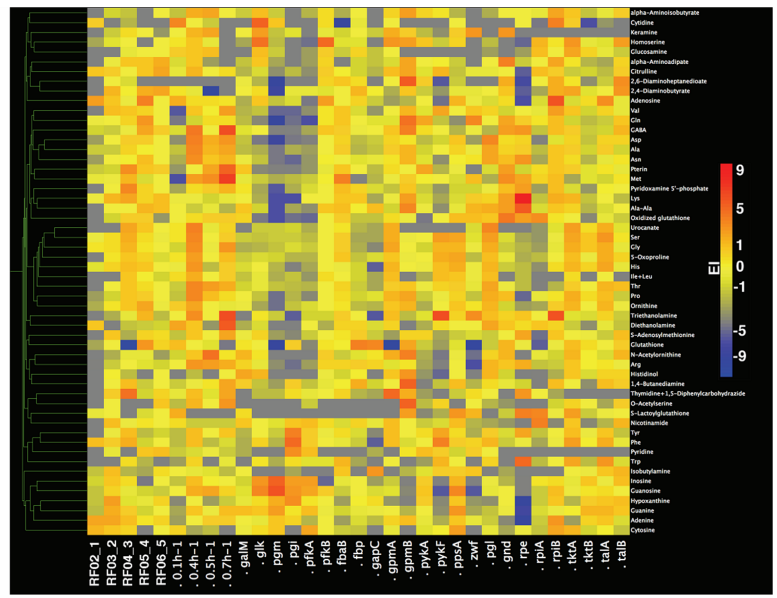
Experimental data

Availability and quality of experimental data is critical for model identification and validation

Measurements of fluxes and metabolite concentrations



Oh *et al.* (2002), *J. Biol. Chem.*, 277(15), 13175-83

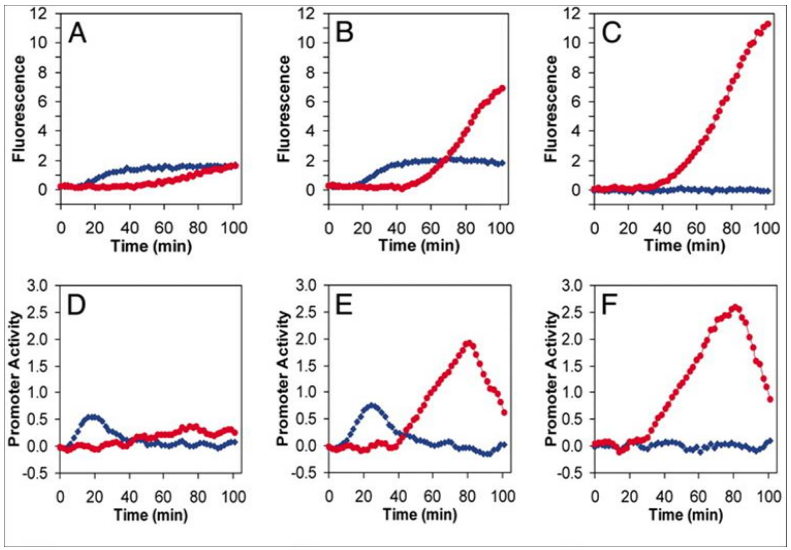


Ishii *et al.* (2007), *Science*, 316(5824):593-7

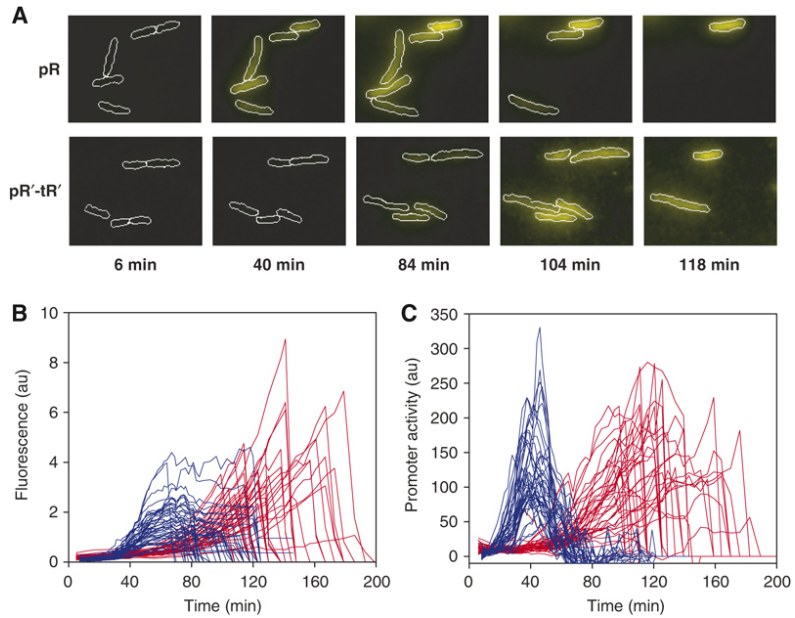
Experimental data

Availability and quality of experimental data is critical for model identification and validation

Reporter gene measurements to follow promoter activity over time, both in populations and individual cells



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



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

Objective of course "Modeling of biological networks"

The objective is to master kinetic modelling as applied to metabolic and gene regulatory networks

- Both the theoretical foundations and concrete applications to diverse systems of biological regulation
- Applications will rely on the practical use of computer tools for the modelling, analysis and simulation of biological networks

Program and teachers

Part 1. Systems biology and kinetic modeling (courses 10 h)

- Reminders on dynamical systems (Hidde de Jong)
- Introduction to regulatory systems (Hans Geiselmann)
- Reminders on kinetic modeling (Daniel Kahn)
- Reminders on enzymology (Daniel Kahn)

Part 2. Metabolic network modeling (courses 8 h, and practicals 9 h)

- Introduction to metabolomics (Daniel Kahn)
- Metabolic Control Theory (Daniel Kahn)
- Practical on the modeling of a metabolic system using COPASI (Daniel Kahn)

Program and teachers

Part 3. Gene regulatory network modeling (courses 12 h, and practicals 6 h)

- Introduction to recent techniques for measuring gene expression (Hidde de Jong)
- Kinetic models of gene expression and dynamics of gene regulatory networks (Hidde de Jong)
- Identification and inference of gene network models (Eugenio Cinquemani)
- Practical on the qualitative modeling of *E. coli* regulatory networks, using GNA (Hidde de Jong)

Part 4. Towards integrated models of regulatory networks (courses 2 h)

- MetaGenoReg project (Daniel Kahn and Hidde de Jong)

Evaluation

Metabolic network modeling:

- Exercises handed out during course

Gene regulatory network modeling:

- Synthesis of articles, guided by specific questions on articles handed out during course