Modeling of biochemical networks

David Swigon

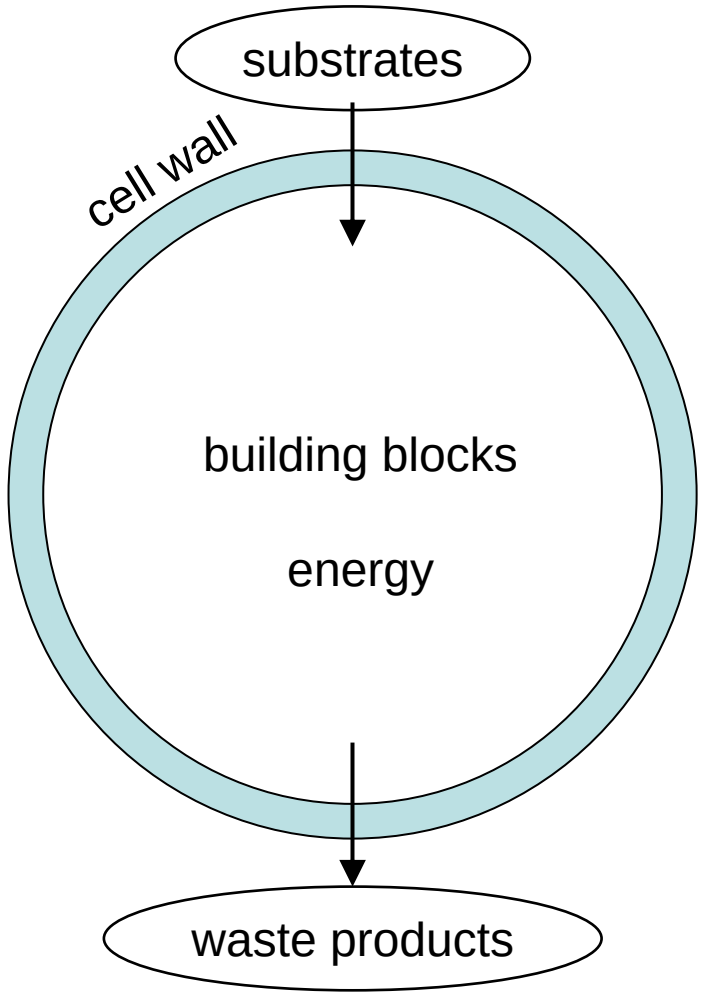
Cell

Cell = collection of coupled and tightly controlled biochemical reactions

Building blocks

- Nucleotides (RNA, DNA, ATP)
- Amino Acids (proteins)
- Carbohydrates (cytoskeleton, chitin, energy storage)
- Lipids (membranes, energy storage)

