269020 VO Computational Concepts in Biology II
Part2: Biological Networks – Dynamic Aspects

Christoph Flamm[†] and Stefanie Widder[‡]

[†]Institute for Theoretical Chemistry, University of Vienna
[‡]Infection Biology, Medical University of Vienna

xtof@tbi.univie.ac.at
stefanie.widder@medunwien.ac.at

SS2022

Reconstruction of Metabolic Networks

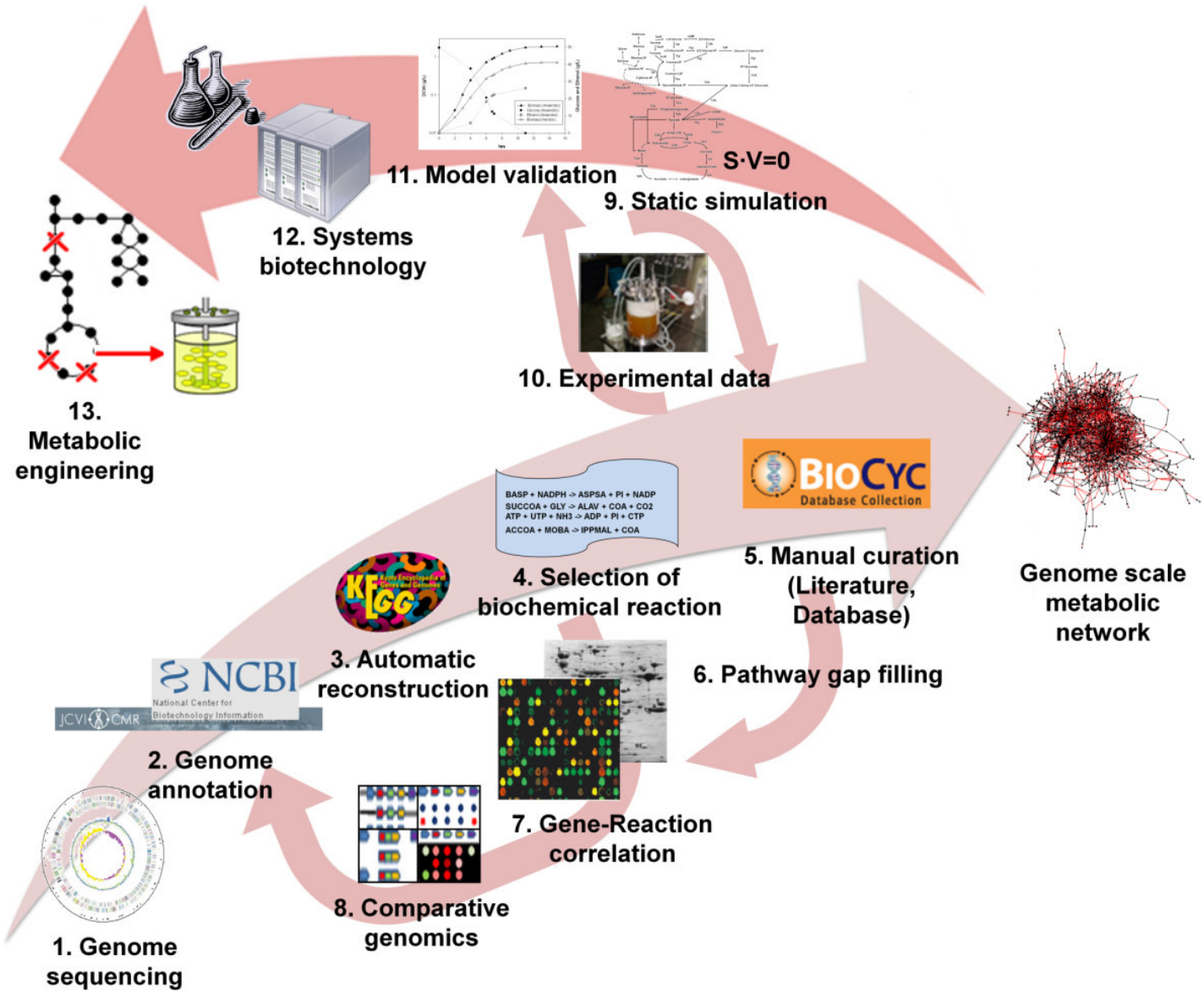
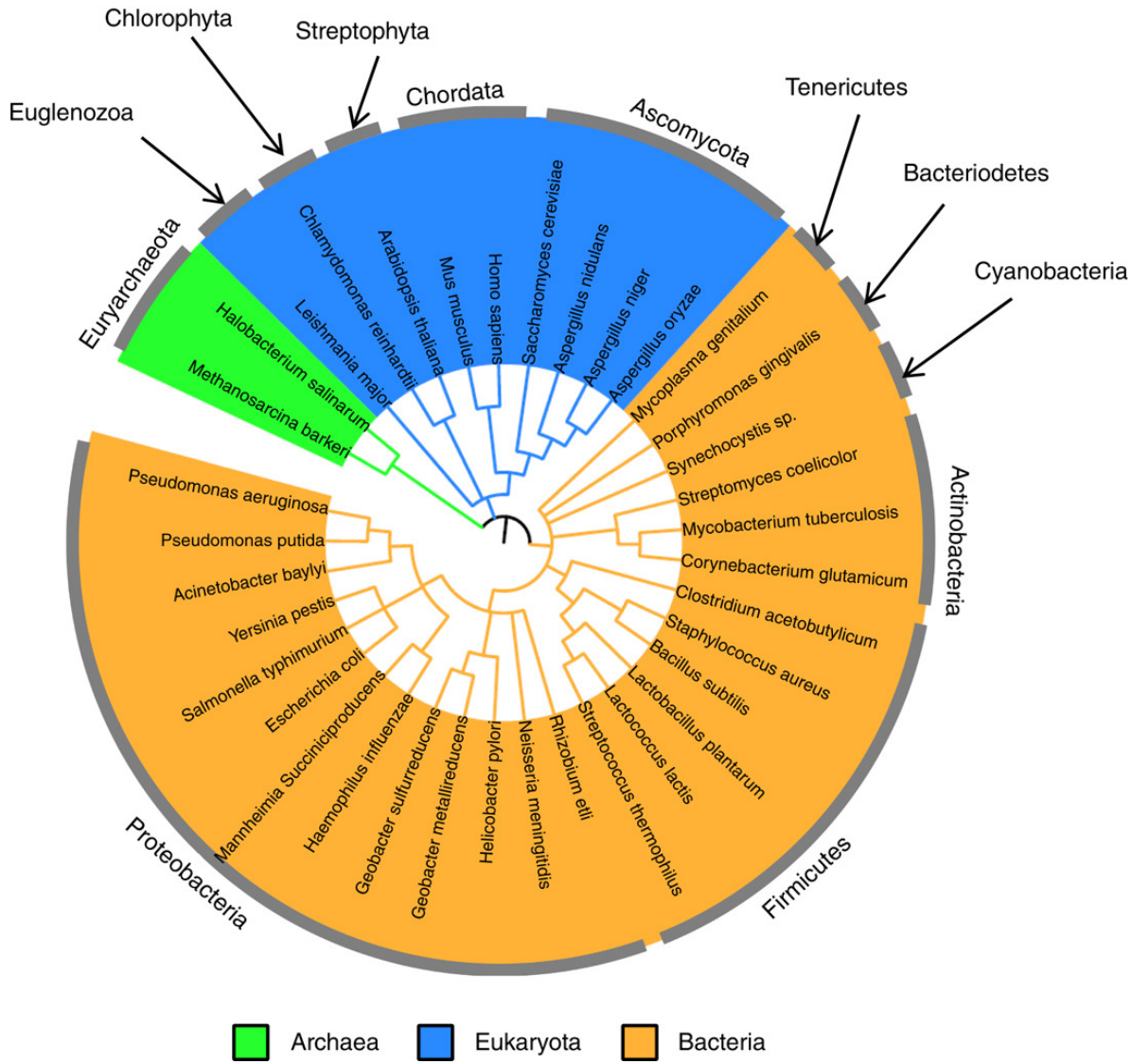


