Basic modeling approaches for biological systems -II

Mahesh Bule
Modeling Metabolism

• Metabolism is a general term for two kinds of reactions:
  - Catabolic reactions (breakdown of complex compounds to get energy and building blocks)
  - Anabolic reactions (construction of complex compounds used in cellular function)
Illustration of metabolism and its role in bottom up modeling approach
Approaches for modeling metabolic process
Modeling relation to real living world

REAL WORLD

Biological System: Cellular Metabolism

new data and novel experimental tools

decoding

Mathematical World

Formal System: Mathematical Model

Computational methods and rules of inference

encoding
Scheme of metabolic model construction

A METABOLIC NETWORK

Step 1: Assemble the building blocks
A list of participating metabolites and reactions, distinguishing between external and internal metabolites, is assembled.

Step 2: Define connectivity and interactions
The metabolites are transformed and translocated by reactions and transport processes, both modified by allosteric regulation.

Step 3: Express each interaction quantitatively
Each interaction (reactions and allosteric modifications) is associated with a rate equation and kinetic parameters.

Model Validation: Comparison to experimental data
The model is validated by comparison with experimental data not used in the construction process.

Model Interrogation: Analysis of emergent properties
Study of metabolism

• Generally metabolism is studied at three levels of abstraction:

  – Enzyme kinetic investigates the dynamic properties of the individual reactions in isolation

  – The network character of metabolism is studied with stoichiometric analysis considering the balance of compound production and degradation

  – Metabolic control analysis quantifies the effect of perturbations in the network employing the individual dynamics of concentration changes and their integration in network
Example: Modeling glucose utilization by microorganism

• Glycolysis pathway
Metabolic Network

• Basic elements of a metabolic network model are
  – Substances with their concentrations
  – The reaction or transport processes changing the concentration of substance
Modeling network function

- **Kinetic models**
  - Dynamical systems
  - Requires kinetic constants (mostly unknown)

- **Approx. kinetics**

- **Constraint-based models**
  - Optimization theory
  - Constrained space of possible, steady-state network behaviors

- **Conventional functional models**
  - Probabilistic models, discrete models, etc'

- **Topological analysis**
  - Graph theory
  - Structural network properties: degree distribution, centrality, clusters, etc’
Kinetic models for metabolism

• Dynamics of metabolic behavior over time
  – Metabolite concentrations
  – Enzyme concentrations
  – Enzyme activity rate – depends on enzyme concentrations and metabolite concentrations
  – Solved using a set of differential equations

• Impossible to model large-scale networks
  – Requires specific enzyme rates data
  – Too complicated
Basic definition in modeling

\[ r_1 = \frac{v_{\text{max}1} S}{K_{m1} + S} \]

Low \( K_m \) will be the path with the higher flux (all other factors being equal).

Low \( K_m \) also means a strong interaction between substrate and enzyme.

These two curves have the same \( v_{\text{max}} \), but their \( K_m \) values differ by a factor of 2.
Basic definition in modeling

- **Total flux**: 
  \[
  F_{\text{tot}} = \frac{v_{\text{max}1} S}{K_{m1} + S} + \frac{v_{\text{max}2} S}{K_{m2} + S}
  \]

- **Selectivity**
  \[
  \frac{F_1}{F_2} = \frac{\frac{v_{\text{max}1} S}{K_{m1} + S}}{\frac{v_{\text{max}2} S}{K_{m2} + S}}
  \]
  \[
  \frac{r_1}{r_2} = \frac{v_{\text{max}1}}{v_{\text{max}2}} \left( \frac{K_{m2} + S}{K_{m1} + S} \right)
  \]
Simplified metabolism - upstream end of glycolysis

Glucose → Glucose 6-Phosphate → Fructose 6-Phosphate → Fructose 1,6-Bisphosphate → Pyruvate

Additional reactions:
- ADP + ATP → AMP + ATP (v8)
- ATP → ADP (v7)
- ADP → ATP (v6)
How do you model this?

