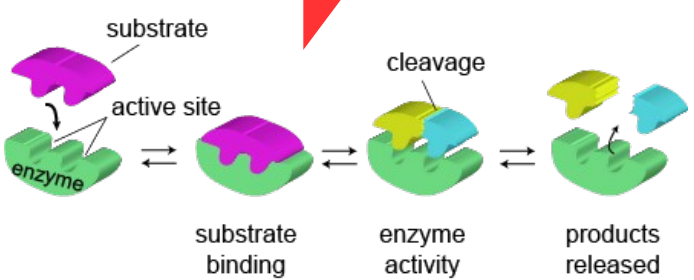
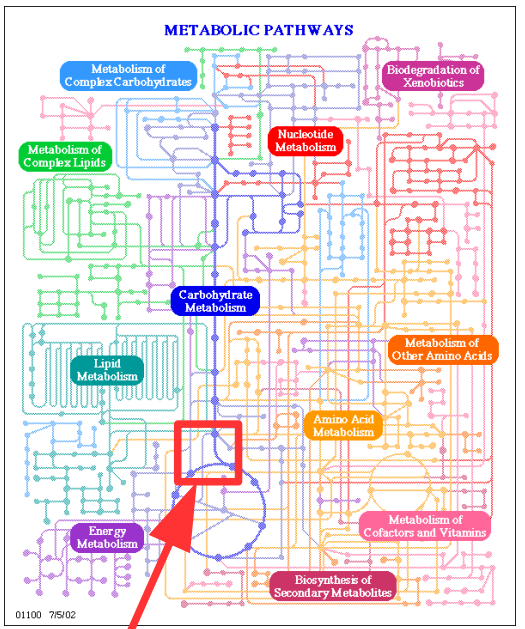

CONSTRAINT-BASED MODELS

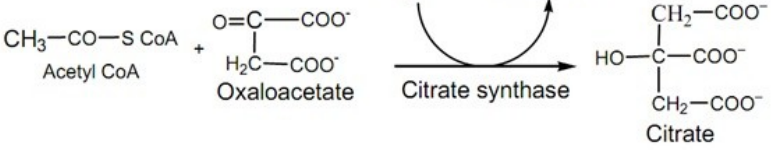
Outline

- Stoichiometric network reconstruction
- Flux-balance analysis
- Optimization using Linear Programming

Stoichiometric network reconstruction

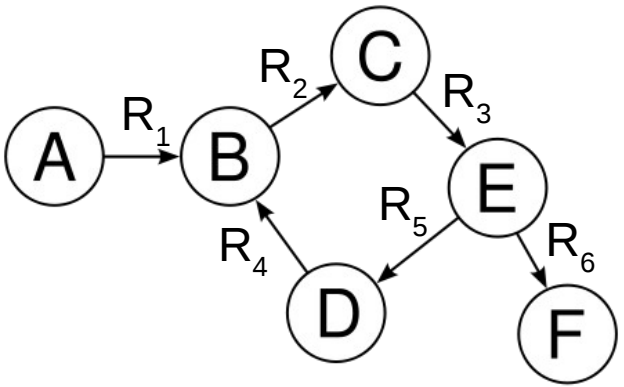




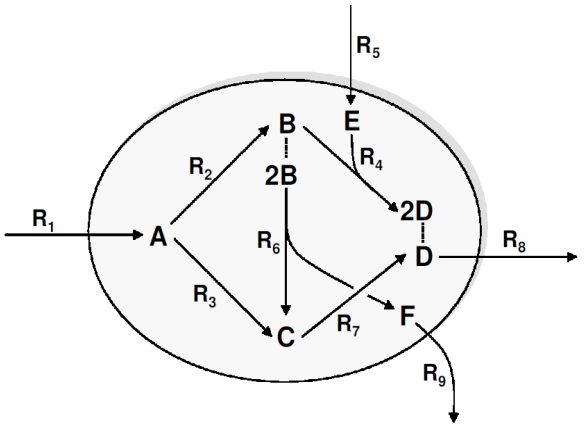


Metabolic network are not simple graphs

simple directed graph



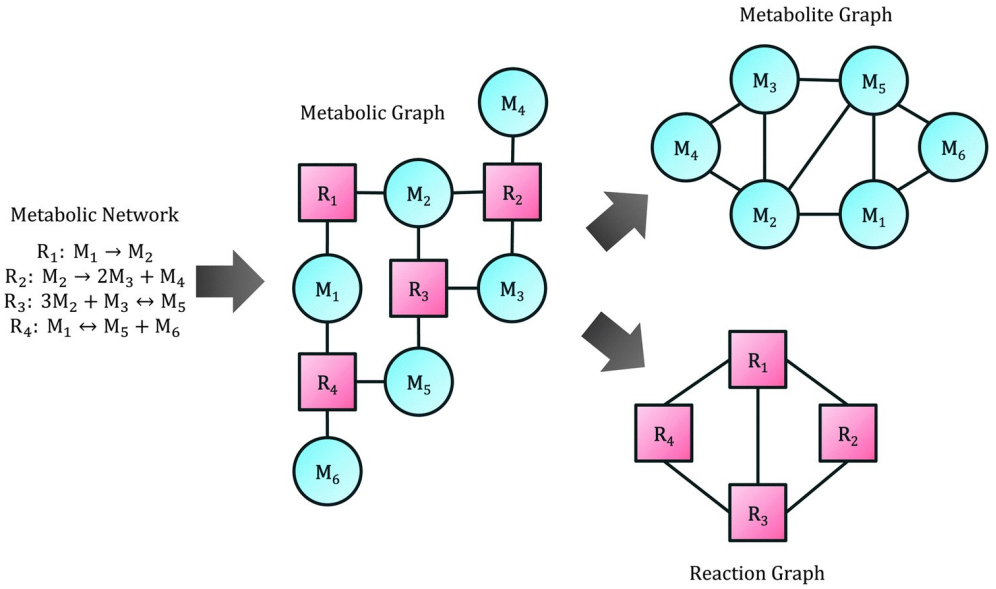
metabolic graph



- Two types of entities (metabolites, enzymes)
- Possible representations:
  - hypergraph - enzymes are hyper-edges (not always 1:1)
  - bipartite graph - enzymes are “special” nodes
- Stoichiometry (encoded as weights in a bipartite graph)