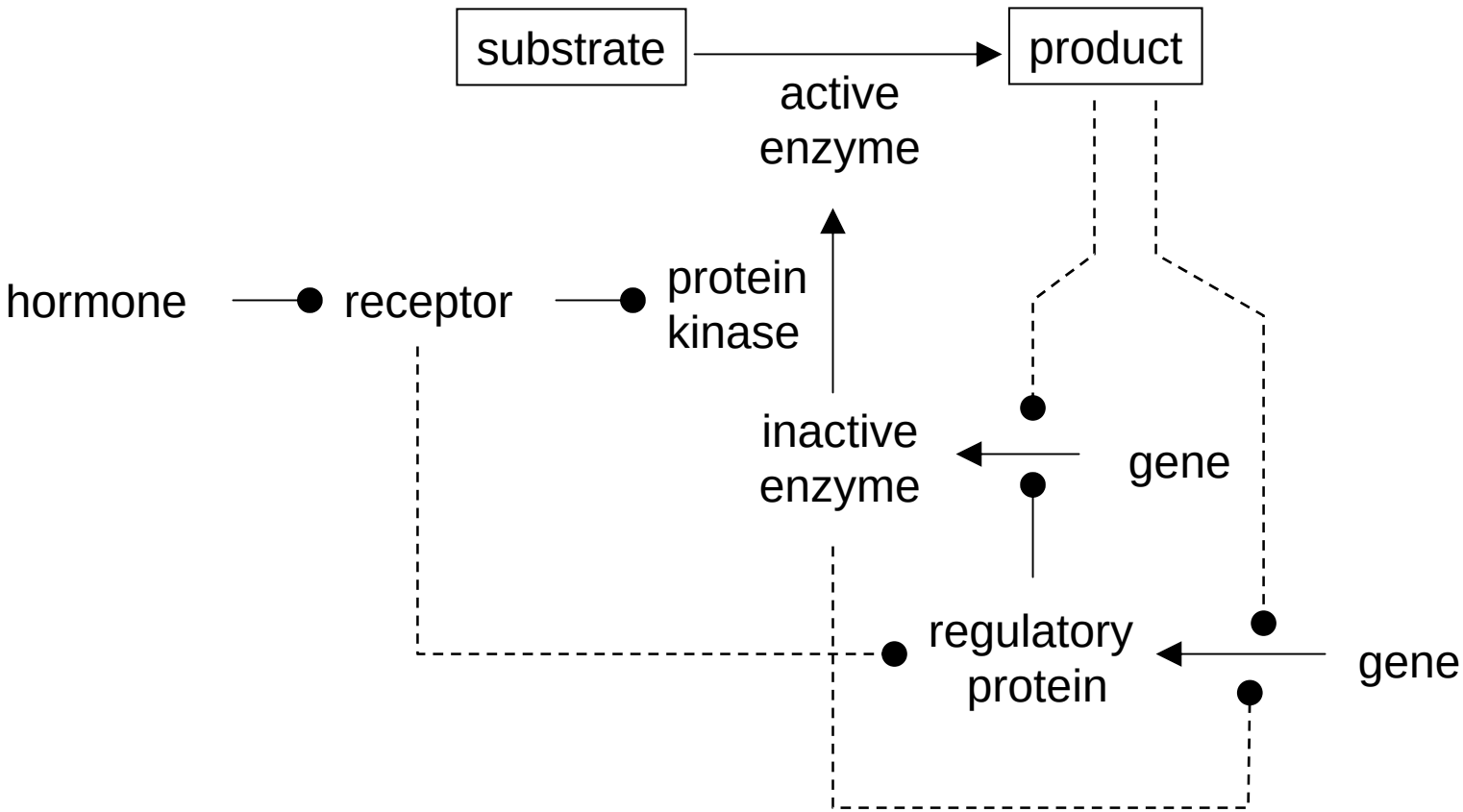
Reactions are catalyzed by enzymes (proteins) that are encoded in DNA.



Biochemical processes

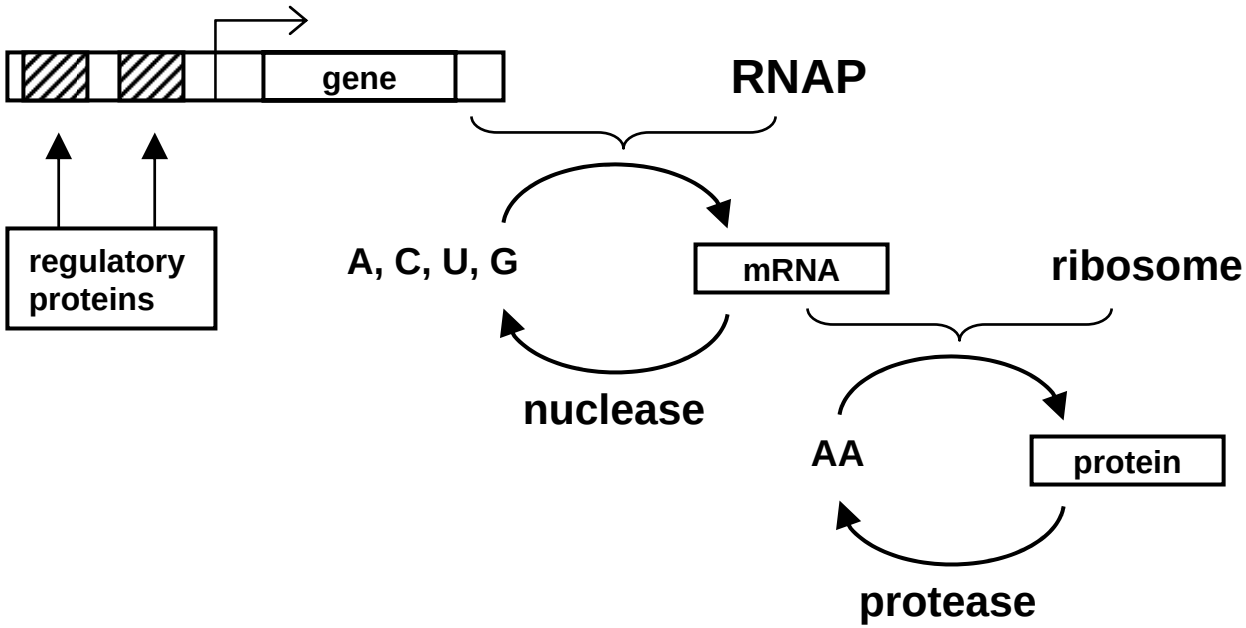
Processes	**Purpose**	**Reactions**	**Examples**
Metabolic processes	• Conversion of food into energy • production of building blocks	Enzymatic reactions	• Glycolysis • Synthesis of Lys, Met, Iso, Thr
Genetic processes	• Production of enzymes • Regulation of metabolic reactions in response to cell needs • Replication of DNA	Transcription Translation	• Lysis-lysogeny switch • Cell cycle
Signaling processes	• Regulation of genetic and metabolic processes in response to extracellular signals • Cell differentiation	Phosphorylation	• MAPK • A-kinase
Mechano-chemical	• Transport of molecules • Chemotaxis	Power stroke	• Kinesin motor • Flagella