Figure from Lee KY et al (2010), The genome-scale metabolic network analysis of Zymomonas mobilis ZM4 explains physiological features and suggests ethanol and succinic acid production strategies, *Microbial Cell Factories* **9**:94 | doi:10.1186/1475-2859-9-94

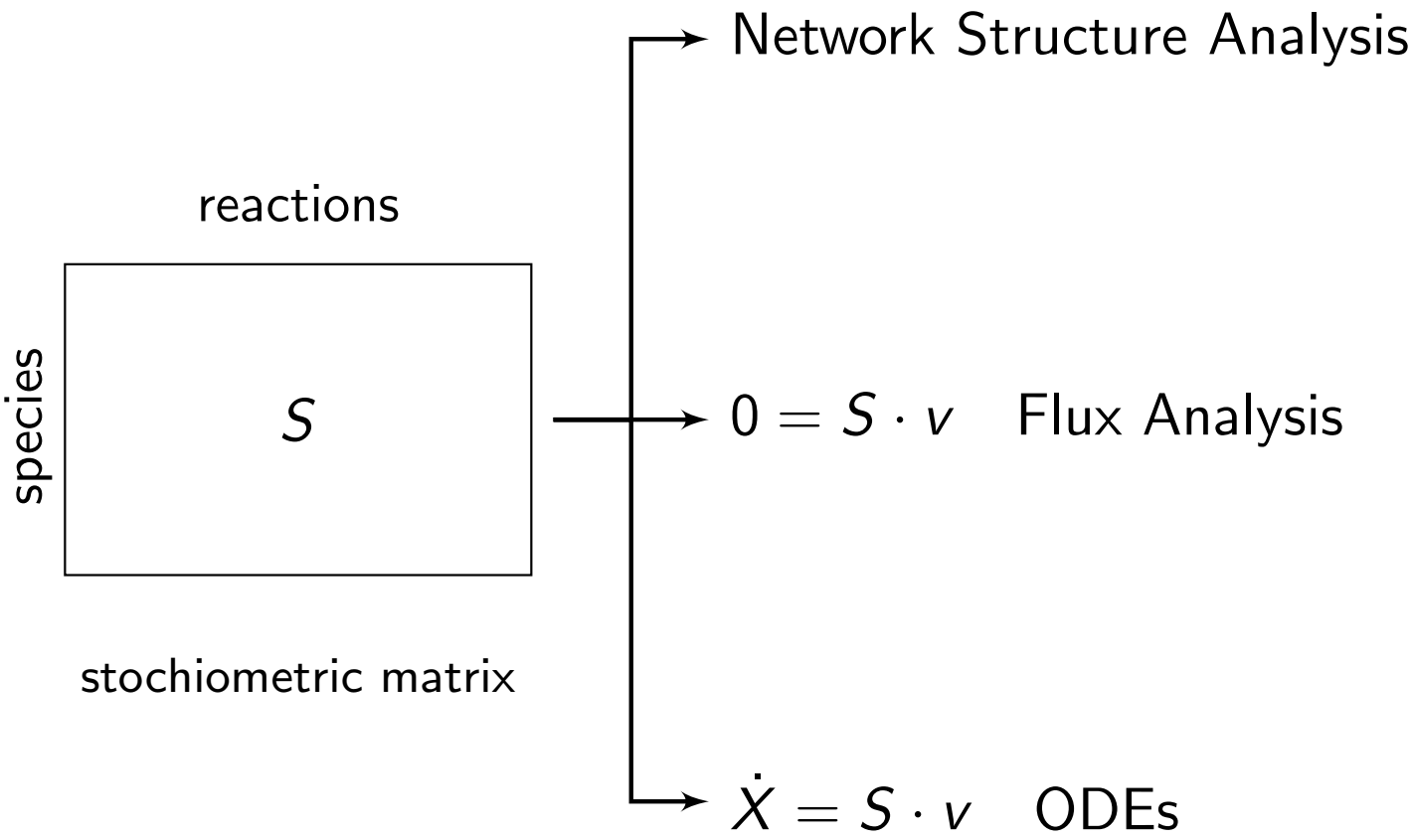
Phylogenetic tree of metabolic reconstractions



Our metabolic knowledge is srongly biased towards cultivatable bacteria.