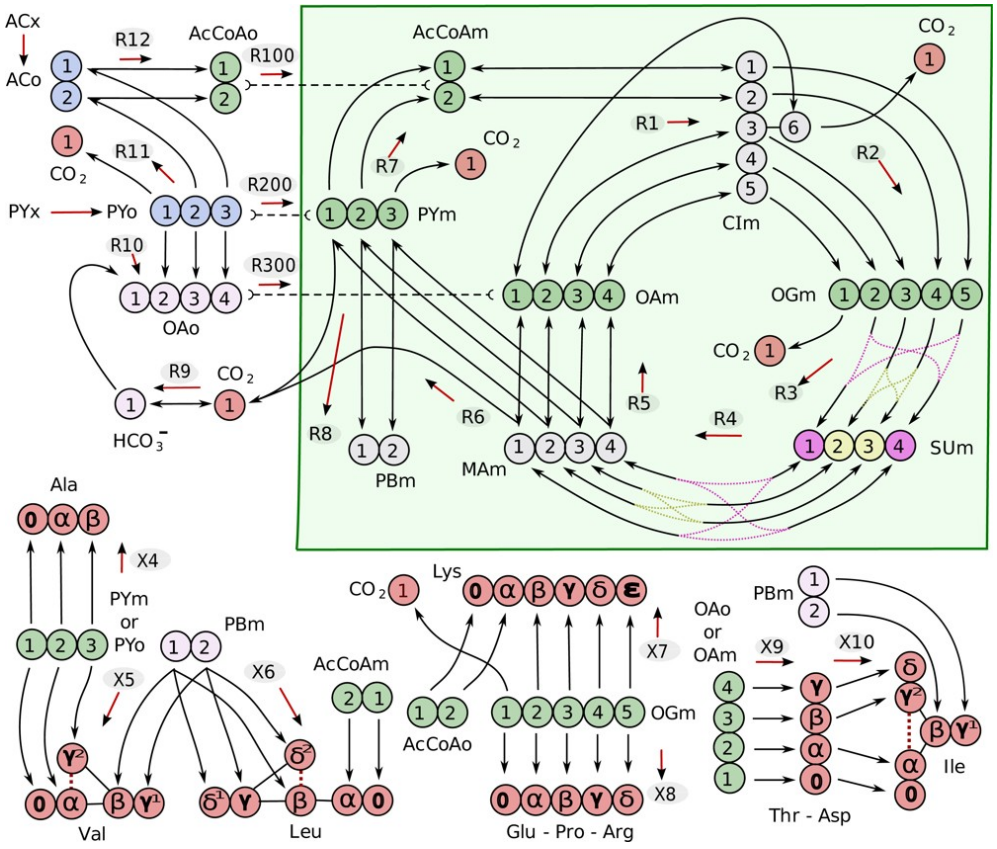
Measuring distance in a metabolic network

- Standard* definition of distance in a bipartite graph



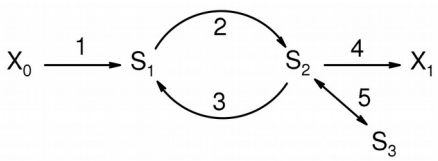
* Ignore currency metabolites: H2O, H^+, CO_2, P_i, PP_i, NH_4^+, ATP, ADP, AMP, NAD(P)(H)

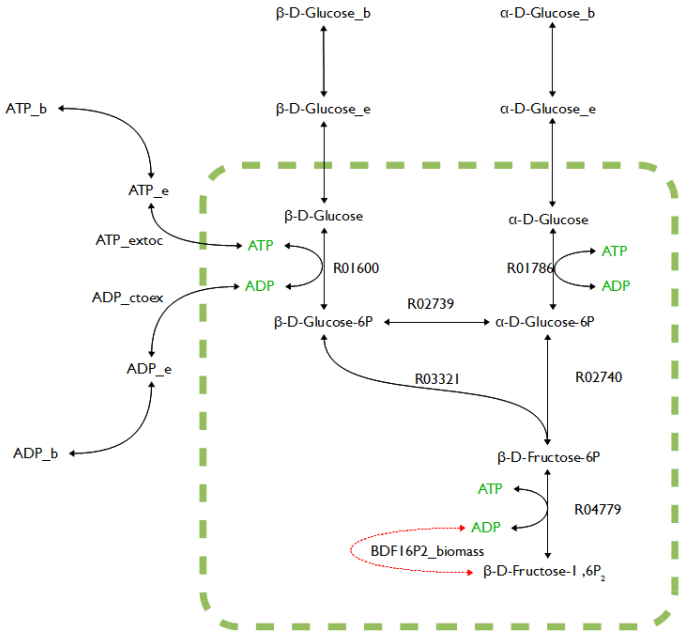
Higher level of detail: atom-mapping

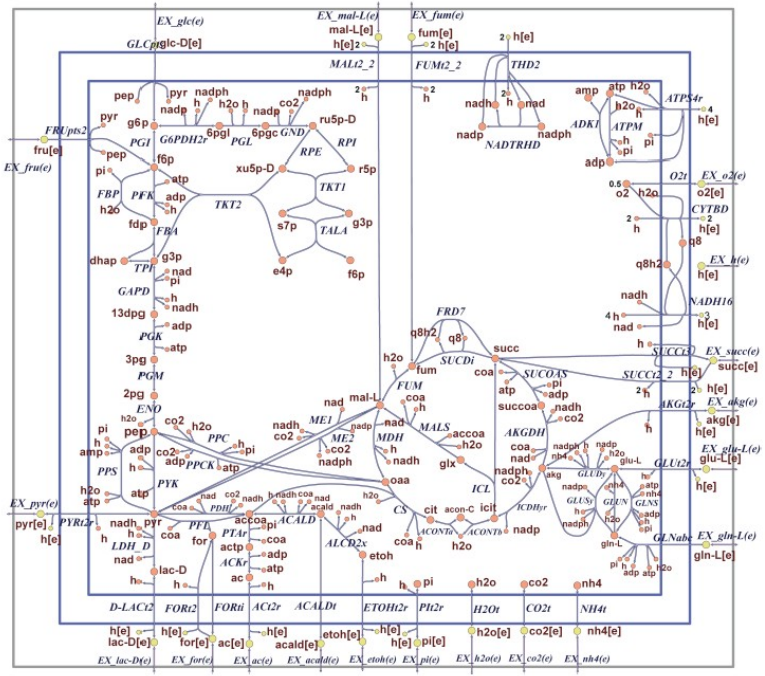


Schryer, David W., et al. "Bidirectionality and compartmentation of metabolic fluxes are revealed in the dynamics of isotopomer networks." International journal of molecular sciences 10.4 (2009): 1697-1718.