• What information is needed?
  – equations for each v
  – initial concentrations of each metabolite
Mass balances

\[
\frac{d\text{Gluc6P}}{dt} = v1 - v2 - v3
\]
\[
\frac{d\text{Fruc6P}}{dt} = v3 - v4
\]
\[
\frac{d\text{Fruc1,6P2}}{dt} = v4 - v5
\]
\[
\frac{d\text{ATP}}{dt} = -v1 - v2 - v4 + v6 - v7 - v8
\]
\[
\frac{d\text{ADP}}{dt} = v1 + v2 + v4 - v6 + v7 + 2v8
\]
\[
\frac{d\text{AMP}}{dt} = -v8
\]
Defining reaction rate equations

\[
v_1 = \frac{v_{max,1}[ATP(t)][Glucose]}{1 + \frac{[ATP(t)]}{K_{ATP1}} + \frac{[Glucose]}{K_{Glucose1}} + \left[ \frac{[ATP(t)]}{K_{ATP1}} \times \frac{[Glucose]}{K_{Glucose1}} \right]}
\]

\[
v_2 = k_2[ATP(t)][Gluc6P]
\]

\[
v_4 = \frac{v_{max,4}(Fruc6P(t))^2}{K_{Fruc6P4} \left( 1 + \kappa \left( \frac{ATP(t)}{AMP(t)} \right)^2 \right) + (Fruc6P(t))^2}
\]

\[
v_5 = k_5[Fruc1,6P2]
\]
Defining reaction rate equations

\[ v_6 = k_6 \text{ADP}(t) \]

\[ v_7 = k_7 \text{ATP}(t) \]

\[ v_8 = k_{8f} \text{ATP}(t) \cdot \text{AMP}(t) - k_{8r} (\text{ADP}(t))^2 \]
Selected packages for kinetic analysis

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copasi</td>
<td><a href="http://www.copasi.org/">http://www.copasi.org/</a></td>
</tr>
<tr>
<td>CellDesigner</td>
<td><a href="http://www.celldesigner.org/">http://www.celldesigner.org/</a></td>
</tr>
<tr>
<td>E-CELL</td>
<td><a href="http://www.e-cell.org/">http://www.e-cell.org/</a></td>
</tr>
<tr>
<td>Cellware</td>
<td><a href="http://www.cellware.org/">http://www.cellware.org/</a></td>
</tr>
<tr>
<td>JDesigner/Jarnac</td>
<td><a href="http://sbw.kgi.edu/sbwWiki/sysbio/jdesigner/">http://sbw.kgi.edu/sbwWiki/sysbio/jdesigner/</a></td>
</tr>
<tr>
<td>SB Toolbox</td>
<td><a href="http://www.sbtoolbox2.org/">http://www.sbtoolbox2.org/</a></td>
</tr>
<tr>
<td>SBRT</td>
<td><a href="http://www.bioc.uzh.ch/wagner/software/SBRT/">http://www.bioc.uzh.ch/wagner/software/SBRT/</a></td>
</tr>
</tbody>
</table>
Constraint Based Models for Metabolism

- Provides a steady-state description of metabolic behavior
  - A single, constant flux rate for each reaction
  - Ignores metabolite concentrations
  - Independent of enzyme activity rates
- Assume a set of constraints on reaction fluxes
- Genome scale models

Flux rate:
\[ \mu \text{-mol} / (\text{mg} \times \text{h}) \]
Constraint Based Models for Metabolism

• Find a steady-state flux distribution through all biochemical reactions

• Under the constraints:
  – Mass balance: metabolite production and consumption rates are equal
  – Thermodynamic: irreversibility of reactions
  – Enzymatic capacity: bounds on enzyme rates
  – Availability of nutrients
Additional Constraints

• Transcriptional regulatory constraints
  – Boolean representation of regulatory network
• Energy balance analysis
  – Loops are not feasible according to thermodynamic principles
• Reaction directionality
  – Depending on metabolite concentrations
Mathematical Representation of Constraint Based models