Figure from Oberhardt MA et al (2009), Applications of genome-scale metabolic reconstructions, *Mol Syst Biol* **5**:320 | doi:10.1038/msb.2009.77

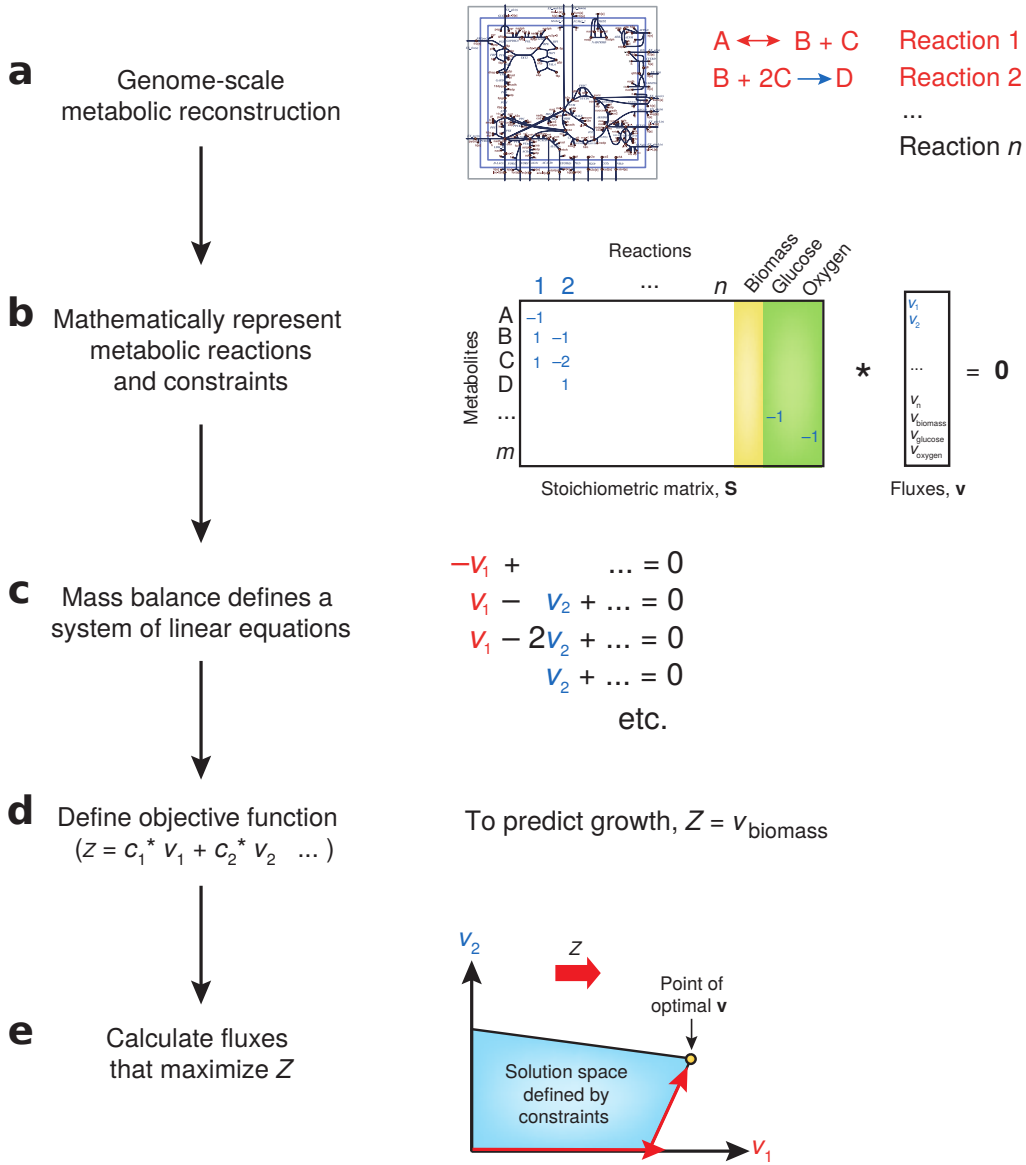
What to do next?



Analysis of the metabolic network depends on:

1. amount and quality of the available information.
2. desired description level of the system.

Steps to set up a Flux Balance Analysis (FBA)



a List of stoichiometrically balanced biochemical reactions.

b Mathematically model reaction network as stoichiometric matrix S.

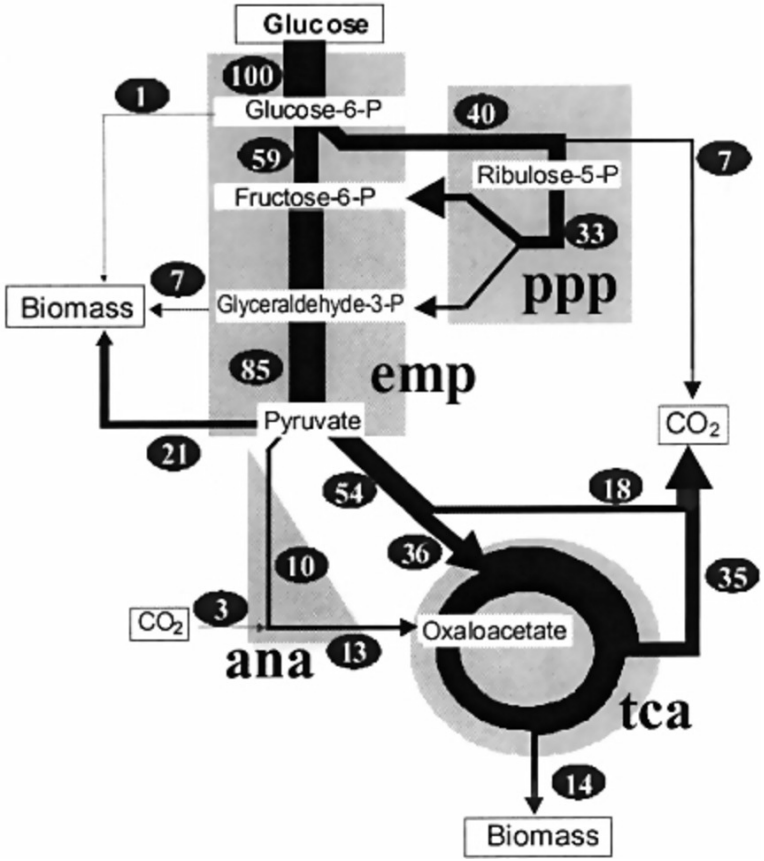
c Set up system of linear equations for steady state conditions $S \cdot v = 0$.

d Define a linear objective function Z, describing the growth of the system.

