Systems biology and complexity research

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA

Interdisciplinary Challenges for Complexity Sciences

Brussels, 27.– 28.05.2009
Web-Page for further information:

http://www.tbi.univie.ac.at/~pks
1. Complex networks in cellular regulation
2. Experimental data and modeling in biology
3. Parameter determination and reverse engineering
4. Gene regulation dynamics
5. Inverse bifurcation analysis
6. Current challenges in biology
1. Complex networks in cellular regulation
2. Experimental data and modeling in biology
3. Parameter determination and reverse engineering
4. Gene regulation dynamics
5. Inverse bifurcation analysis
6. Current challenges in biology
Replication: DNA ⇒ 2 DNA

Transcription: DNA ⇒ RNA

Metabolism

Food

Waste

Nucleotides
Amino Acids
Lipids
Carbohydrates
Small Molecules

Ribosom

mRNA

Protein

Translation: RNA ⇒ Protein
A model genome with 12 genes

Sketch of a genetic and metabolic network
The reaction network of cellular metabolism published by Boehringer-Mannheim.
The citric acid or Krebs cycle (enlarged from previous slide).

The reaction network of cellular metabolism published by Boehringer-Mannheim.
**E. coli:** Genome length $4 \times 10^6$ nucleotides
Number of cell types 1
Number of genes 4 460

Four books, 300 pages each

**Man:** Genome length $3 \times 10^9$ nucleotides
Number of cell types 200
Number of genes $\approx 30 000$

A library of 3000 volumes,
300 pages each

Complexity in biology
1. Complex networks in cellular regulation

2. Experimental data and modeling in biology

3. Parameter determination and reverse engineering

4. Gene regulation dynamics

5. Inverse bifurcation analysis

6. Current challenges in biology
From qualitative data to quantitative modeling

Genomics, transcriptomics, proteomics

Metabolomics, functional genomics

Computational systems biology
Analysis by gel electrophoresis

Jeff Rogers, Gerald F. Joyce.

RNA 7:395-404, 2001
Gene expression DNA microarray representing 8613 human genes used to study transcription in the response of human fibroblasts to serum.

V.R.Iyer et al., *Science* **283**: 83-87, 1999
Embryonic stem cell

SOM-based “GEDI maps”
(Eichler, G.S. et al., *Bioinformatics* 2003)

Hsiao, L.L. et al., *Physiol.Genomics* 2001
Affymetrix, ~ 7000 genes

Drawings by Stuart A. Kauffman, 2009
A pH-modulated, self-replicating peptide
Shao Yao, Indraneel Ghosh, Reena Zutshi, Jean Chmielewski.
The elements of the simulation tool MiniCellSim

*SBML: Bioinformatics* 19:524-531, 2003;
1. Complex networks in cellular regulation
2. Experimental data and modeling in biology
3. **Parameter determination and reverse engineering**
4. Gene regulation dynamics
5. Inverse bifurcation analysis
6. Current challenges in biology
**Kinetic differential equations**
\[ \frac{dx}{dt} = f(x; k); \ x=(x_1, \ldots, x_n); \ k=(k_1, \ldots, k_m) \]

**Reaction diffusion equations**
\[ \frac{\partial x}{\partial t} = D \nabla^2 x + f(x; k) \]

**Parameter set**
\[ k_j(T, p, pH, I, \ldots); \ j=1, 2, \ldots, m \]

**General conditions**: \( T, p, pH, I, \ldots \)
**Initial conditions**: \( x(0) \)

**Boundary conditions**: boundary \( \partial \), normal unit vector \( \hat{u} \)
- Dirichlet: \( x^S = g(r, t) \)
- Neumann: \( \frac{\partial x}{\partial u} = \hat{u} \cdot \nabla x^S = g(r, t) \)

The forward problem of chemical reaction kinetics (Level I)
The forward problem of biochemical reaction kinetics (Level I)
The inverse problem of biochemical reaction kinetics (Level I)

