Introduction to Biological Modeling

Lecture 3: Metabolism
Oct. 6, 2010

Steve Andrews

Brent lab, Basic Sciences Division, FHCRC

Last week
- modeling cellular dynamics
- minimizing necessary parameters
- eukaryotic cell cycle
- positive feedback, oscillations
- databases

Copasi software

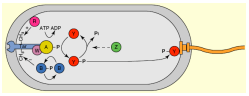
Reading

Covert, Schilling, Famili, Edwards, Goryanin, Selkov, and Palsson "Metabolic modeling of microbial strains *in silico*" *TRENDS in Biochemical Sciences* 26:179-186, 2001.

Klamt, Stelling, "Stoichiometric and constraint-based modeling" in *Systems Modeling in Cell Biology* edited by Szallasi, Stelling, and Periwal, MIT Press, Cambridge, MA, 2006.

Credit: Alberts, et al. *Molecular Biology of the Cell*, 3rd ed., 1994; http://www.aspencountry.com/product.asp?dept_id=460&pfid=35192.

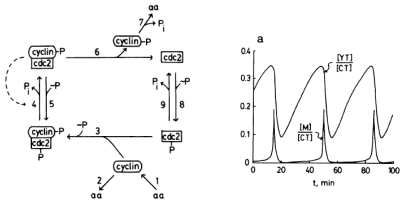
Small biochemical networks



E. coli chemotaxis

few chemical species
few reactions
enough known parameters

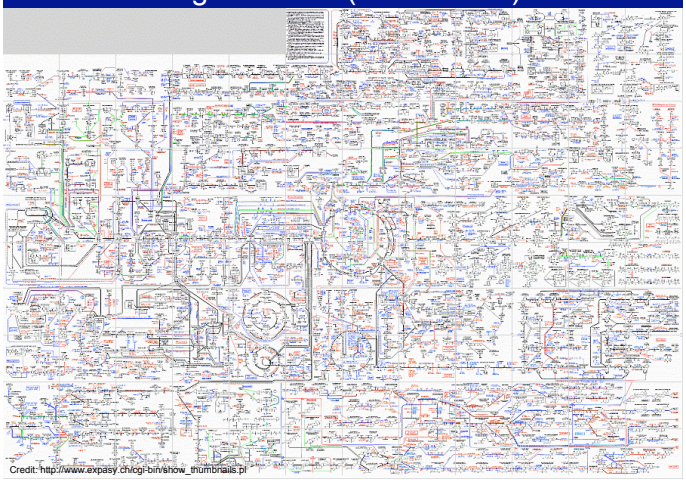
simple dynamics
can build model by hand
can understand intuitively



eukaryotic cell cycle

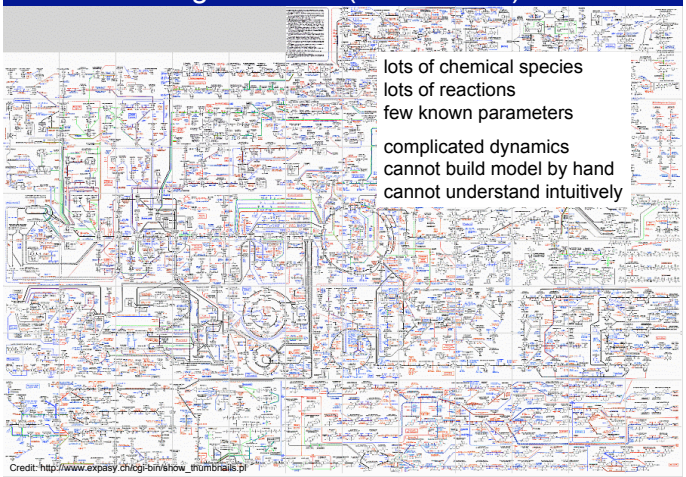
Credit: Andrews and Arkin, Curr. Biol. 16:R523, 2006; Tyson, *Proc. Natl. Acad. Sci. USA* 88:7328, 1991.