Metabolic networks: from small scale to genome scale





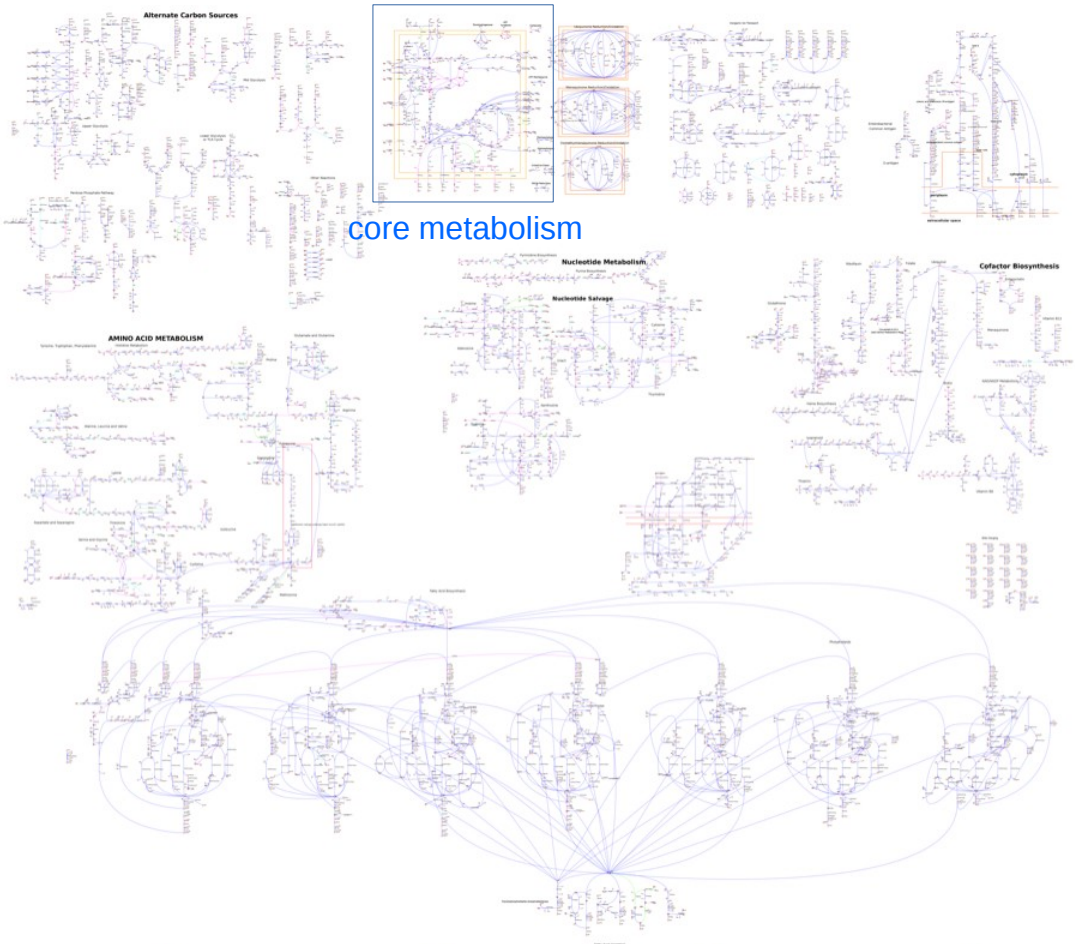


toy model

hexose transport model

core metabolism model

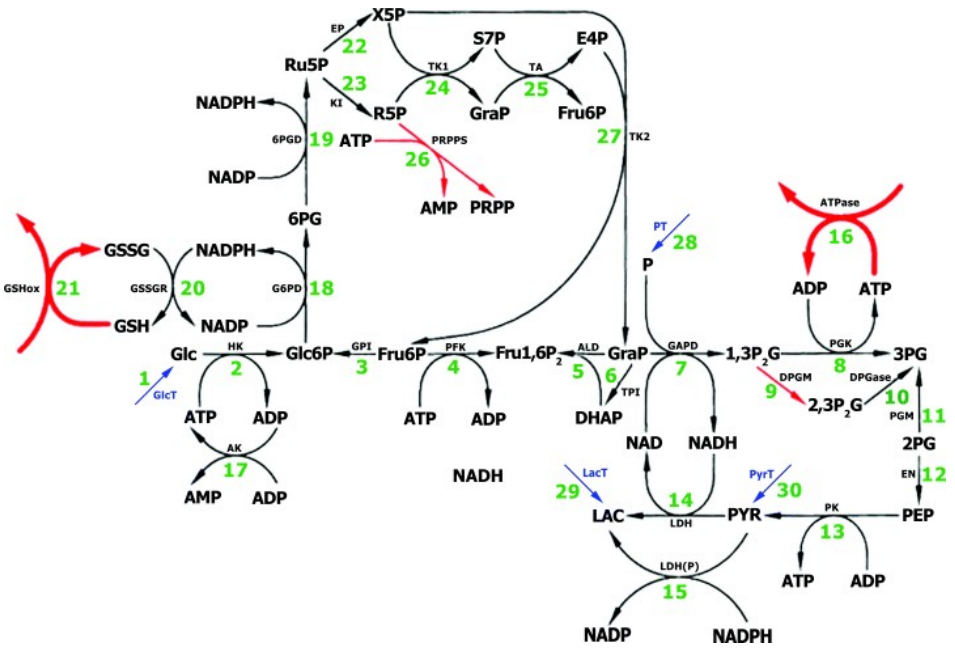
Genome scale model (*E. coli*)



Model	Year	Reactions	Metabolites
iJE660	2000	627	438
iJR904	2003	931	625
iAF1260	2007	1260	1039
iJO1366	2011	2077	1136
core model	2007	95	72

http://systemsbiology.ucsd.edu/InSilicoOrganisms/Ecoli/EcoliSBML

Realistic example: erythrocyte metabolism



metabolites

reactions

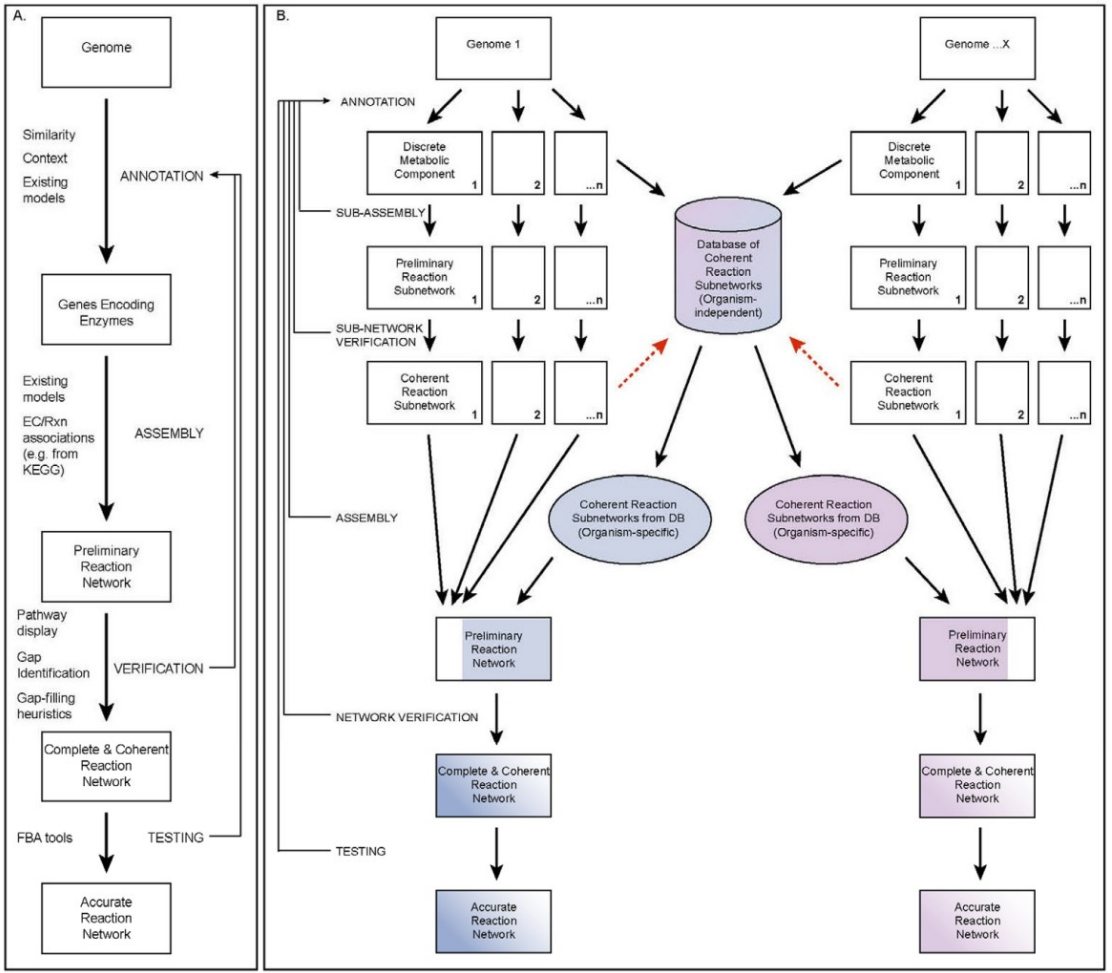
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
		Glct	HK	GPI	PFK	ALD	TPI	GAPDH	PGK	DPGM	DPGase	PGM	EN	PK	LDH	LDH(P)	ATPase	AK	G6PD	6PGD	GSSGR	GSHox	EP	KI	TK1	TA	PRPPS	TK2	Pt	Lact	Pyrt
1	Glcin	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Glc6P	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
3	Fru6P	0	0	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
4	Fru16P2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	GraP	0	0	0	0	-1	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	1	0	0	0
6	DHAP	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	13P2Gri	0	0	0	0	0	0	-1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	23P2Gri	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	3PGri	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	2PGri	0	0	0	0	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	PEP	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	ATP	0	-1	0	-1	0	0	0	1	0	0	0	0	1	0	0	-1	1	0	0	0	0	0	0	0	0	-1	0	0	0	0
13	ADP	0	1	0	1	0	0	0	-1	0	0	0	0	-1	0	0	1	-2	0	0	0	0	0	0	0	0	0	0	0	0	0
14	6PGlcA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0	0
15	NADP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	-1	-1	1	0	0	0	0	0	0	0	0	0	0
16	GSH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	-2	0	0	0	0	0	0	0	0	0
17	Rul5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	-1	-1	0	0	0	0	0	0	0
18	Xul5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	-1	0	0	0
19	Rib5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	-1	0	0	0	0
20	Sed7P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	0
21	E4P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	0
22	NAD	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	Pi	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
24	Lac	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
25	Pyr	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