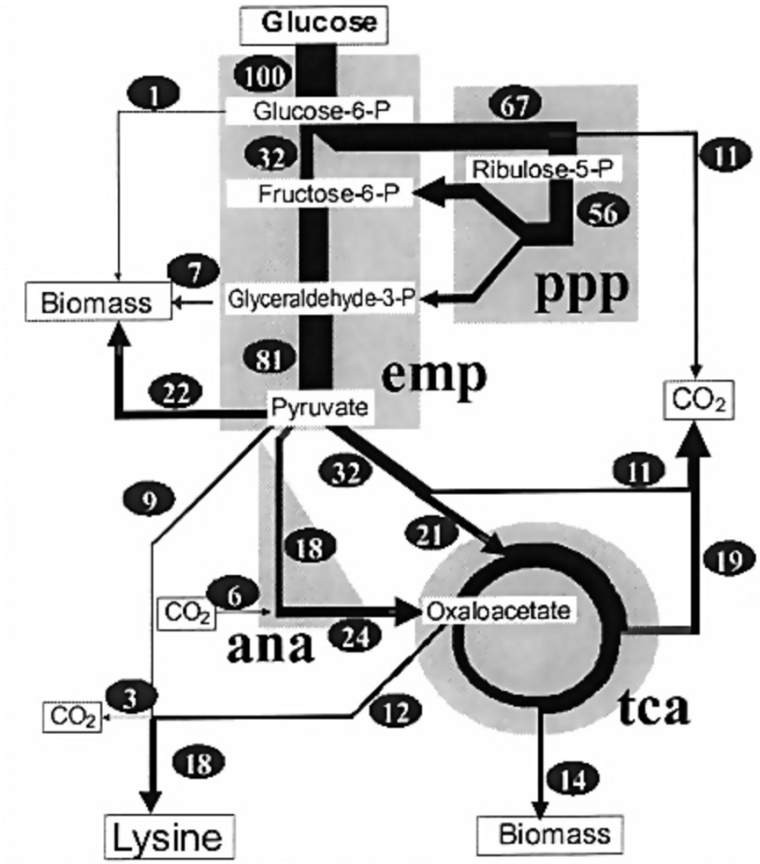
e Calculate a flux distribution using linear programming.

Figure adapted from Orth JD, Thiele I & Palsson BØ (2010), What is flux balance analysis?, *Nature Biotech* **28**(3):245-248 | doi:10.1038/nbt.1614

Information gained from Flux Analysis



normal growth condition.



Lys-production condition.

The goal is the quantitative description of cellular fluxes in relation to environmental conditions.

Figure adapted from Marx A (1997), Bestimmung des Kohlenstoffflusses im Zentralstoffwechsel von *Corynebacterium glutamicum* mittels ^{13}C-Isotopenanalyse, PhD-thesis Uni Düsseldorf.

Objective functions for Flux Balance Analysis

Formulate as an linear programing problem with additional constraints (capacity of enzymes, external fluxes, ...) and an appropriate optimization function.

Obj Function	Explanation
$\max \frac{v_{\text{Biomass}}}{v_{\text{Glucose}}}$	biomass yield (same as groth rate)
$\max \frac{v_{\text{ATP}}}{v_{\text{Glucose}}}$	ATP yield
$\min \frac{\sum v_{\text{NADH}}}{v_{\text{Glucose}}}$	redox potential
$\min \sum \delta_i$	reaction steps
$\max \frac{v_{\text{Biomass}}}{\sum v_i^2}$	biomass yield per flux unit

Schuetz R et al (2007), Systematic evaluation of objective functions for predicting intracellular fluxes in E. coli, Mol Sys Biol 3:119 | doi:10.1038/msb4100162

Primer: Linear Programming (LP)

Linear programming is an optimization method requiring 2 inputs:

1. A linear objective function.
2. A set of linear constraints.

Example: Production planning problem

Product	machine 1	machine 2	machine 3
A	40	24	0
B	24	48	60

Total machine running time is 8 hours/day.
Profit: 10 €/A and 40 €/B.

Question: How many units of product A and B need to be manifactured in order to maximize profit?

Expressed as LP Problem

1. maximize profit:

$$z = F(x_1, x_2) = 10 \cdot x_1 + 40 \cdot x_2$$

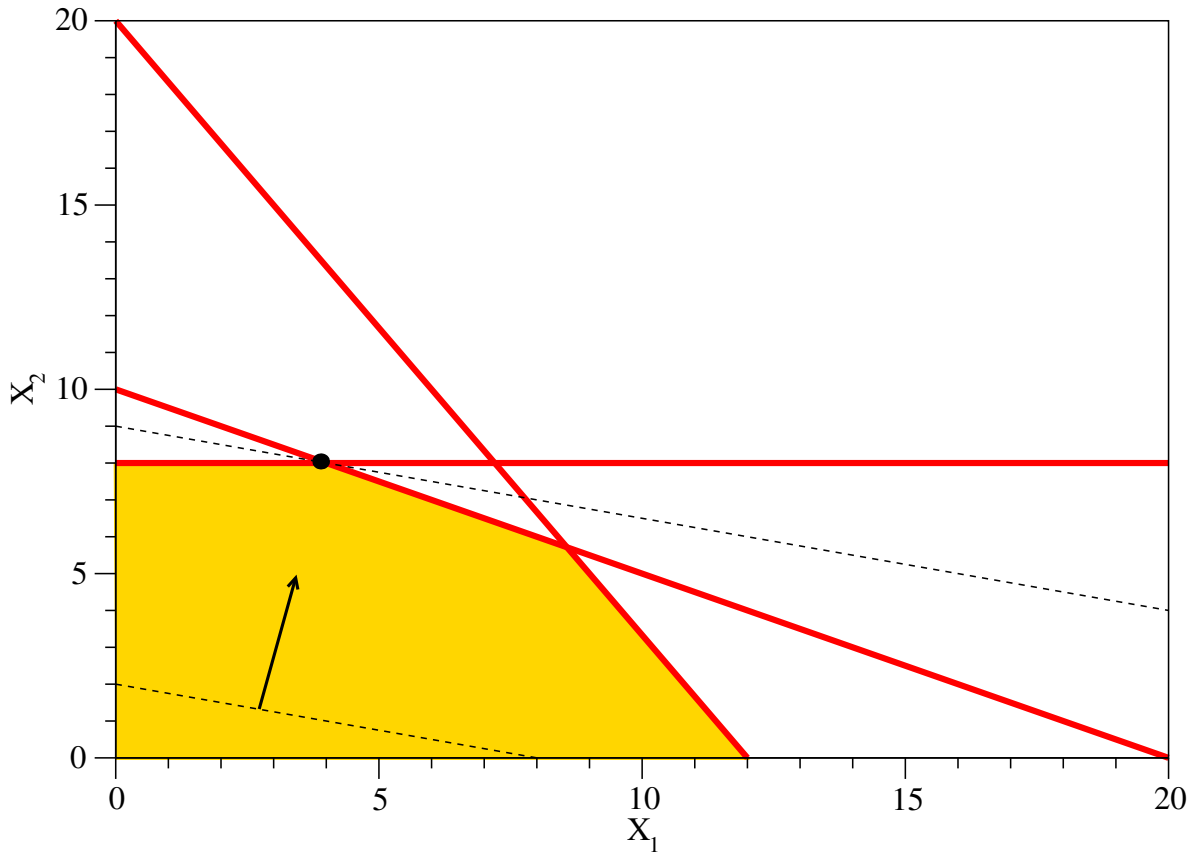
2. subject to the linear constraints:

$$\begin{aligned} 40 \cdot x_1 + 24 \cdot x_2 &\leq 480 \\ 24 \cdot x_1 + 48 \cdot x_2 &\leq 480 \\ 60 \cdot x_2 &\leq 480 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Admissible solutions:

- $x_1 = 0 \wedge x_2 = 0 \implies z = 0$
- $x_2 = 0 \curvearrowright x_1 = 12 \implies z = 120$
- $x_1 = 0 \curvearrowright x_2 = 8 \implies z = 320$

LP Problem: Graphical Solution



$$
\begin{aligned}
40 \cdot x_1 + 24 \cdot x_2 \leq 480 &\iff x_2 \leq -\tfrac{5}{3} \cdot x_1 + 20 \\
24 \cdot x_1 + 48 \cdot x_2 \leq 480 &\iff x_2 \leq -\tfrac{1}{2} \cdot x_1 + 10 \\
60 \cdot x_2 \leq 480 &\iff x_2 \leq 8
\end{aligned}
$$

$$
z = 10 \cdot x_1 + 40 \cdot x_2 \iff x_2 = -\tfrac{1}{4} \cdot x_1 + \tfrac{1}{40} \cdot z
$$

LP Problem: Formal Formulation

The linear objective function is generally a sum of terms that contain weighted measurable elements from a metabolic model.

Maximize:

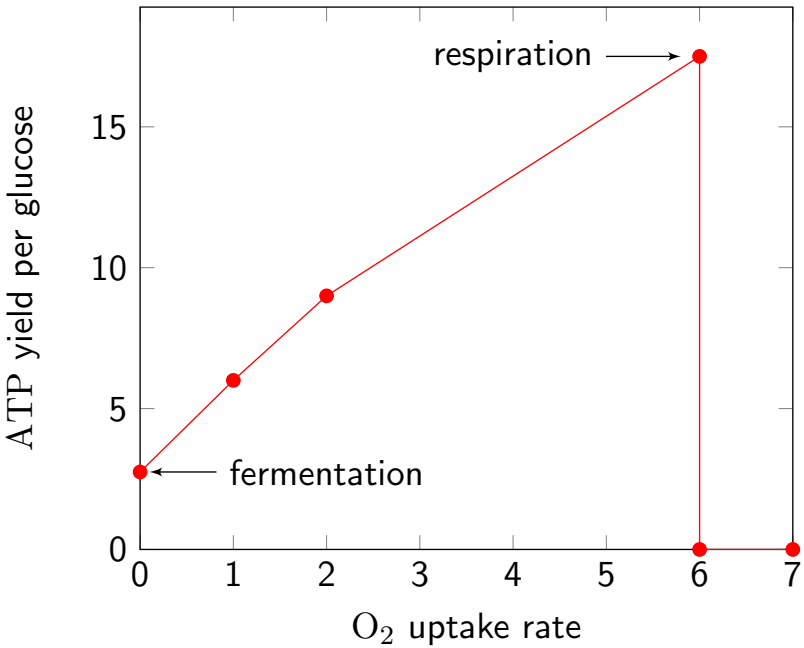
$$Z = c_1 \cdot x_1 + c_2 \cdot x_2 + \cdots = \mathrm{c}^T \cdot \mathrm{x}$$

