A simple open system

- Two metabolites: A and B
- Four reactions:
  - Two internal reactions: R2 and R3
  - Two boundary (or external) reactions: R1 (source) and R4 (sink)
  - Three reactions are irreversible: R1, R2 and R4
  - One reaction is reversible: R3
- A simple system, without compartmentalization
iJO1366 - *Escherichia coli* K12

- Three compartments: cytoplasm, periplasmic space, growth medium
- Chemical compounds: 1135
- Localized metabolites: 1804
- 2578 reactions:
  - 1472 "enzymatic" reactions (one compartment)
  - 775 transport reactions (two compartment)
  - 329 boundary reactions (source or sink)
  - 1933 reactions are irreversible
  - 633 reactions are reversible

Mass balances

\[
\frac{dc_A}{dt} = +r_1 - r_2 - r_3 \\
\frac{dc_B}{dt} = +r_2 + r_3 - r_4
\]

- The concentration of A and B, \(c_A(t)\) and \(c_B(t)\) are the solution of the above ODE system
- The rate of reaction \(i\): \(r_i\) has the following form in the general case
  \[
r_i = r_i(c_A, c_B, t, \theta)
\]
Enzymatic kinetics

- Michaelis-Menten kinetics - possibly the simplest model for the kinetics of an irreversible reaction catalyzed by an enzyme

\[ S + E \leftrightarrow SE \rightarrow S + P \]

\[ r(S) = \frac{V_{\text{max}}S}{K_m + S} \quad 0 \leq r(S) < R_{\text{max}} \]

- This is a non-linear function.
- The value of the $V_{\text{max}}$ and $K_m$ parameters must be available...
- Note that the reaction rate is bounded (good news)

Enzymatic kinetics

- The Monod-Wyman-Changeux model for allosteric enzymes

\[ \nu_0(s) = \frac{V_{\text{max}}s}{(1 + K_Rs)} \left( \frac{K_R + K_L}{1 + K_Rs} \right)^{n-1} \]

- This is a non-linear function
- The values of the five parameters must be available
- Note that the reaction rate is bounded (good news again)
Matrix notation

\[
\frac{dc_A}{dt} = r_1 - r_2 - r_3 \\
\frac{dc_B}{dt} = r_2 + r_3 - r_4
\]

\[
c = \begin{pmatrix} c_A \\ c_B \end{pmatrix} \quad N = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix} \quad r = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}
\]

The stoichiometric matrix

- The stoichiometric matrix capture the topology of the network
- The stoichiometric coefficients are
  - negative for substrates
  - positive for products
- Reaction directions are NOT in the matrix!
- For large network, the stoichiometric matrix is usually sparse
A realistic stoichiometric matrix (yeast)

Simplifying the problem

\[
\frac{d}{dt} c = Nr
\]

Time-dependent concentration changes

Stoichiometric matrix time-invariant

Flux distribution time-variant

- The major difficulties in establishing a kinetics model lies in
  - Finding formulations for the non-linear enzymatic kinetics
  - Obtaining numerical values for the kinetics parameter

- Possible simplification: focus on the invariant \( N \) and ignore the details of kinetics => **stoichiometric and constraint-based models**
  - This is a serious simplification, however
Steady state hypothesis

\[ \frac{d}{dt}c = Nr \rightarrow 0 = Nr \]

- At steady state, the concentrations \( c \) and the rates \( r \) become constant: the consumption of every metabolite equal its production
- The system of ODE simplifies into an homogeneous system of linear equations

\[
\begin{align*}
n_{11}r_1 + n_{12}r_2 + \cdots + n_{1q}r_q &= 0 \\
n_{21}r_1 + n_{22}r_2 + \cdots + n_{2q}r_q &= 0 \\
\vdots & \vdots \\
n_{q1}r_1 + n_{q2}r_2 + \cdots + n_{qq}r_q &= 0
\end{align*}
\]

- (Quasi steady state hypothesis: Metabolic reactions and turnover rate of metabolites are fast with respect to other cellular process like protein synthesis. The above simplification still holds in a context where other "slow moving" variables are taken into account, e.g. regulation of protein synthesis)

Nullspace Analysis

- Trivial solution (not interesting): \( r = 0 \)
- Most often, there exist an infinite number of non-trivial solutions that lie in the right nullspace of the matrix \( N \). These solutions can be written as a linear combinations of the columns of the kernel matrix \( K \) associated to \( N \)

\[
N = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix} \quad \quad \quad K = \begin{pmatrix} 1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}
\]

Computing the kernel matrix is a classical problem of linear algebra

\[ NK = 0 \]

- and the non-trivial solutions can be written as a linear combination:

\[ r = \sum \alpha_i K_i ; \alpha_i \neq 0 \]

- Note that the kernel matrix, in general, is not a unique representation
Kernel matrix

$K = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & -1 \\
0 & 1 & 1
\end{pmatrix}$

- The nullspace analysis ignores the reaction directions !!!