- Stoichiometric matrix – network topology with stoichiometry of biochemical reactions

\[ \text{Glucose} + \text{ATP} \rightarrow \text{Glucose-6-Phosphate} + \text{ADP} \]

\[
\begin{array}{c}
\text{Glucose} \\
\text{ATP} \\
\text{G-6-P} \\
\text{ADP}
\end{array}
\begin{array}{c}
-1 \\
-1 \\
+1 \\
+1
\end{array}
\]

Mass balance
\[ S \cdot v = 0 \]
Subspace of \( \mathbb{R}^n \)

Thermodynamic
\[ v_i > 0 \]
Convex cone

Capacity
\[ v_i < v_{\text{max}} \]
Bounded convex cone
Constraint-based modeling applications

- Phenotype predictions:
  - Growth rates across media
  - Knockout lethality
  - Nutrient uptake/secretion rates
  - Intracellular fluxes
  - Growth rate following adaptive evolution

- Bioengineering:
  - Strain design – overproduce desired compounds

- Biomedical:
  - Predict drug targets for metabolic disorders

- Studying an array of questions regarding:
  - Dispensability of metabolic genes
  - Robustness and evolution of metabolic networks
Phenotype Predictions: Flux Predictions

• Predict metabolic fluxes following gene knockouts
• Search for short alternative pathways to adapt for gene knockouts (Regulatory On/Off Minimization-ROOM)
Strain design: maximizing metabolite production rate

• Identify a set of gene whose knockout increases the production rate of some metabolite
• The knockout of reaction v3 increases the production rate of metabolite F

Suboptimal mutant flux distribution

Opt-knock mutant flux distribution
Minimization Of Metabolic Adjustment (MOMA)

- The flux distribution after a knockout is close to the wild-type’s state under the Euclidian norm

Regulatory On/Off Minimization (ROOM)

- Minimize the number of Boolean flux changes from the wild-type’s state
ROOM vs. MOMA

• ROOM predicts metabolic steady-state after adaptation
  • Provides accurate flux predictions
  • Preserved flux linearity
  • Finds alternative pathways
  • Predicts steady-state growth rates

• MOMA predicts transient metabolic states following the knockout
  • Provides more accurate transient growth rates
Altering Phenotypic Potential

• Explaining gene dispensability
  – Only 32% of yeast genes contribute to biomass production in rich media
  – Considered one arbitrary optimal growth solution

• OptKnock – Identify gene deletions that generate desired phenotype

• OptStrain – Identify strains which can generate desired phenotypes by adding/deleting genes
Modeling Gene Knockouts

- Gene knockout
- Enzyme knockout
- Reaction knockout
Regulatory On/Off Minimization (ROOM)

- Predicts the metabolic steady-state following the adaptation to the knockout
- Assumes the organism adapts by minimizing the set of regulatory changes

- Finds flux distribution with minimal number of Boolean flux changes
Example of network

(a) Stoichiometric matrix

<table>
<thead>
<tr>
<th></th>
<th>v1</th>
<th>v2</th>
<th>v3</th>
<th>v4</th>
<th>v5</th>
<th>v6</th>
<th>b1</th>
<th>b2</th>
<th>b3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>-2</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>byp</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>cof</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(b) Wild-type

(c) MOMA - knockout

(d) ROOM - knockout
Flux Balance Analysis (FBA)

- Finds flux distribution with maximal growth rate
- Biomass production rate represents growth rate
- Solver need to be used is Linear Programming (LP)

Max $v_{\text{gro}}$, - maximize growth
s.t
$S \cdot v = 0$, - mass balance constraints
$v_{\text{min}} \leq v \leq v_{\text{max}}$ - capacity constraints
Topological Methods

• Network based pathways:
  – Extreme Pathways
  – Elementary Flux Modes

• Decomposing flux distribution into extreme pathways

• Extreme pathways defining phenotypic phase planes

• Uniform random sampling
Extreme Pathways and Elementary Flux Modes

- Unique set of vectors that spans a solution space
- Consists of minimum number of reactions
- Extreme Pathways are systematically independent (convex basis vectors)
Flux Balance Analysis-Example