- **Kinetic differential equations**
  \[
  \frac{dx}{dt} = f(x; k); \quad x = (x_1, \ldots, x_n); \quad k = (k_1, \ldots, k_m)
  \]

- **Reaction diffusion equations**
  \[
  \frac{\partial x}{\partial t} = D \nabla^2 x + f(x; k)
  \]

- **General conditions**
  \( T, p, pH, I, \ldots \)

- **Initial conditions**
  \( x(0) \)

- **Boundary conditions**
  - Dirichlet:
    \( x^S = g(r, t) \)
  - Neumann:
    \( \frac{\partial x}{\partial u} = \hat{u} \cdot \nabla x^S = g(r, t) \)

- **Data from measurements**
  \( x(t_j); \quad j = 1, 2, \ldots, N \)

Genome: Sequence \( I_G \)

Parameter set
\( k_j (I_G; T, p, pH, I, \ldots); \quad j = 1, 2, \ldots, m \)
General conditions: $T$, $p$, pH, $I$, ...
Initial conditions: $x(0)$

Boundary conditions:
- Dirichlet: $x^S = g(r, t)$
- Neummann: $\frac{\partial x}{\partial u} = \hat{u} \cdot \nabla x^S = g(r, t)$

Kinetic differential equations
\[ \frac{dx}{dt} = f(x; k); \quad x = (x_1, \ldots, x_n); \quad k = (k_1, \ldots, k_m) \]

Reaction diffusion equations
\[ \frac{\partial x}{\partial t} = D \nabla^2 x + f(x; k) \]

Parameter set
\[ k_j(I_G; T, p, pH, I, \ldots); \quad j = 1, 2, \ldots, m \]

Bifurcation analysis
\[ \gamma(k_i, k_j; k) \]

The forward problem of bifurcation analysis (Level II)
The inverse problem of bifurcation analysis (Level II)

Kinetic differential equations
\[ \frac{dx}{dt} = f(x; k); \quad x = (x_1, \ldots, x_n); \quad k = (k_1, \ldots, k_m) \]

Reaction diffusion equations
\[ \frac{\partial x}{\partial t} = D \nabla^2 x + f(x; k) \]

General conditions: \( T, p, pH, I, \ldots \)
Initial conditions: \( x(0) \)

Boundary conditions:
- Dirichlet: \( x^S = g(r, t) \)
- Neumann: \( \frac{\partial x}{\partial u} = \hat{u} \cdot \nabla x^S = g(r, t) \)

Parameter set
\[ k_j(I_G; T, p, pH, I, \ldots); \quad j = 1, 2, \ldots, m \]

Bifurcation pattern
\[ \mathcal{Y}(k_1, k_2; k) \]
1. Complex networks in cellular regulation
2. Experimental data and modeling in biology
3. Parameter determination and reverse engineering
4. **Gene regulation dynamics**
5. Inverse bifurcation analysis
6. Current challenges in biology
Three states of a gene regulated by activator and repressor

State I: basal state

State II: active state

State III: inactive state
Dynamic patterns of gene regulation I: Simple two-gene systems

Stefanie Widder\textsuperscript{a}, Josef Schicho\textsuperscript{b}, Peter Schuster\textsuperscript{a,c,*}

\textsuperscript{a}Institut für Theoretische Chemie der Universität Wien, Wilhelmsplatz 17, A-1090 Wien, Austria
\textsuperscript{b}RICAM—Research Institute for Computational and Applied Mathematics of the Austrian Academy of Sciences, Altenbergerstraße 69, A-4040 Linz, Austria
\textsuperscript{c}Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Received 24 February 2006; revised version received 7 January 2007; accepted 8 January 2007