Reaction correlations group (null space)

$K = \begin{pmatrix}
1 & 0 & \leftarrow R1 \\
0 & -1 & \leftarrow R2 \\
1 & 1 & \leftarrow R3 \\
1 & 0 & \leftarrow R4
\end{pmatrix}$

- The reaction R1 + R4 form a correlation group
  - It corresponds to the mass conservation in the system at steady state
  - Same flux
Another example (null space)

- **R6** is a zero flux reaction
- The **R4+R5** group has a problem (**a posteriori**):
  - The **M5** metabolite is a dead end!
  - Indeed the null space analysis ignore the direction of the reactions
- Solid arrows: operational sub-network allowing non-zero fluxes
- One coloured group: "same" flux
Flux Balance Analysis (FBA)

- Idea: Incorporate further constraints to limit network behaviour with respect to feasible steady state flux distributions
- Enzymatic reaction are catalyzed by enzymes that are available in a finite amount => place lower and upper bounds on admissible reactions flows

A (very) simple FBA problem

- A simple model with a single metabolite and 3 reactions
- All three reactions are irreversible (lower bound)
  \[
  V_{1,\text{min}} = V_{2,\text{min}} = V_{3,\text{min}} = 0
  \]
- An upper bound is placed on every reaction
  \[
  V_{1,\text{max}} = 6, \quad V_{2,\text{max}} = 8, \quad V_{3,\text{max}} = 10
  \]
- The flux space has 3 dimensions
- The null space has 2 dimensions
The solution space

- The flux space has 3 dimensions
- The nullspace has 2 dimensions
- The solution space has 2 dimensions (i.e. the nullspace) and is further restricted by the bounds placed on the reaction fluxes

Steady state and optimal solutions

- Steady state solutions can be located anywhere within this polygon
  - $V_2$ is max
  - $V_3$ is max

- Optimal solutions are located on the edges
  - $V_1$ is max
  - $V_3$ is max
Linear programming (LP)

\[ n_{11} \cdot r_1 + n_{12} \cdot r_2 + \cdots + n_{1q} \cdot r_q = 0 \quad UB_1 \leq r_i \leq UB_1 \]
\[ n_{21} \cdot r_1 + n_{22} \cdot r_2 + \cdots + n_{2q} \cdot r_q = 0 \quad UB_2 \leq r_2 \leq UB_2 \]
\[ \vdots \quad \vdots \quad \vdots \quad \vdots \quad = 0 \quad \vdots \quad \vdots \quad \vdots \quad \vdots \]
\[ n_{n1} \cdot r_1 + n_{n2} \cdot r_2 + \cdots + n_{nq} \cdot r_q = 0 \quad UB_q \leq r_q \leq UB_q \]

- "Linear programming (LP, or linear optimization) is a mathematical method for determining a way to achieve the best outcome (such as maximum profit or lowest cost) in a given mathematical model for some list of requirements represented as linear relationships."
- LP is heavily used in microeconomics and company management, for planning, production, transportation, technology and other issues. Typically, companies would like to maximize profits or minimize costs with limited resources.

Another example (FBA)

- The problem:
  - Define fluxes boundaries (some reaction can be turned off)
  - ...and compute the maximal flux on E3
Optimize E3 – single knockout

- Max flux on E3?

```
<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
</tr>
</thead>
<tbody>
<tr>
<td>lb</td>
<td>0</td>
<td>-10</td>
<td>0</td>
<td>0</td>
<td>-10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ub</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>
```

Optimize E3 – A source turned off

- Max flux on E3?

```
<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
</tr>
</thead>
<tbody>
<tr>
<td>lb</td>
<td>0</td>
<td>-10</td>
<td>0</td>
<td>0</td>
<td>-10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ub</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>
```
Optimize E3 – double knockouts

- Knockout of R3
- E1 turned off
- Max flux on E3?

The objective function (FBA)

- Any linear combination of reaction fluxes should be the function to maximize (or minimize)
- Maximizing the biomass reaction is quite common and often produces perfectly meaningful results at the light of biology
Summary

Credits

• Most graphics and concepts presented here were taken from Joerg Stelling ETHZ course in computational system biology and from Mathias Ganter PhD thesis