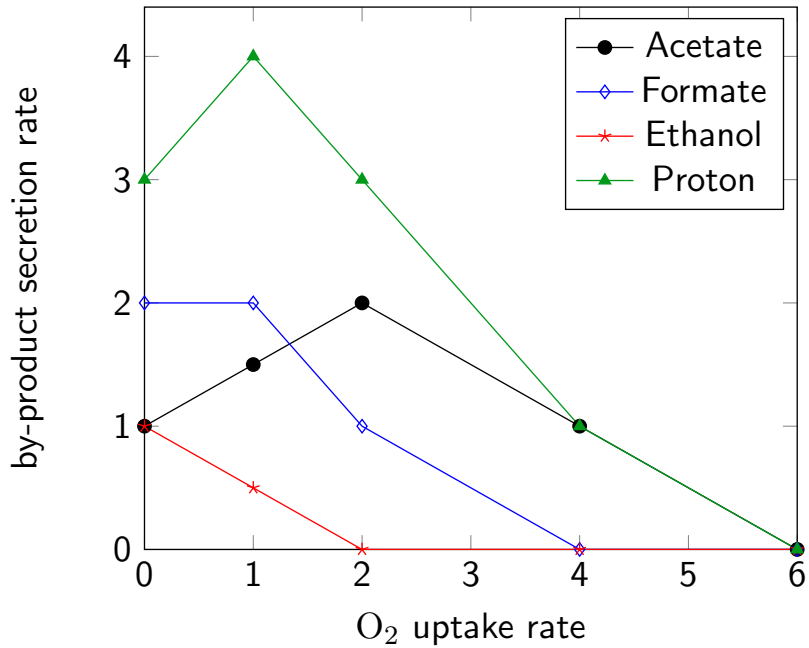
Subject to:

$$\begin{array}{ccc} a_{11} \cdot x_1 + a_{12} \cdot x_2 + & \cdots & a_{1n} \cdot x_n \leq b_1 \\ a_{12} \cdot x_1 + a_{22} \cdot x_2 + & \cdots & a_{2n} \cdot x_n \leq b_2 \\ \vdots & & \vdots \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + & \cdots & a_{mn} \cdot x_n \leq b_m \end{array}$$

$$\mathrm{A} \cdot x \leq b$$

Variation Analysis: Single Parameter



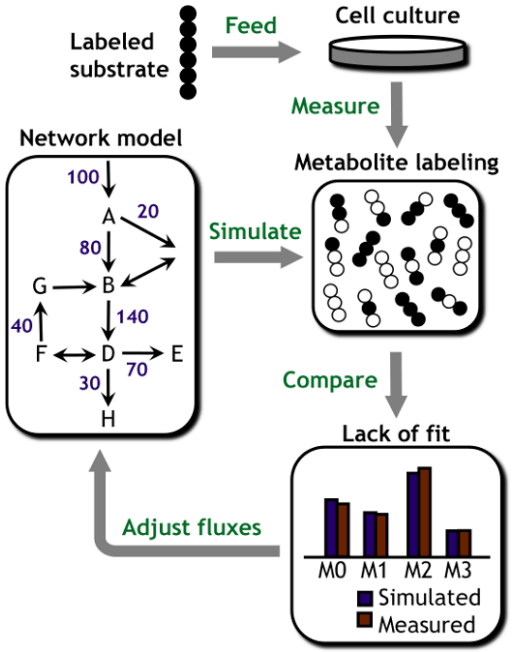


1. ATP yield curve is piecewise linear.
2. ATP yield varies between 2.75 (anaerobic) and 17.5 (aerobic).
3. kinks correspond to changes in by-product secration rates.
4. 3 distinct optimal phenotypes (0-1, 1-2, 2-6 O_2 uptake rates).
5. The ATP yield relies on different pathways.

There is no solution for oxygen uptake rates above 6, because the equations cannot be balanced.

Edwards JS, Palsson BØ (2000), Robustness Analysis of the *Escherichia coli* Metabolic Network, *Biotech Progress* **16**(6):927-939 | doi:10.1021/bp0000712

Isotope Labeling Experiments



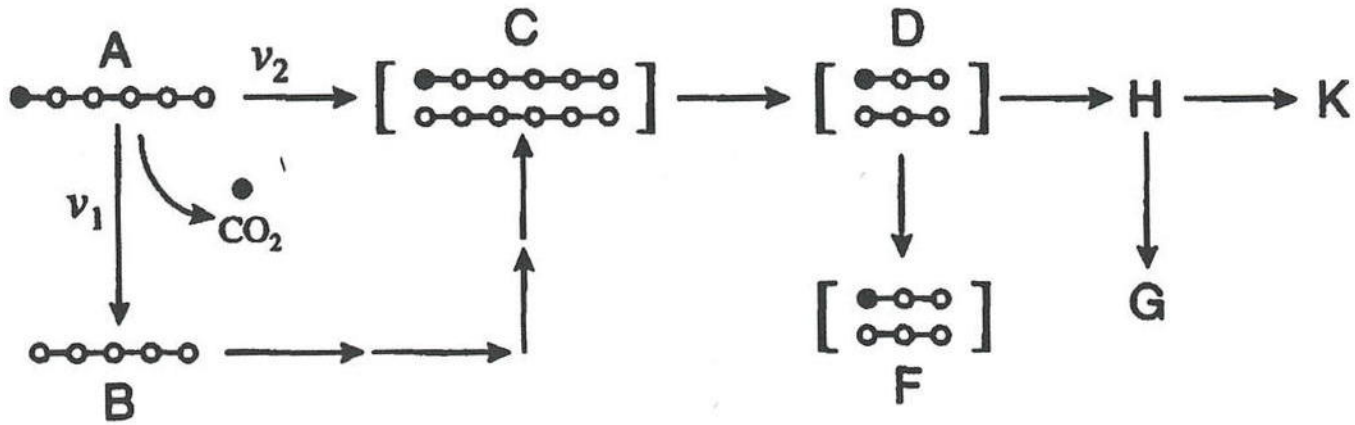
1. Introduce a isotope labeled substrate into the cell culture at metabolic steady state.
2. Allow the system to reach an isotopic steady state.
3. Measure (e.g. NMR, MS) relative labeling in metabolic intermediates and by products.
4. Estimate fluxes from these measurements.