Biochemical regulation

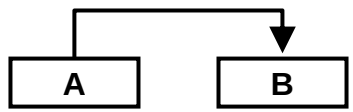


Genetic network

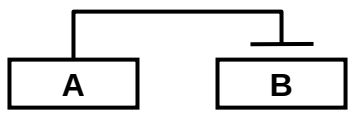
Gene expression



Activation (upregulation)



Repression (downregulation)



Regulatory motifs and modules

Switches
- logical control, computation, signal integration, memory

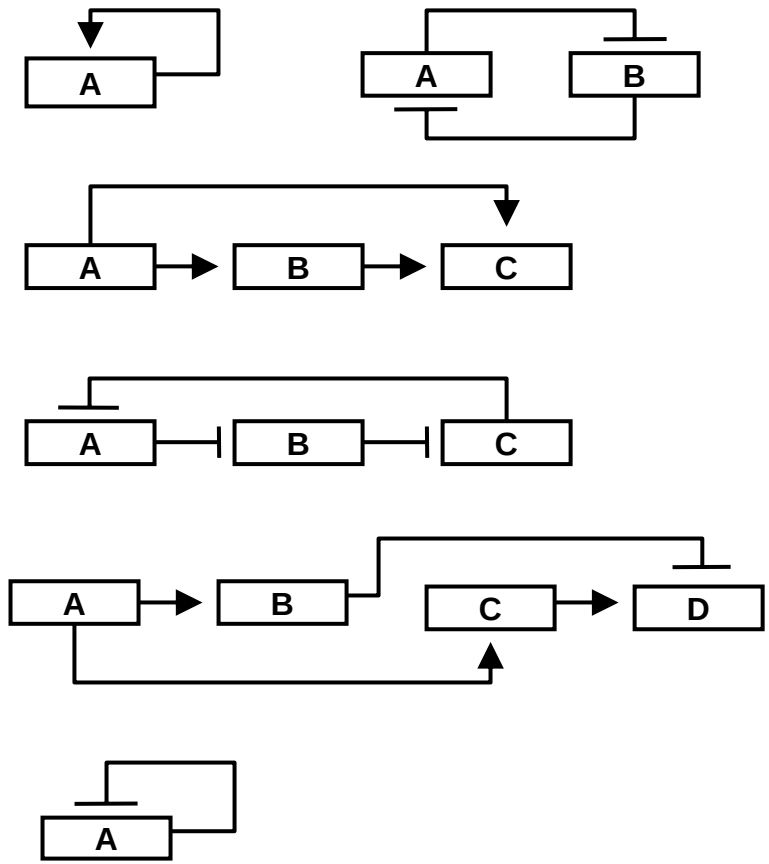
Control elements
- continuous adjustment: feedback and feed-forward loops

Oscillators
- periodic transitions between states: synchronization, carry signal

Amplitude filters
- amplify signals of intermediate strength: autoregulation

Noise filters or amplifiers
- precise regulation from noisy components, source of heterogeneity

Fundamental question – is decomposition possible ?