Big networks (metabolism)



Credit: http://www.expasy.ch/cgi-bin/show_thumbnails.pl

Big networks (metabolism)



lots of chemical species
lots of reactions
few known parameters

complicated dynamics
cannot build model by hand
cannot understand intuitively

Credit: http://www.expasy.ch/cgi-bin/show_thumbnails.pl

What is and isn't known

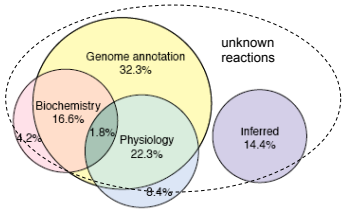
A lot is known
- 100s of enzymes
- 100s of reactions
- 100s of metabolites

Databases

KEGG
Kyoto Encyclopedia of Genes and Genomes

BRENDA
Braunschweig Enzyme Database

A lot is not known
- lots of enzymes, reactions, and metabolites
- most kinetic parameters
- most gene regulation



reaction sources in a genome-scale *H. pylori* metabolism model

Credit: Covert *et al. TRENDS in Biochemical Sciences* 26:179, 2001

Why study metabolism?

Basic science
- how do cells work?
- structure of complex networks

Medical
- metabolic disorders
- biosynthetic drugs (e.g. insulin)

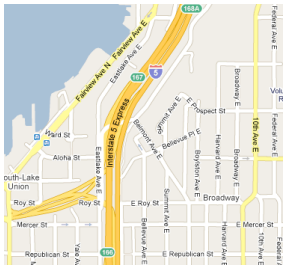
Bioengineering
- biofuels
- bioremediation of waste
- enzyme production

Credits: http://www.autobloggreen.com/2006/05/29/; https://isbibbio.wikispaces.com/Laundry+Detergents+and+Enzymes

Other complex networks

biological
food webs
gene regulatory networks
signaling networks

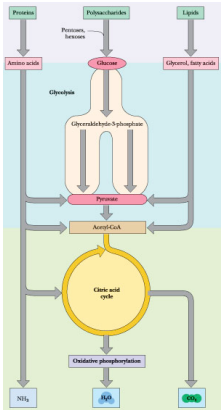
non-biological
road maps
physical internet (IP addresses)
internet websites
electronic circuits



Metabolism quick overview

Anabolism
- biosynthesis of proteins, polysaccharides, lipids, etc.

Catabolism
- breakdown of proteins, polysaccharides, lipids, etc. to make energy

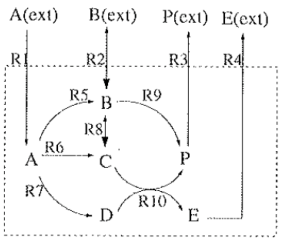


credit: http://web.virginia.edu/Heidi/chapter18/chp18.htm

Nomenclature
Constraint-based modeling
Pathway analysis
Metabolic control analysis
Copasi
Summary

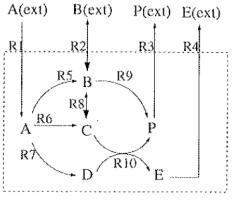
Terminology

stoichiometry
internal, external metabolites
reversible, irreversible reactions
enzyme catalysis
flux



Credit: Klamt and Stelling in *System Modeling in Cellular Biology* ed. Szallasi *et al.*, p. 73, 2006.

Stoichiometric matrix



Reaction network can be summarized in a matrix

Columns list reaction stoichiometry
Example for reaction R10:

$$C + D \rightarrow P + E$$

reactions

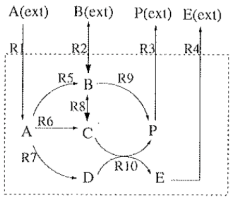
Rows list internal metabolite sources and sinks