Abstract

Regulation of gene activities is studied by means of computer assisted mathematical analysis of ordinary differential equations (ODEs) derived from binding equilibria and chemical reaction kinetics. Here, we present results on cross-regulation of two genes through activator and/or repressor binding. Arbitrary (differential) binding functions can be used but systematic investigations are presented for gene-regulator complexes with integer valued Hill coefficients up to \( n = 4 \). The dynamics of gene regulation is derived from bifurcation patterns of the underlying systems of kinetic ODEs. In particular, we present analytical expressions for the parameter values at which one-dimensional (transcritical, saddle-node or pitchfork) and/or two-dimensional (Hopf) bifurcations occur. A classification of regulatory states is introduced, which makes use of the sign of a 'regulatory determinant' \( D \) (being the determinant of the block in the Jacobian matrix that contains the derivatives of the regulator binding functions) \( 4 \) systems with \( D < 0 \), observed, for example, if both proteins are activators or repressors, to give rise to one-dimensional bifurcations only and lead to bistability for \( n > 2 \) and \( 6 \) systems with \( D > 0 \), found for combinations of activation and repression, sustain a Hopf bifurcation and undamped oscillations for \( n > 2 \). The influence of basal transcription activity on the bifurcation patterns is described. Binding of multiple subunits can lead to richer dynamics than pure activation or repression states (if intermediates between the unbound state and the fully saturated DNA initiate transcription).

Then, the regulatory determinant \( D \) can adopt both signs, plus and minus.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Basal transcription; Bifurcation analysis; Cooperative binding; Gene regulation; Hill coefficient; Hopf bifurcation

1. Introduction

Theoretical work on gene regulation goes back to the 1960s (Monod et al., 1963) soon after the first repressor protein had been discovered (Jacob and Monod, 1961). A little later the first paper on oscillatory states in gene regulation was published (Goodwin, 1965). The interest in gene regulation and its mathematical analysis never ceased (Tiwari et al., 1974; Tyson and Othmer, 1978; Smith, 1987) and saw a great variety of different attempts to design models of genetic regulatory networks that can be used in systems biology for computer simulation of genetic and metabolic networks. Most models in the literature aim at a minimalistic dynamic description which, nevertheless, tries to account for the basic regulatory functions of large networks in the cell in order to provide a better understanding of cellular dynamics. A classic in general regulatory dynamics is the monograph by Thomas and D’Ari (1990). The currently used mathematical methods comprise application of Boolean logic (Thomas and Kaufman, 2001b; Savageau, 2001; Albert and Othmer, 2003), stochastic processes (Hume, 2000) and deterministic dynamic models, examples are Cherry and Adler (2000), Bindschadler and Sneyd (2001) and Kobayashi et al. (2003) and the recent elegant analysis of bistability (Craciun et al., 2003).
Cross-regulation of two genes

\[
\begin{align*}
G_1 + \text{monomers} & \rightarrow G_1 + Q_1 \\
Q_1 + \text{monomers} & \rightarrow Q_1 + P_1 \\
G_2 + \text{monomers} & \rightarrow G_2 + Q_2 \\
Q_2 + \text{monomers} & \rightarrow Q_2 + P_2
\end{align*}
\]

synthesis \hspace{5cm} \text{degradation}

\[
\begin{align*}
Q_1 & \rightarrow \emptyset \\
Q_2 & \rightarrow \emptyset \\
P_1 & \rightarrow \emptyset \\
P_2 & \rightarrow \emptyset
\end{align*}
\]
Activation: \[ F_i(p_j) = \frac{p_j^n}{K + p_j^n} \]

Repression: \[ F_i(p_j) = \frac{K}{K + p_j^n} \]

\[ i, j = 1, 2 \]

Gene regulatory binding functions
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S , E</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>2</td>
<td>E , B(E,P)</td>
<td>S</td>
<td>S , B(P₁,P₂)</td>
</tr>
<tr>
<td>3</td>
<td>E , B(E,P)</td>
<td>S , O</td>
<td>S , B(P₁,P₂)</td>
</tr>
<tr>
<td>4</td>
<td>E , B(E,P)</td>
<td>S , O</td>
<td>S , B(P₁,P₂)</td>
</tr>
</tbody>
</table>