H.G. Holzhütter (2004)
The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks
European Journal of Biochemistry / FEBS, 271(14), 2905–22
doi:10.1111/j.1432-1033.2004.04213.x

Automated genomic reconstructions

- www.theseed.org

Goal:
Annotating 1000 genomes and reconstructing the metabolic networks

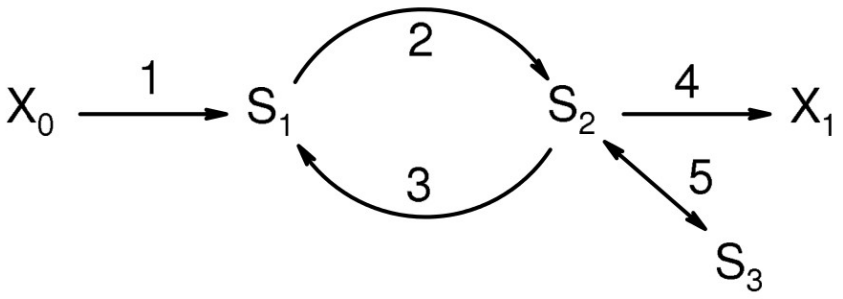


DeJongh et al.
BMC Bioinfo
2007

Uses for whole-cell stoichiometric models

- Framing a linear problem:
  - Simulating exponential growth by assuming steady-state:
    - i.e. all internal metabolites have a constant concentration
  - External fluxes are measured (or at least bounded)
  - Some reactions are considered to be irreversible
- Together, the flux solution space is constrained enough to answer some questions, e.g.
  - What is the maximal possible growth rate?
  - Which reactions are essential for growth?
  - Which external conditions can support growth (anaerobic, carbon sources, nitrogen sources, etc.)?

Metabolic network representations



Graph

$$\begin{aligned}\frac{dS_1}{dt} &= v_1 - v_2 + v_3\\ \frac{dS_2}{dt} &= v_2 - v_3 - v_4 - v_5\\ \frac{dS_3}{dt} &= v_5\end{aligned}$$

Kinetic model (ODE)

reactions

metabolites

$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \end{array} \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Stoichiometric matrix

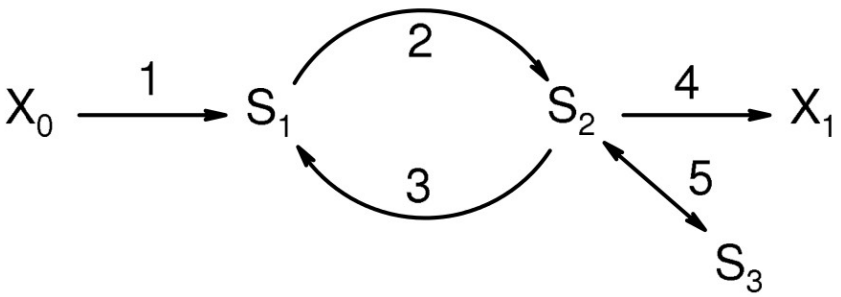
$$\begin{bmatrix} \frac{dS_1}{dt} \\ \frac{dS_2}{dt} \\ \frac{dS_3}{dt} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}$$

$$\dot{S} = N \cdot v$$

System of differential equations

Assumption of steady-state

- A kinetic model would have to contain rate laws for each reaction



$$\frac{dS_1}{dt} = v_1 - v_2 + v_3$$
$$\frac{dS_2}{dt} = v_2 - v_3 - v_4 - v_5$$
$$\frac{dS_3}{dt} = v_5$$

$$v_1 = V_{max,1}\frac{X_0}{K_1 + X_0}$$

$$v_2 = V_{max,2}\frac{S_1}{K_2 + S_1}$$

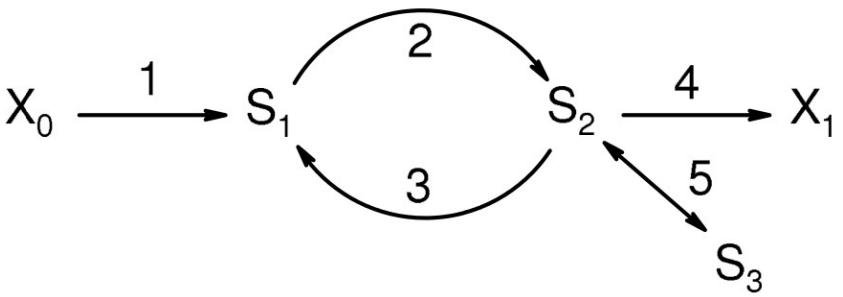
$$v_3 = V_{max,3}\frac{S_2}{K_3 + S_2}$$

$$v_4 = V_{max,4}\frac{S_2}{K_4 + S_2}$$

$$v_5 = E_5\frac{k_{f,5}S_2/K_{f,5} - k_{b,5}S_3/K_{b,5}}{1 + S_2/K_{f,5} + S_3/K_{b,5}}$$

Assumption of steady-state

- A kinetic model would have to contain rate laws for each reaction
- However, one can also do quite a lot without it:



$$\frac{dS_1}{dt} = v_1 - v_2 + v_3 \quad = 0$$

$$\frac{dS_2}{dt} = v_2 - v_3 - v_4 - v_5 \quad = 0$$

$$\frac{dS_3}{dt} = v_5 \quad = 0$$

$$v_1 = V_{max,1}\frac{X_0}{K_1 + X_0}$$

$$v_2 = V_{max,2}\frac{S_1}{K_2 + S_1}$$

$$v_3 = V_{max,3}\frac{S_2}{K_3 + S_2}$$

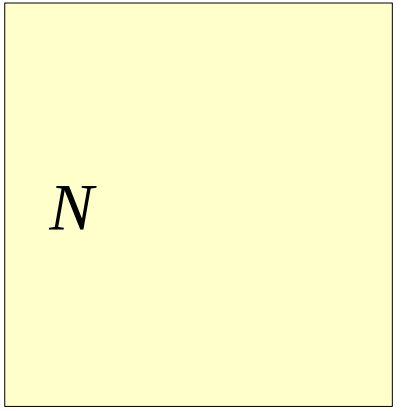
$$v_4 = V_{max,4}\frac{S_2}{K_4 + S_2}$$

$$v_5 = E_5\frac{k_{f,5}S_2/K_{f,5} - k_{b,5}S_3/K_{b,5}}{1 + S_2/K_{f,5} + S_3/K_{b,5}}$$

What is flux balance analysis (FBA)?

- First, define the scope: define a stoichiometric network ***N***

$$\begin{matrix} S_1 \\ S_2 \\ S_3 \end{matrix} \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$



What is flux balance analysis (FBA)?