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{matrix} \leftarrow A \\ \leftarrow B \\ \leftarrow C \\ \leftarrow D \\ \leftarrow E \\ \leftarrow P \end{matrix}$$

internal metabolites

R1 R2 R3 R4 R5 R6 R7 R8 R9 R10

Dynamical model



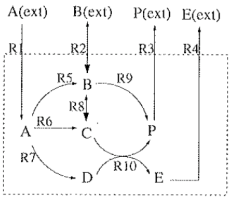
Dynamics for A

$$\frac{d[A]}{dt} = 1 \cdot r_1(t) - 1 \cdot r_5(t) - 1 \cdot r_6(t) - 1 \cdot r_7(t)$$

reaction rates (fluxes)

$r_1(t)$ = rate of reaction R1
$r_2(t)$ = rate of reaction R2
...
$r_{10}(t)$ = rate of reaction R10

Dynamical model



Dynamics for A

$$\frac{d[A]}{dt} = 1 \cdot r_1(t) - 1 \cdot r_5(t) - 1 \cdot r_6(t) - 1 \cdot r_7(t)$$

reaction rates (fluxes)

$r_1(t)$ = rate of reaction R1
$r_2(t)$ = rate of reaction R2
...
$r_{10}(t)$ = rate of reaction R10

math

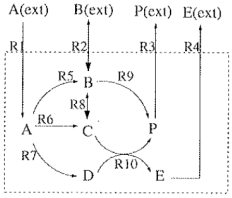
$$\frac{d}{dt}\begin{bmatrix}[A]\\[B]\\[C]\\[D]\\[E]\\[P]\end{bmatrix} = \begin{bmatrix}1&0&0&0&-1&-1&-1&0&0&0\\0&1&0&0&1&0&0&-1&-1&0\\0&0&0&0&0&1&0&1&0&-1\\0&0&0&0&0&0&1&0&0&-1\\0&0&0&-1&0&0&0&0&0&1\\0&0&-1&0&0&0&0&0&1&1\end{bmatrix} \begin{bmatrix}r_1(t)\\r_2(t)\\r_3(t)\\r_4(t)\\r_5(t)\\r_6(t)\\r_7(t)\\r_8(t)\\r_9(t)\\r_{10}(t)\end{bmatrix}$$

stoichiometric matrix

$$\frac{d\mathbf{c}}{dt} = \mathbf{N} \cdot \mathbf{r}(t)$$

c = metabolite concentrations
N = stoichiometric matrix
r(*t*) = reaction rate vector

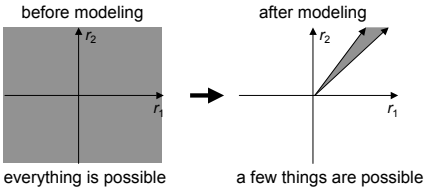
Constraint-based modeling



Main idea
We can infer a lot about the fluxes, just from the diagram (and assumptions).

- can predict metabolite production rates
- can infer missing reactions
- improves network understanding

flux space
1 dimension for each reaction (10 here)



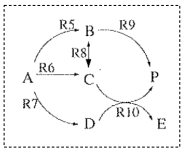
everything is possible

a few things are possible

Constraint 0: conservation relations

Some metabolite concentrations always change together

Biology



$[A]+[B]+[C]+[D]+[E]+[P] = \text{constant}$

6 metabolites
5 degrees of freedom

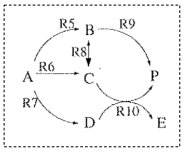
traffic analogy

total number of cars on island is constant

Conservation relations are always true
(regardless of dynamics, reaction directionality, etc.)