- Isotope measurements provide many additional independent constraints for MFA.
- The cell's **flux state** and the administered isotope label fully and uniquely determine the **isotopomere patterns** of metabolites at steady state.
- **Quantitative interpretation** requires a mathematical model, which relates metabolic flux to isotopomere abundance.

Figure adapted from Duckwall CS, Murphy TA & Young JD (2013), Mapping cancer cell metabolism with ^{13}C flux analysis: Recent progress and future challenges, *J Carcinog* **12**:13 | doi:10.4103/1477-3163.115422 (PMC3746411).

Determination of Flux Split Ratios

Method works only at both **metabolic** and isotopic steady state.
In isotopic steady state the relative population all the isotopomeres of a metabolite is constant.



1. A is a six carbon compound whose first atom is labeled.
2. Competing/alternative pathways must introduce asymmetries.
3. The pathway via intermediate B produces unlabeled C.
4. Via the "direct" pathway the label is retained.
5. The label enrichment in C is directly proportional to the rate of ν_2 relative to the total rate of A consumtion $(\nu_1 + \nu_2)$.

Isotopomers

Are defined as isomeres of a metabolit that differ only in the labeling state of their individual atoms

(e.g. carbon [^{12}C, ^{13}C], hydrogen [^{1}H, ^{2}H] or oxygen [^{16}O, ^{17}O, ^{18}O].

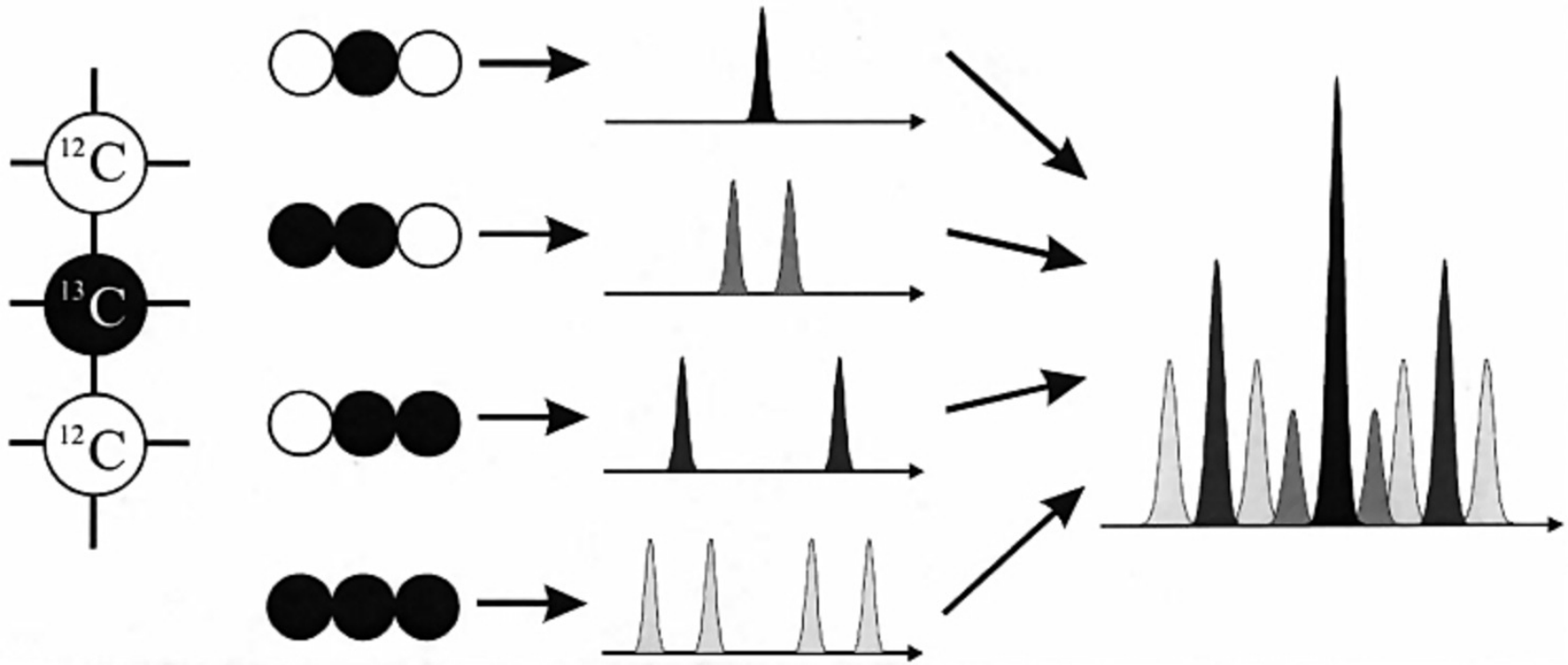
2^N isotopomeres are possible for a metabolite with N atoms that may be in one of two states (unlabeled or labeled).

Example (glucose $\mathrm{C_6H_{12}O_6}$)

Atoms	# of Isotopomeres	
C	6.400×10^1	$(2^6 = 64)$
O	7.260×10^2	$(3^6 = 726)$
H	4.096×10^3	$(2^{12} = 4096)$
C, H	2.621×10^5	$(2^6 \times 2^{12} = 262144)$
C, H, O	1.911×10^8	$(2^6 \times 2^{12} \times 3^6 = 191102976)$

Measuring Isotopomere Distribution

Any experimental technique capable of detecting differences between isotopomeres can be used to measure the labeling state.