- First, define the scope: define a stoichiometric network ***N***
- Apply a **steady state assumption***, i.e. all internal metabolite concentrations are constant (mass balance)

* In realistic models, ***N*** has more reactions than metabolites, which means that this system of linear equations is under-determined

$$\frac{dS_1}{dt} = v_1 - v_2 + v_3 = 0$$
$$\frac{dS_2}{dt} = v_2 - v_3 - v_4 - v_5 = 0$$
$$\frac{dS_3}{dt} = v_5 = 0$$

$$N \cdot v = 0$$

What is flux balance analysis (FBA)?

- First, define the scope: define a stoichiometric network ***N***
- Apply a **steady state assumption**, i.e. all internal metabolite concentrations are constant (mass balance)
- Add individual constraints* for each reaction flux

* there is still usually a large solution space

$$N \cdot v = 0$$
$$a_i \leqslant v_i \leqslant b_i$$

What is flux balance analysis (FBA)?

- First, define the scope: define a stoichiometric network ***N***
- Apply a **steady state assumption**, i.e. all internal metabolite concentrations are constant (mass balance)
- Add individual constraints for each reaction flux
- Maximize an objective function, typically biomass production rate

c·v – a linear combination of biomass precursor synthesis fluxes

$$\max_{v} \quad c \cdot v$$

$$N \cdot v = 0$$

$$a_i \leqslant v_i \leqslant b_i$$

What is flux balance analysis (FBA)?

- First, define the scope: define a stoichiometric network ***N***
- Apply a **steady state assumption**, i.e. all internal metabolite concentrations are constant (mass balance)
- Add individual constraints for each reaction flux
- Maximize an objective function, typically biomass production rate

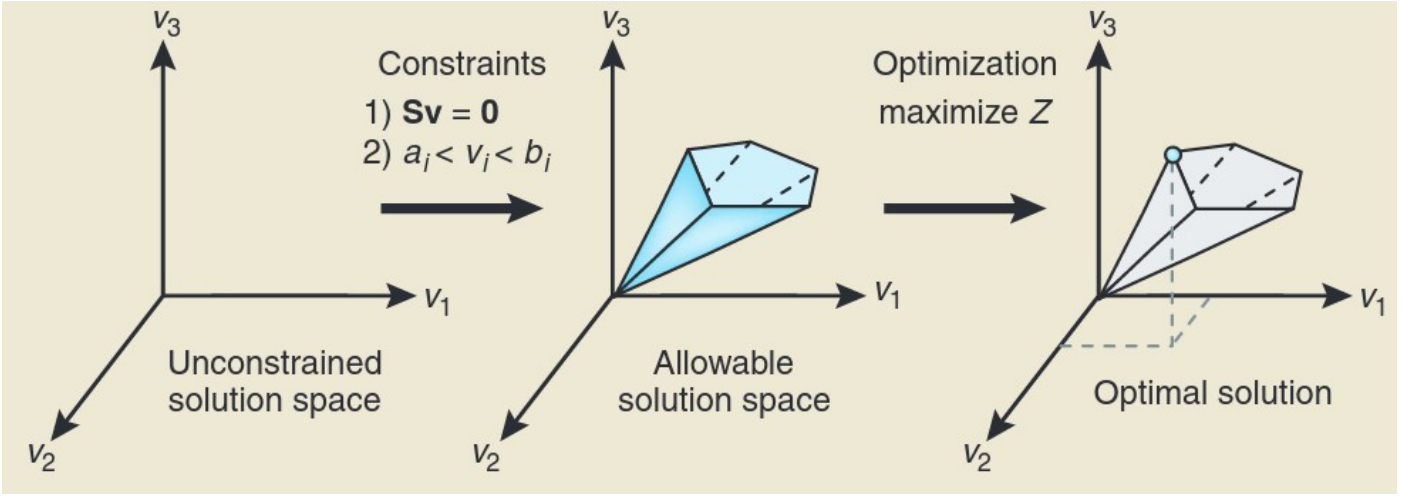
$$\max_{v} \quad c \cdot v$$

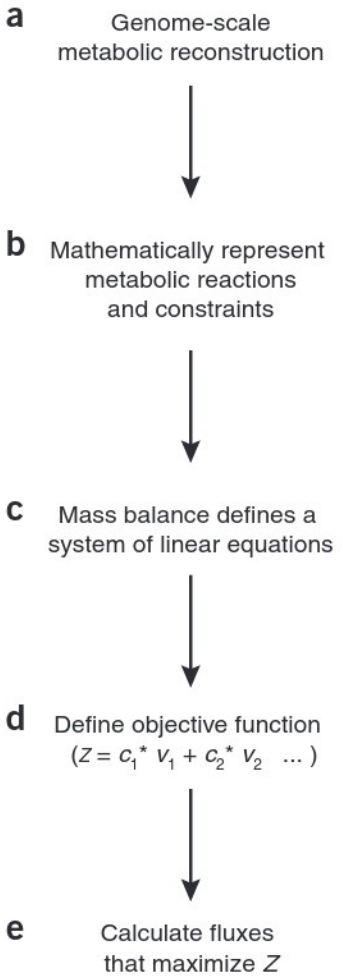
$$N \cdot v = 0$$

$$a_i \leqslant v_i \leqslant b_i$$

Solve using linear programming

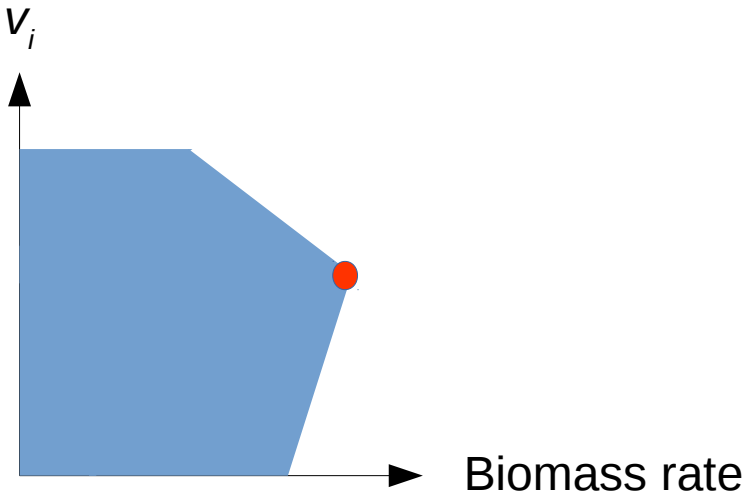
The conceptual basis of constraint-based models





Uses of FBA

- What is the maximal possible yield in different conditions?
- Which media can support growth at all?
- What is the effect of a single gene knockout (for enzymes)?
- Drawing a phenotypic phase plane (PPP)

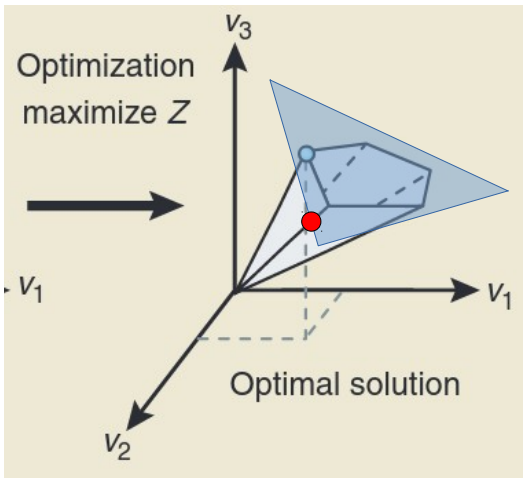


Additional constraints

- ATP maintenance
- Loopless-COBRA
- Thermodynamic constraints (TMFA)
- FBAwMC (with Molecular Crowding)
- Flux minimization