Constraint 0: conservation relations - math

Example



$[A]+[B]+[C]+[D]+[E]+[P] = \text{constant}$

6 metabolites
5 degrees of freedom

Conservation relations arise when rows of the stoichiometric matrix are linearly dependent

$$\mathbf{N} = \begin{array}{c}\begin{matrix}R5&R6&R7&R8&R9&R10\end{matrix}\\\begin{bmatrix}-1&-1&-1&0&0&0\\1&0&0&-1&-1&0\\0&1&0&1&0&-1\\0&0&1&0&0&-1\\0&0&0&0&0&1\\0&0&0&0&1&1\end{bmatrix}\end{array}\begin{matrix}A\\B\\C\\D\\E\\P\end{matrix}$$

i.e. there exists $\mathbf{y}^T$ for which:

$$\mathbf{y}^T\mathbf{N} = \mathbf{0}$$

In this case, the only solution is

$$\mathbf{y}^T = \begin{bmatrix}1&1&1&1&1&1\end{bmatrix}$$

which implies

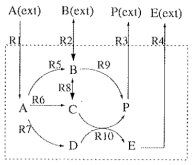
$$[A]+[B]+[C]+[D]+[E]+[P] = \text{constant}$$

$\mathbf{y}^T$ is the *left null-space* of **N**

Constraint 1: steady-state assumption

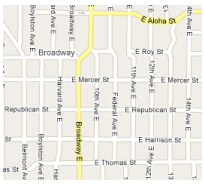
If system is at steady state, fluxes into and out of each metabolite are equal.

biology



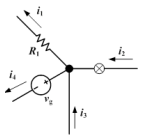
true when metabolic reactions are much faster than:
- external metabolite changes
- internal gene regulation

traffic analogy



the same number of cars enter and leave each intersection.

electronics

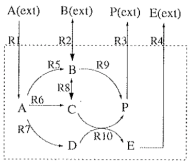


current entering and leaving a junction are equal (Kirchoff's law)

Constraint 1: steady-state assumption

If system is at steady state, fluxes into and out of each metabolite are equal.

biology



math

$$\frac{d[A]}{dt} = 0 \qquad 0 = \frac{d\mathbf{c}}{dt} = \mathbf{N} \cdot \mathbf{r}(t)$$

$$\frac{d[B]}{dt} = 0 \qquad \mathbf{N} \cdot \mathbf{r} = \mathbf{0}$$

$$\vdots$$

$$\frac{d[P]}{dt} = 0 \qquad \mathbf{r} \text{ is } \textit{null-space} \text{ of } \mathbf{N}$$

$$\mathbf{r} = \begin{bmatrix} 2 & 0 & -1 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \text{R1} \\ \text{R2} \\ \text{R3} \\ \text{R4} \\ \text{R5} \\ \text{R6} \\ \text{R7} \\ \text{R8} \\ \text{R9} \\ \text{R10} \end{matrix}$$

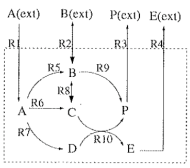
Only possible fluxes are proportional to a column of **r**, or a sum of columns.

Uses:
- identifies "dead-end" metabolites, which show network mistakes
- in computer modeling

Constraint 2: reaction direction, capacity

flux *signs* are known for irreversible reactions
flux *values* may be limited

biology



irreversible reactions
V_{max} for enzymes
max. transport rates

e.g. $r_1 > 0$

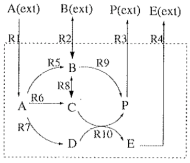
traffic analogy

traffic is limited by:
one-way roads
maximum road capacity

Constraint 2: reaction direction, capacity

flux *signs* are known for irreversible reactions
flux *values* may be limited

biology



irreversible reactions
V_{max} for enzymes
max. transport rates

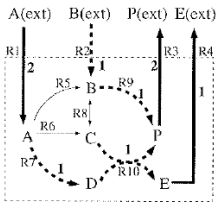
math

$$\begin{bmatrix} 0 \\ -\infty \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -\infty \\ 0 \\ 0 \end{bmatrix} \leq \mathbf{r} \leq \begin{bmatrix} 1 \\ 0 \\ \infty \\ \infty \\ \infty \\ \infty \\ \infty \\ \infty \\ \infty \\ \infty \end{bmatrix} \begin{matrix} \text{R1} \\ \text{R2} \\ \text{R3} \\ \text{R4} \\ \text{R5} \\ \text{R6} \\ \text{R7} \\ \text{R8} \\ \text{R9} \\ \text{R10} \end{matrix}$$