- **Step (I) : definition**

**Vertex** - substrate/metabolite concentration.

**Edge** - flux (conversion mediated by enzymes of one substrate into the other)

- **Internal flux edge**
- **External flux edge**
Flux Balance Analysis

• Step (II)- Dynamic mass balance

\[ \frac{dx}{dt} = S \cdot \nu \]

\[
\begin{bmatrix}
\frac{dA}{dt} \\
\frac{dB}{dt} \\
\frac{dC}{dt}
\end{bmatrix} = \begin{bmatrix}
-1 & -1 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & -1 & 0 \\
0 & 1 & -1 & -1 & 0 & 0 & -1
\end{bmatrix}
\begin{bmatrix}
\nu_1 \\
\nu_2 \\
\nu_3 \\
\nu_4 \\
b_1 \\
b_2 \\
b_3
\end{bmatrix}
\]

\[
\frac{dA}{dt} = -\nu_1 - \nu_2 + \nu_3 + b_1
\]

\[
\frac{dB}{dt} = \nu_1 + \nu_4 - b_2
\]

\[
\frac{dC}{dt} = \nu_2 - \nu_3 - \nu_4 - b_3
\]
Flux Balance Analysis

- Step (III)- Assumption of steady state

\[ \frac{dx}{dt} = S \cdot v \quad \Rightarrow \quad 0 = S \cdot v \]
Minimal model of glycolysis

**A: Reaction scheme**

- **G** to **G**
- **G** to 2 **TP**
- **TP** to **P**
- **G** to **G**

**B: Stoichiometry**

\[
N = \begin{bmatrix}
+1 & -1 & 0 & 0 & 0 & 0 \\
0 & +2 & -1 & 0 & -1 & 0 \\
0 & 0 & +1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 \\
0 & +2 & -2 & 0 & 0 & +1 \\
\end{bmatrix}
\]

**C: Reaction list:**

- v0: G_x → G
- v1: G + 2 ATP → 2 TP + 2 ADP
- v2: TP + 2 ADP → P + 2 ATP
- v3: P → P_x
- v4: TP → G
- v5: ATP → ADP

**D: The mass-balance equations**

\[
\frac{d}{dt} \begin{bmatrix}
G \\
TP \\
P \\
ATP \\
ADP
\end{bmatrix} = \begin{bmatrix}
+1 & -1 & 0 & 0 & 0 \\
0 & +2 & -1 & 0 & -1 \\
0 & 0 & +1 & -1 & 0 \\
0 & 0 & 0 & 0 & -1 \\
0 & +2 & -2 & 0 & 0 & +1 \\
\end{bmatrix} \begin{bmatrix}
v_0 \\
v_1 \\
v_2 \\
v_3 \\
v_4 \\
v_5
\end{bmatrix}
\]
Structural kinetic modeling

**CONSTRANINT (FLUX-BALANCE) ANALYSIS**

**PROS:**
- Requires no information on kinetic parameters
- Applicable to large 'genome-scale' systems

**CONS:**
- Allows no description of dynamic properties
- Allows no analysis of allosteric regulation

**DETAILED KINETIC MODELS**

**PROS:**
- Gives quantitative description of the dynamics
- Description includes regulatory properties

**CONS:**
- Estimation of kinetic parameters not feasible
- Computationally demanding for large systems

An intermediate approach: Structural Kinetic Modeling

Combines the advantages of a stoichiometric description with kinetic properties:

- Requires no knowledge of kinetic parameters
- Applicable to large 'genome-scale' system
- Gives quantitative description of the dynamics
- Description includes regulatory properties
Proposed flow for structural kinetic modeling

CELLULAR METABOLISM

Stoichiometric Analysis of Metabolism
Flux balance analysis
Thermodynamic constraints

$J_x = \Lambda \theta^\mu_x$

$^{13}$C-Flux Analysis
Metabolomics

Metabolic States
steady-state concentrations
flux distribution
thermodynamic feasibility

From Structure to Dynamics

Statistical Evaluation of Systems Dynamics
Identification of Stabilizing Sites
Comparison of Metabolic States

evaluate eigenvalues of the Jacobian

sample iteratively