S ...... stable point attractor  
E ...... extinction  
O ...... oscillations  
B ...... bistability
An example analyzed and simulated by MiniCellSim

Stable stationary state

Limit cycle oscillations

Fading oscillations caused by a stable heteroclinic orbit

Increasing inhibitor strength

Hopf bifurcation

Bifurcation to May-Leonhard system
The repressilator limit cycle
The repressilator heteroclinic orbit
The repressilator heteroclinic orbit (logarithmic time scale)
The repressilator limit cycle
1. Complex networks in cellular regulation
2. Experimental data and modeling in biology
3. Parameter determination and reverse engineering
4. Gene regulation dynamics
5. Inverse bifurcation analysis
6. Current challenges in biology
The bifurcation manifold
Definition of the forward operator $F(p)$
Iterative solution for $\min J(p)$
\[ \begin{align*}
\dot{x}_i &= \beta_i (y_i - x_i) \\
\dot{y}_i &= \alpha_i \left( \frac{1 - \delta_i}{1 + x_{i-1}^{h_i} \mod n} + \delta_i \right) - y_i, \quad i = 0, \ldots, n - 1 \\
\alpha_i &= \alpha, \quad \beta_i = \beta, \quad h_i = h, \quad \delta_i = \delta
\end{align*} \]

\[ p_i = (\alpha, \beta) \quad \text{and} \quad (10^{-4}, 0) \leq (\delta, h) \leq (10^{-1}, 2) \]

Inverse bifurcation analysis of the repressilator model

Inverse bifurcation analysis of the repressilator model

\[
\frac{d}{dt}[\text{pRB}] = k_1 \frac{[\text{E2F1}]}{K_{m_1} + [\text{E2F1}]} \frac{J_{11}}{J_{11} + [\text{pRB}]} - \phi_{\text{pRB}}[\text{pRB}]
\]

\[
\frac{d}{dt}[\text{E2F1}] = k_p + k_1 \frac{a^2 + [\text{E2F1}]^2}{K_{m_2}^2 + [\text{E2F1}]^2} \frac{J_{12}}{J_{12} + [\text{pRB}]} - \phi_{\text{E2F1}}[\text{E2F1}]
\]

\[
\frac{d}{dt}[\text{AP1}] = F_m + k_25 [\text{E2F1}] \frac{J_{15}}{J_{15} + [\text{pRB}]} \frac{J_{65}}{J_{11} + [\text{pRB}']} - \phi_{\text{AP1}}[\text{AP1}]
\]

A simple dynamical cell cycle model

*Bioessays* 24:1095-1109, 2002
A simple dynamical cell cycle model

Inverse bifurcation analysis of a dynamical cell cycle model

1. Complex networks in cellular regulation
2. Experimental data and modeling in biology
3. Parameter determination and reverse engineering
4. Gene regulation dynamics
5. Inverse bifurcation analysis
6. *Current challenges in biology*
Explanation of important global properties

homeostasis

robustness

stability against mutation

self-repair or regeneration

.........
The bacterial cell as an example for a simple form of autonomous life

Escherichia coli genome:
4 million nucleotides
4460 genes

The structure of the bacterium *Escherichia coli*
Evolution does not design with the eyes of an engineer, evolution works like a tinkerer.

Gene is not a typical four-letter word. It is not offensive. It is not difficult. It is not a difficult term to define. And yet, it is one of the most widely used in biology, and yet, it is one of the most difficult to define.

The difficulty to define the notion of "gene". Helen Pearson, Nature 441: 399-401, 2006
ENCODE stands for **ENCyclopedia Of DNA Elements**.

**ENCODE Project Consortium.**
Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.
Web-Page for further information:

http://www.tbi.univie.ac.at/~pks