assumes only A is available

Constraint 3: experimental data

Flux measurements can constrain system



Measure (bold lines):
R1 = R3 = 2, R4 = 1

Infer (dashed lines):
R2 = R7 = R9 = R10 = 1

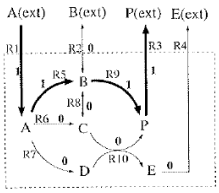
Don't know (thin lines):
R5, R6, or R8

traffic analogy

"Constraint" 4: optimization

Assume network has evolved to be "optimal"
Popular choices:
- maximum growth rate per substrate consumed
- total flux is minimized (to minimize enzyme synthesis)
- for mutants, least change from wild type

biology



We believe the network evolved to maximize P output. Thus, if it's just fed A, the fluxes must be as shown (except for R5, R6, R8 uncertainty).

traffic analogy

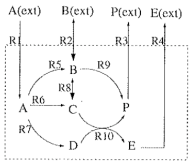
Assume everyone drives the shortest route possible.

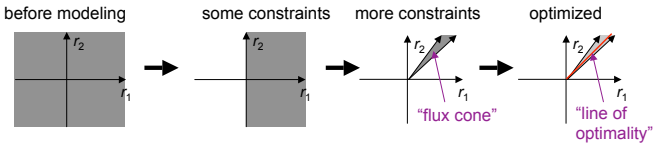
i.e. total flux is minimized

Constraint-based modeling summary

Constraints

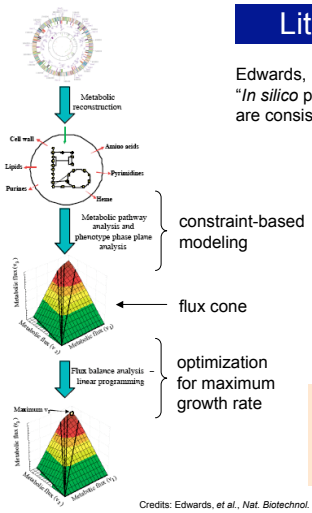
0. conservation relations
1. steady state assumption
2. max. and min. fluxes
3. experimental data
4. optimization



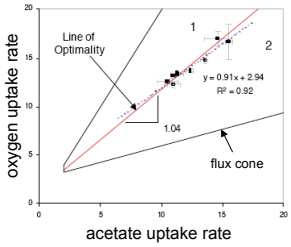


From the reaction network, and some assumptions, we can estimate most reaction fluxes.

Literature example



Edwards, Ibarra, and Palsson, 2001
"*In silico* predictions of *E. coli* metabolic capabilities are consistent with experimental data"



Result

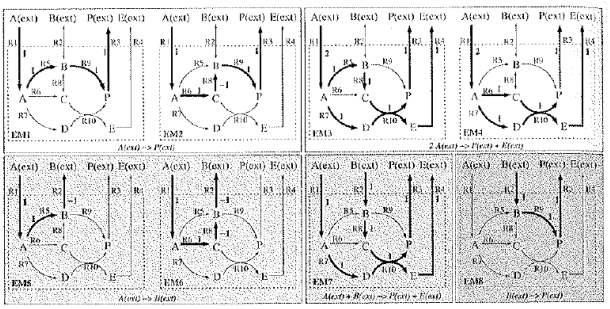
• constraint-based modeling and optimization based on growth rate yields fluxes that agree with experiment

Credits: Edwards, *et al.*, *Nat. Biotechnol.* 19:125, 2001.

Pathways

Elementary flux mode

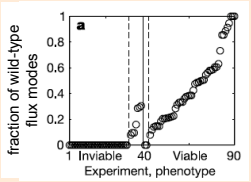
a path through the network that cannot be simplified (and obeys constraints like steady-state, reaction directionality, etc.)



Pathway applications

Removing all essential pathways leads to inviability
• helpful for understanding mutants
• good for designing drug targets

Stelling et al. (2002) showed high correlation between fraction of flux modes available in mutants and viability. No flux modes implied inviable.