Modeling approaches

Metabolic networks

- steady state assumption, network flow analysis, extreme fluxes, perturbation analysis

Genetic & signaling networks

- graph theory and clustering techniques
- bayesian networks
- boolean networks
- ODEs
  - mass-action kinetics
  - piecewise linear kinetics
  - delay equations
- reaction-diffusion equations

Bottom up

- decomposition into chemical reactions
- application of mass-action or Michaelis-Menten kinetics

Top down

- captures essential features of incompletely known networks
- use of empirical laws (Hill equation, power law)

Common problem:
incomplete information about

- components
- interactions
- dynamical behavior

Graph theory and clustering

Genetic network = directed graph

Operations on graph

- Search for path connecting two genes – clues about missing regulatory interaction or redundancy
- Search for cycles – indication of feedback loops
- Global connectivity indicates complexity
- Loosely connected subgraphs indicate regulatory modules
- Comparison across taxa identifies evolutionary conservation of modules

Clustering techniques

- Used to group together genes with similar expression patterns: likely to regulate each other or be corregulated by the same factor
- Methods based on euclidean distance, linear correlation, etc.

Bayesian networks

Components

- Directed acyclic graph $G = <V, E>$
- Expression levels X_i ,
- Conditional probability $p(X_i \mid \text{parents}(X_i))$

$$p(\mathbf{X}) = \prod_{i \in V} p(X_i \mid \text{parents}(X_i))$$

For a given a set of expression data **X** learning techniques are used to find networks (G, p) that provide the best match

- NP-hard problem
- Heuristic algorithms are available

Boolean networks

Assumptions

- Gene is active or inactive, $x_i = \{0, 1\}$, $i = 1,\ldots,N$
- Boolean function b_i determines change in time
- Deterministic transitions, synchronous update

$$x_i(t+1) = b_i(\boldsymbol{x}(t))$$

Properties

- State-space has 2^N elements
- Cycles or point attractors

Logical networks

- Generalization to multiple values and asynchronous updates

Nonlinear ODEs

Continuum description $x_i = x_i(t)$

$$\frac{dx_i}{dt} = f_i(\boldsymbol{x})$$

Rate functions f_i

- Mass-action chemical kinetics $f_i(\boldsymbol{x}) = P(x_1, x_2, \ldots)$
- Michaelis-Menten, Hill, Shea & Ackers kinetics $f_i(\boldsymbol{x}) = \frac{P(x_1, x_2, \ldots)}{Q(x_1, x_2, \ldots)}$
- Piecewise linear $f_i(\boldsymbol{x}) = \sum_j \kappa_{ij} b_{ij}(\boldsymbol{x}) - \gamma_i x_i$

Key problem – lack of knowledge about kinetic constants

Chemical kinetics

$mA + nB \rightarrow C$ $\quad$ m, n – stoichiometric coefficients

Continuum description $\quad$ $a = a(t)$, etc.

mass-action kinetics

$$\dot{a} = -mka^m b^n$$
$$\dot{b} = -nka^m b^n$$
$$\dot{c} = ka^m b^n$$

Results

Horn & Jackson 1972 – conditions for existence of chemical equilibrium

Feinberg 1995 – deficiency zero theorem: Lyapunov stability, global equilibrium
– deficiency one algorithm for multistationarity