FBA with Molecular Crowding

- Catalyzing a reaction at a certain rate requires a minimal concentration of enzyme
- Physiological constraints on protein concentrations put an upper bound on the sum of all enzyme concentrations:



$$\Sigma\, \alpha_i\, E_i \leq V/M$$

enzyme *i* specific volume

enzyme *i* concentration

total cell dry weight

total cell volume

Beg, Q. K., Vazquez, A., Ernst, J., de Menezes, M. A., Bar-Joseph, Z., Barabási, A.-L. A.-L., & Oltvai, Z. N. (2007)
Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity
PNAS
104(31), 12663–8. doi:10.1073/pnas.0609845104

Principle of Flux Minimization

- Sometimes called parsimonious FBA (pFBA)
- Rather than maximizing the biomass flux, minimize the sum of all fluxes*

$$\underset{v}{minimize} \quad \sum_{j=1}^{r} |v_j|$$

* Some people use the sum of squared fluxes

$$\underset{v}{minimize} \quad \sum_{j=1}^{r} |v_j|^2$$

H.G. Holzhütter (2004)
The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks
European Journal of Biochemistry / FEBS, 271(14), 2905–22
doi:10.1111/j.1432-1033.2004.04213.x

Loopless-COBRA and Thermodynamic constrains

- Loopless only avoids futile cycles which are thermodynamically infeasible
- TMFA (Thermodynamic Metabolic Flux Analysis) applies the second law of thermodynamics on all active reactions: $\Delta G < 0$
- Add new variables to represent the metabolite concentrations $ln(x_i)$
- Add constraints, $\Delta G = \Delta G^{\circ} - RT \Sigma ln(x_i) + RT \Sigma ln(x_j)$
  - where *i* are the substrates and *j* are the products
- Add constraints, for each active reaction, the ΔG must be negative

Henry, C. S., Broadbelt, L. J., & Hatzimanikatis, V. (2007)
Thermodynamics-Based Metabolic Flux Analysis
Biophysical Journal, 92(5), 1792–1805
doi:10.1529/biophysj.106.093138

Alternative optimization goals

- Instead of biomass yield:
  - Minimize glucose uptake rate
  - Maximize ATP production rate
  - Minimal sum of fluxes (l_0, l_1 or l_2 norms)
- Related to genetic manipulations:
  - MoMA (Minimal Metabolic Adjustment)
  - OptKnock

Alternative optimization goals

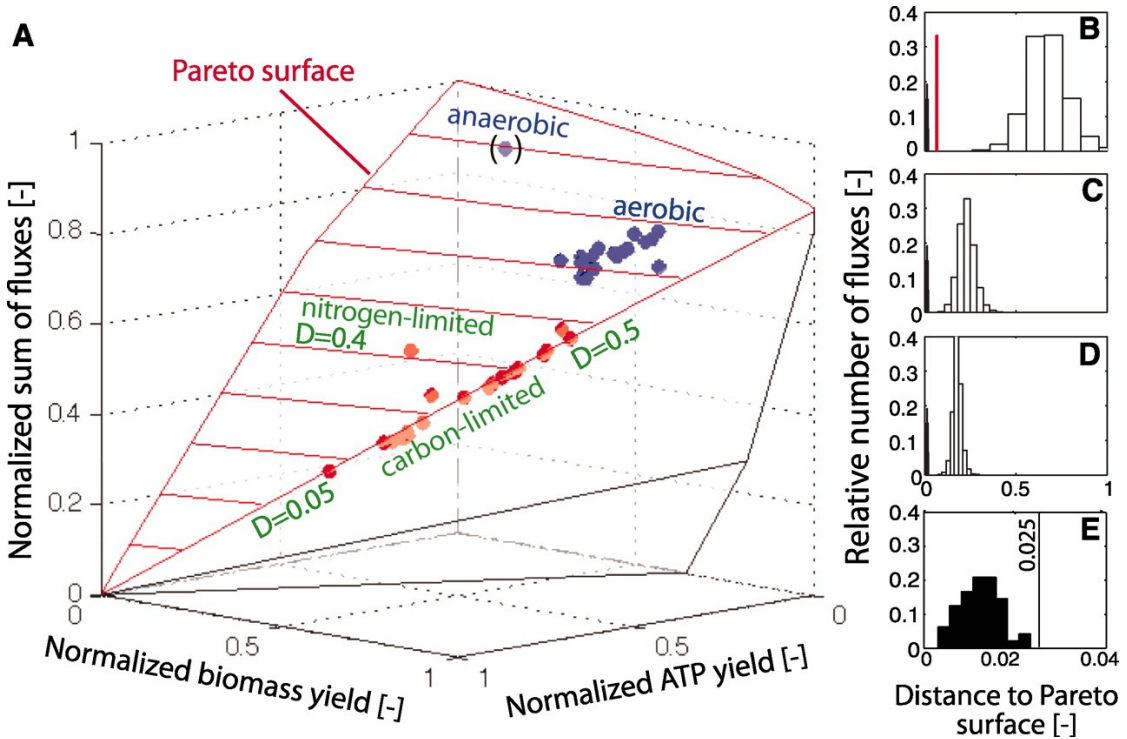
both perform best in nutrient-limited continuous cultures

performs best in unlimited aerobic growth on glucose

Max biomass[b]	$\max \frac{v_{biomass}}{v_{glucose}}$	Maximization of biomass yield
Max ATP	$\max \frac{v_{ATP}}{v_{glucose}}$	Maximization of ATP yield
$\text{Min}\sum v^{2c}$	$\min \sum_{i=1}^{n} v_i^2$	Minimization of the overall intracellular flux
Max ATP per flux unit[c]	$\max \frac{v_{ATP}}{\sum_{i=1}^{n} v_i^2}$	Maximization of ATP yield per flux unit
Max biomass per flux unit[c]	$\max \frac{v_{biomass}}{\sum_{i=1}^{n} v_i^2}$	Maximization of biomass yield per flux unit
Min glucose	$\min \frac{v_{glucose}}{v_{biomass}}$	Minimization of glucose consumption
Min reaction steps[c]	$\min \sum_{i=1}^{n} y_i^2, y_i \in \{0, 1\}$	Minimization of reaction steps
Max ATP per reaction step[c]	$\min \frac{v_{ATP}}{\sum_{i=1}^{n} y_i^2}, y_i \in \{0, 1\}$	Maximization of ATP yield per reaction step
Min redox potential[d,e]	$\min \frac{\sum_{n} v_{NADH}}{v_{glucose}}$	Minimization of redox potential[f]
Min ATP production[d,e]	$\min \frac{\sum_{n} v_{ATP}}{v_{glucose}}$	Minimization of ATP producing fluxes[g]
Max ATP production[d,e]	$\max \frac{\sum_{n} v_{ATP}}{v_{glucose}}$	Maximization of ATP producing fluxes[h]

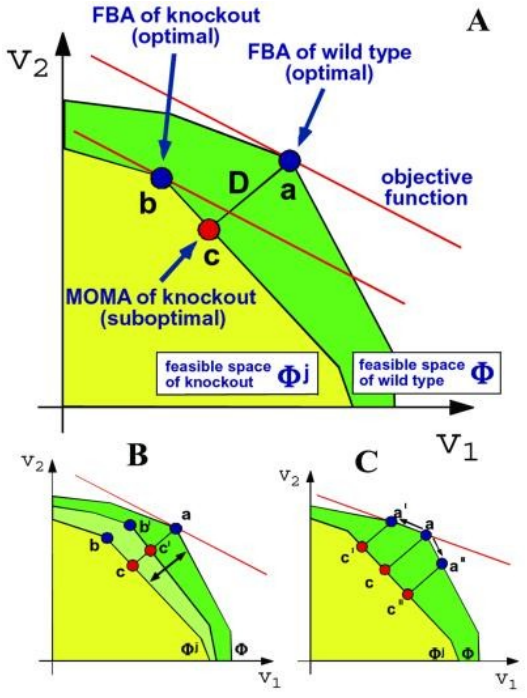
Schuetz, R., Kuepfer, L., & Sauer, U. (2007)
Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli
Molecular Systems Biology, 3(119), 119
doi:10.1038/msb4100162