Pathways help build intuition
• in all biochemistry texts

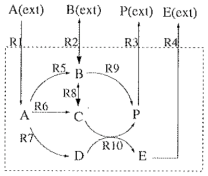
For further analysis
• The minimal set of elementary flux modes are the "eigenvectors" of the network

Credit: Stelling *et al.* Nature 420:190, 2002.

Metabolic control analysis

Metabolic control analysis is sensitivity analysis of the reaction network. Same constraints (steady-state, reaction directionality, etc.)

biology



we want to make E from A,
will doubling enzyme 10 help?
what about knocking out enzyme 9?

traffic analogy

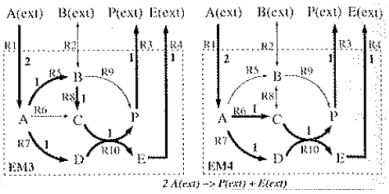
if Mercer Street is widened, will that fix congestion? Or just move it to the next traffic light?

Credit: http://www.cityofseattle.net/transportation/ppmp_mercer.htm

Metabolic control analysis

Common misconception
- There is one rate-limiting step

Truth
- Lots of reactions affect total production rate
  - upstream reactions in pathway
  - downstream reactions due to product inhibition



Math

Flux control coefficient is effect of enzyme amount on flux:

$$C_{R10}^{R4} = \frac{\partial \ln r_4}{\partial \ln [E_{10}]}$$

r_4 is flux in reaction R4, $[E_{10}]$ is enzyme amount in R10.

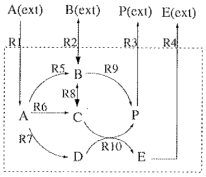
Flux control coefficients are usually between 0 and 1:
$C_{R10}^{R4} = 0$ R10 has no effect
$C_{R10}^{R4} = 1$ R10 is rate-limiting

Typical flux control coefficients are 0 to 0.5, with *several enzymes sharing most of the control on any flux.*

Metabolic control analysis - substrate conc.

What happens if an external metabolite concentration changes?

biology



we want to make E from A,
should we increase [A]?
does [P] matter?

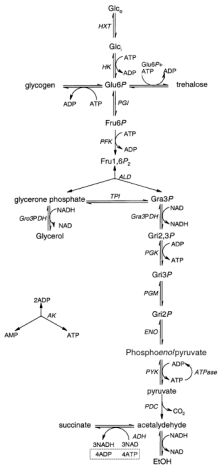
Response coefficient: $R_{[A(ext)]}^{R4} = \frac{\partial \ln r_4}{\partial \ln [A]}$

Found by summing control coefficients and "enzyme elasticities" for each enzyme

traffic analogy

how will traffic change after a football game ends?

Metabolic modeling example



Pritchard and Kell (2002) investigated flux control in yeast glycolysis

They used Gepasi (predecessor to Copasi). This is a Copasi example file: YeastGlycolysis.cps

In Copasi, it's easy to find:
- stoichiometric matrix
- constraint 0: mass conservation
- steady-state concentrations
- elementary flux modes
- metabolic control analysis coefficients

Credit: Pritchard, Kell, *Eur. J. Biochem.* 269:3894, 2002.

Summary

- Big networks
- Useful databases for metabolism
  - KEGG, BRENDA
- Stoichiometric matrix
  - math representation of network
- Constraint-based modeling
  - 0. mass conservation
  - 1. steady-state assumption
  - 2. reaction min. & max. fluxes
  - 3. experimental data
  - 4. optimize
- Metabolic Control Analysis
  - sensitivity of fluxes to parameters
- Simulation tools
  - Copasi does many of these tasks

Homework

Next week's class is on gene regulatory networks.

Class will be in room B1-072/074

Read

Milo, Shen-Orr, Itzkovitz, Kashtan, Chklovskii, and Alon, "Network motifs: Simple building blocks of complex networks" *Science* 298:824, 2002.