Clarke 1980 – necessary conditions for stability, Hopf bifurcation

Willamowski & Rossler 1980 – chaos from 5 bimolecular reactions

Magnasco 1997 – capacity for universal computation

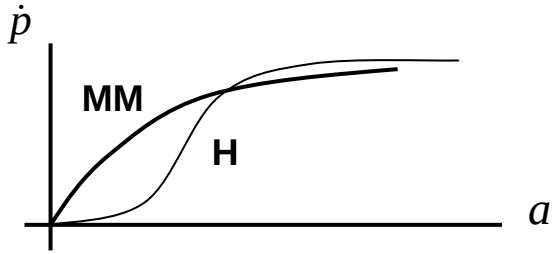
Enzyme kinetics

Michaelis-Menten kinetics

quasi-steady state assumption

A + E ↔ AE → AE + P

$$\dot{p} = \frac{V_{max} a}{K + a} - \gamma p$$



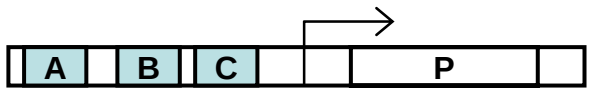
Hill function

- binding cooperativity, sigmoidal dependence, $h > 1$

$$\dot{p} = \frac{V_{max} a^h}{K + a^h} - \gamma p$$

Shea & Ackers

- Binding of regulatory factors approaches equilibrium



$$\dot{p} = \frac{P(a,b,c)}{Q(a,b,c)} - \gamma p$$

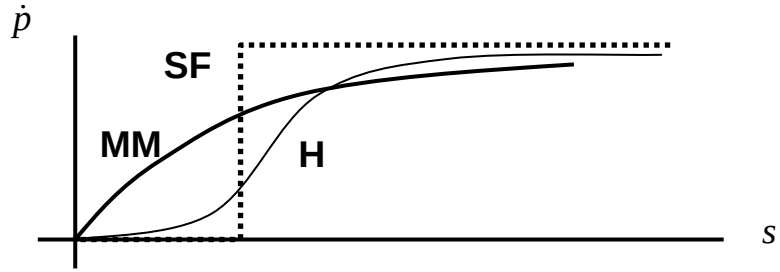
Results

- Conditions for multistability and oscillations in systems with feedback (Griffith 1968, Tyson & Othmer 1978, Thomas 1981, Thieffry et al. 1995, etc.)
- Relation between multistability and structure of regulatory region (Keller 1994, Wolf & Eeckman 1998)
- Modeling of phage λ–system, *lac* operon, *trp* operon, segment polarity networks, etc.

Piecewise-linear kinetics

Step function

$$\dot{p} = V_{\max} h(s, s^{\#}) = \begin{cases} V_{\max} & s > s^{\#} \\ 0 & s < s^{\#} \end{cases}$$



In general

$$\dot{x}_i = \sum_j \kappa_{ij} b_{ij}(\boldsymbol{x}, \boldsymbol{x}_{ij}^{\#}) - \gamma_i x_i \qquad b_{ij}(\boldsymbol{x}, \boldsymbol{x}_{ij}^{\#}) \in \{0,1\}$$

Results

Snoussi & Thomas 1993 – analytical determination of steady states

Mestl et al. 1996 – analysis of periodic orbits, classification of chaotic dynamics

De Jong 2003 – qualitative dynamics

Applications to number of prokaryotic and eukaryotic systems

Delay equations

- Required to account for transport phenomena

$$\frac{d}{dt}x(t) = \frac{kx^2(t-\tau)}{K + x^2(t-\tau)} - \gamma x(t) + b$$

Smolen 1998 – delayed self-activation exhibits multistability, delayed two-gene system with activation and repression gives oscillations

PDE models

$$\frac{\partial x_i}{\partial t} = f_i(\boldsymbol{x}) + \delta_i \frac{\partial^2 x_i}{\partial l^2}$$

- Pattern formation in reaction-diffusion systems
- Applications to embryo development and segmentation

Combined models

McAdams & Shapiro 1995 – combine boolean logic with ODEs to model phage λ–system

Stochastic effects

Assumptions

- Concentrations are discrete-valued random variables X_i
- Requires knowledge of detailed chemical steps
- Transitions given by master equation

$$\frac{\partial}{\partial t} P(\boldsymbol{X}, t) = \sum_r \left(\beta_r P(\boldsymbol{X} - \boldsymbol{v}_r, t) - \alpha_r P(\boldsymbol{X}, t) \right)$$

- Numerical simulations use Gillespie algorithm

Results

Thattai & Oudenaarden 2001 – control of noise by feedback loops, cascades

Barkai & Leibler – fluctuations can destabilize steady states or induce oscillations

Paulsson, Berg, & Ehernberg 2000 – stochastic focusing

Kepler & Elston 2001 – stochasticity induces bistability

Summary

Cons

- Multitude of levels of description available – difficult to gauge the correct amount of detail needed
- Lack of information about the parameters of the system
- Lack of experimental data for quantitative comparison of models and data
- Qualitative evaluation of models is subjective

Pros

- Amount of data is growing exponentially
- Time resolved *in situ* measurements possible
- Opportunity to do interesting and applicable mathematics

References

H. De Jong, Modeling and simulation of genetic regulatory systems: a literature review, *J. Comp. Bio.* **9**, 67-103 (2002).

P. Smolen, D.A. Baxter, & J.H. Byrne, Modeling transcriptional control in gene networks – methods, recent results, and future directions, *Bull. Math. Biol.* **62**, 247-292 (2000).

J. Hasty, D. McMillen & J.J. Collins, Engineered gene circuits, *Nature* **420**, 224-230 (2002).

D.M. Wolf, & A.P. Arkin, Motifs, modules and games in bacteria, *Curr. Opin. Microbiol.* **6**, 125-134 (2003).

H.H. McAdams & A.P. Arkin, Simulation of prokaryotic genetic circuits, *Annu. Rev. Biophys. Biomol. Struct.* **27**, 199-224 (1998).

M. Ptashne & A. Gann, *Genes & Signals*, Cold Spring Harbor Laboratory Press, 2002.

B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, & J.D. Watson, *Molecular biology of the cell*, Garland Publishing, 2002.