Alternative optimization goals



Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., & Sauer, U. (2012)
Multidimensional Optimality of Microbial Metabolism
Science, 336(6081), 601–604
doi:10.1126/science.1216882

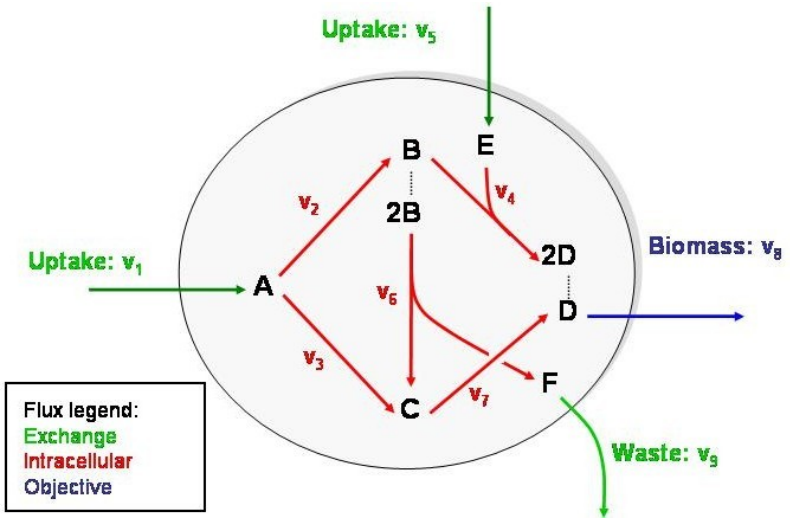
MoMA - Minimal Metabolic Adjustment



- Idea: we cannot assume optimality for knockout strains
- Assumption: fluxes in mutants would be closest to original optimum, within the feasible space
- Implementation: minimize the l_2 distance between knockout solution and wild-type optimum

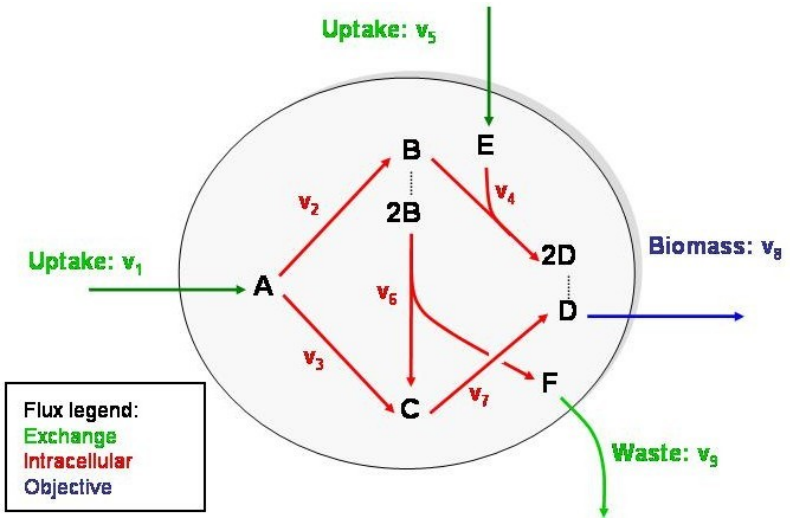
Metabolic engineering using OptKnock

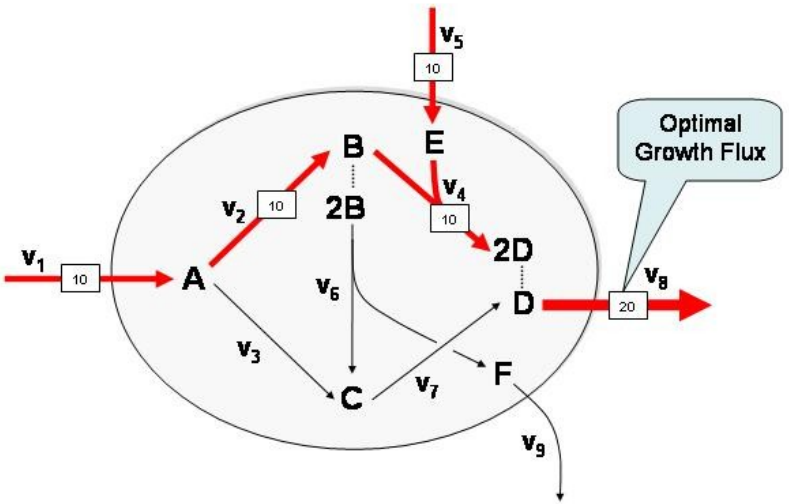
- Question: which genetic knockouts should one do to maximize production of byproduct (e.g. ethanol)?



Metabolic engineering using OptKnock

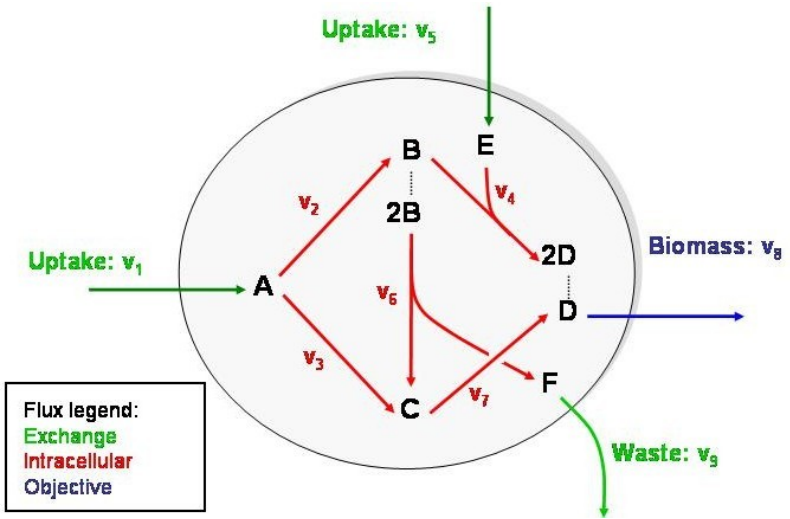
- Question: which genetic knockouts should one do to maximize production of byproduct (e.g. ethanol)?

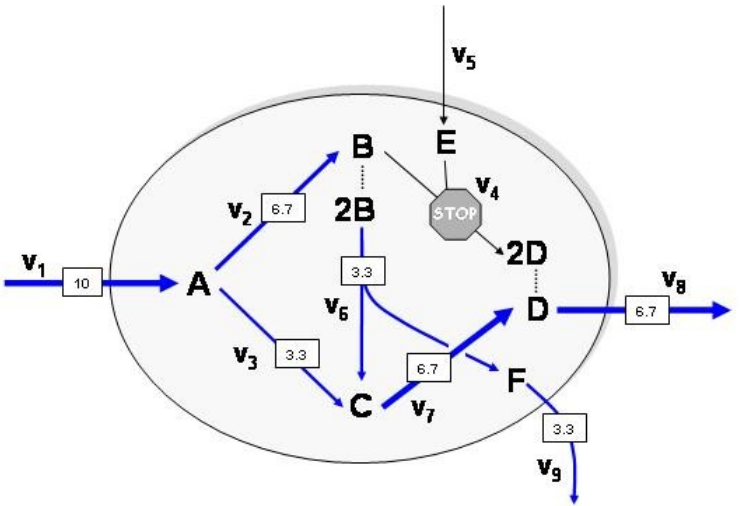




Metabolic engineering using OptKnock

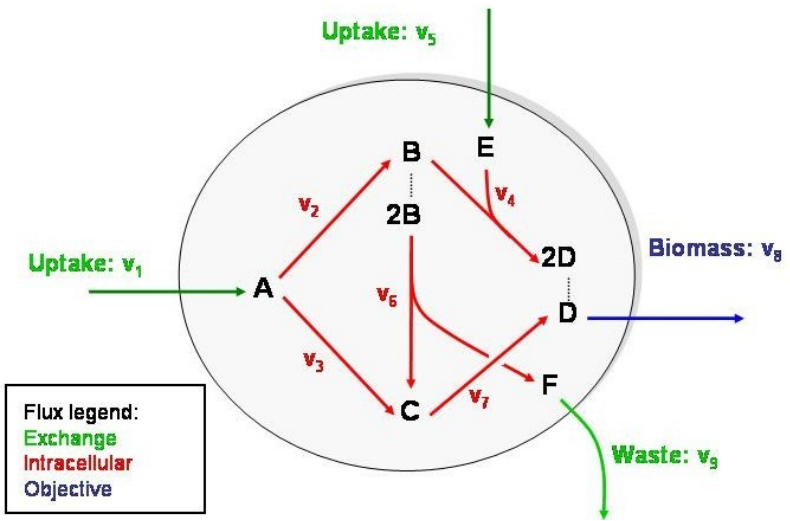
- Question: which genetic knockouts should one do to maximize production of byproduct (e.g. ethanol)?

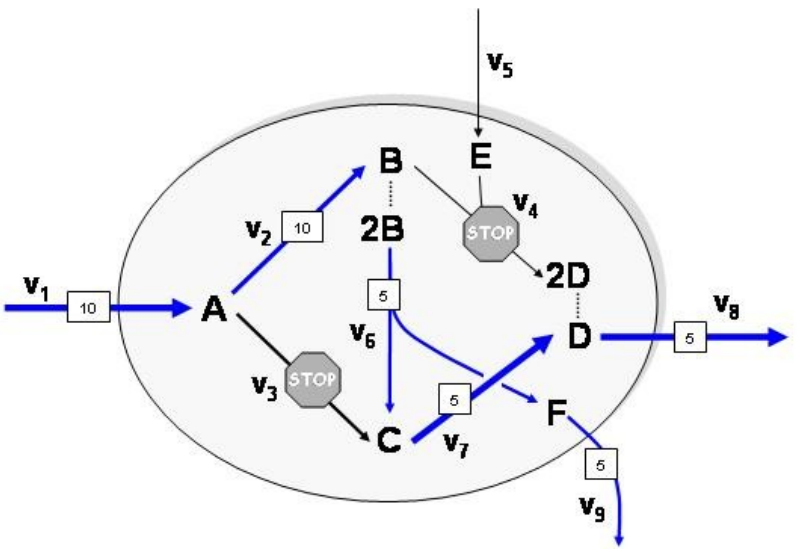




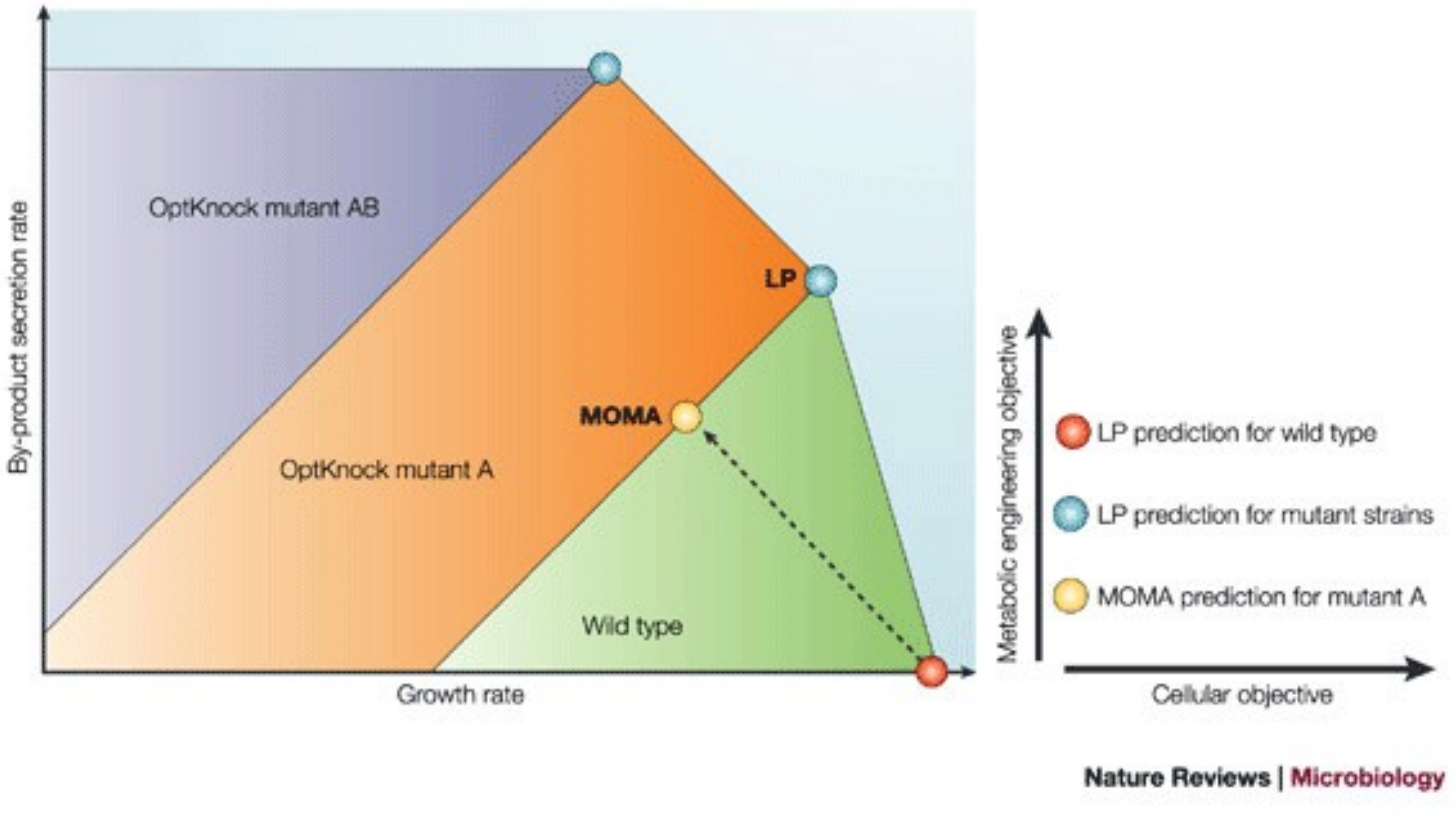
Metabolic engineering using OptKnock

- Question: which genetic knockouts should one do to maximize production of byproduct (e.g. ethanol)?





Metabolic engineering using OptKnock



What is at the boundary of FBA?

- Metabolism doesn't end in precursor metabolites
- All following processes are “lumped” into one step called the biomass function:
  - transcription
  - translation
  - protein modification / assembly / trafficking
  - DNA replication
  - membrane assembly / division
  - macromolecule degradation
- What can we gain by extending the model to encompass the entire cell?