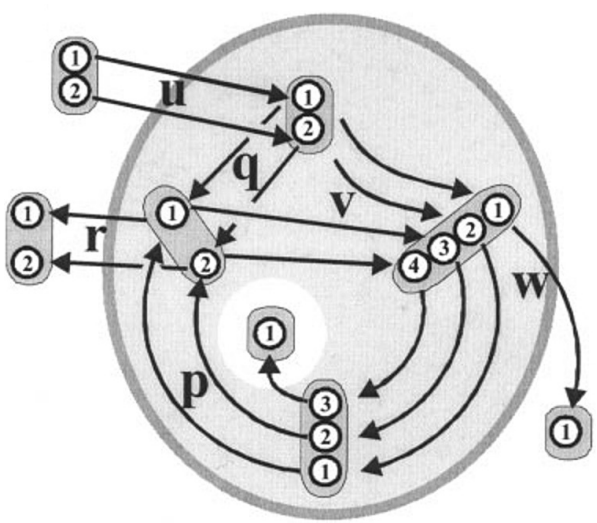
The two dominating technologies are:

1. Nuclear magnetic resonance spectroskopy (NMR).
2. Mass spectrometry (MS).

Figure adapted from [Wiechert, 2001]

Atom Transition Network

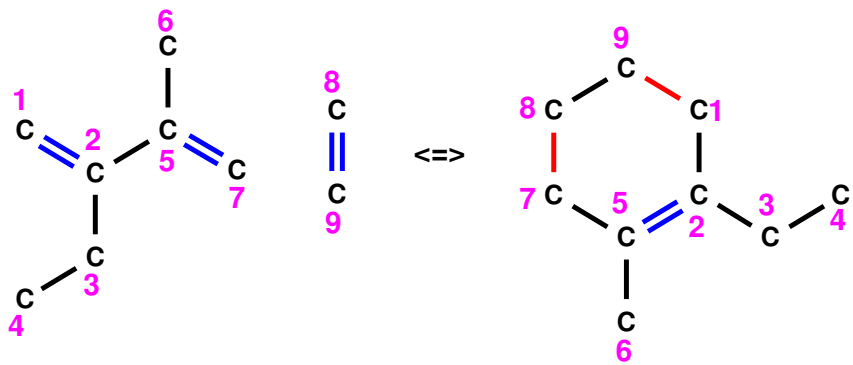
Estimating fluxes from isotopomere patterns is an inverse problem.



The *atom-atom mapping* between reaction educts and products must be known.

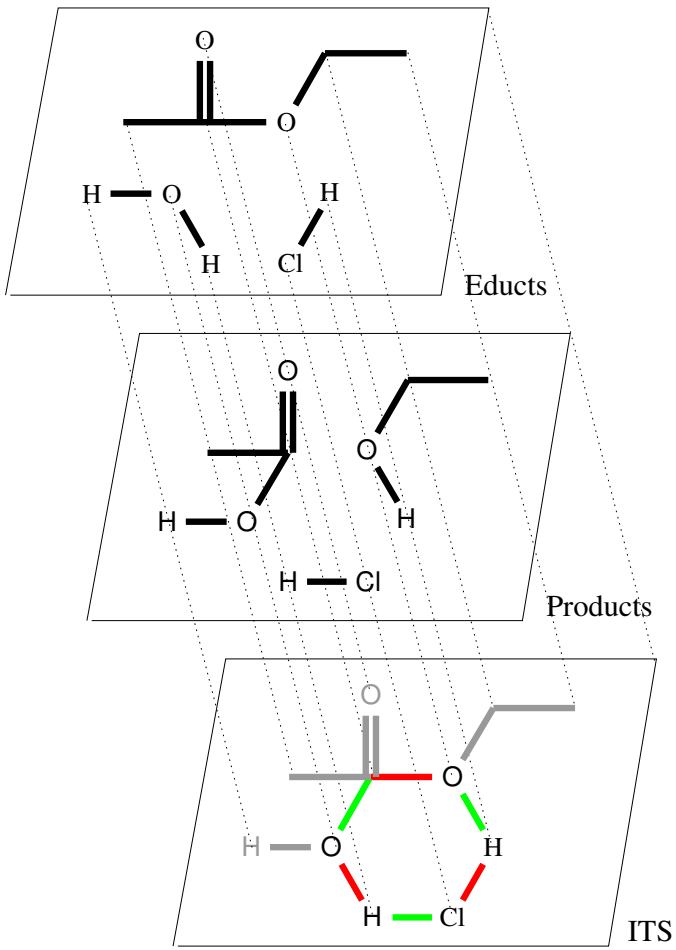
Getting this information is an *NP*-hard problem. Therefore most approaches use the heuristic principle of

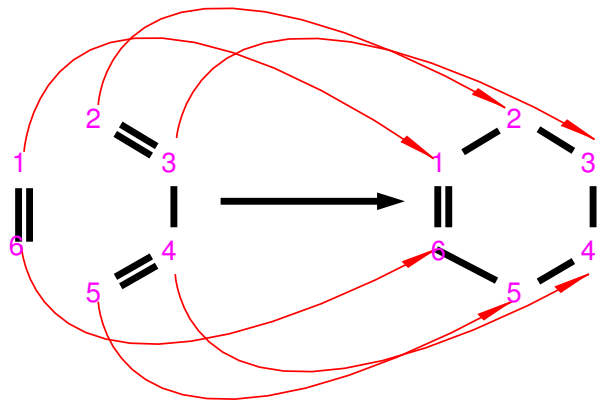
“minimal chemical distance”.



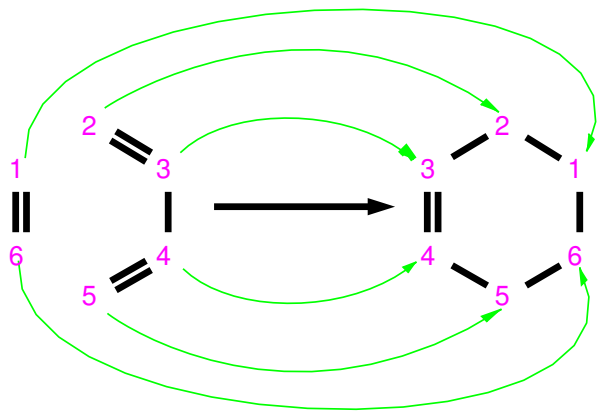
Chemical Reactions and Atom-to-Atom Mapping

$$CH_3CO_2Et + H_2O \overset{HCL}{\rightleftharpoons} CH_3CO_2H + EtOH$$





chemically wrong!



chemically correct

Fujita S (1988), A Novel Approach t o Systematic Classification of Organic Reactions, *J Chem Soc Perkin Trans 2* **5**:597-616 | 10.1039/P29880000597

Recap: Central Carbon Metabolism

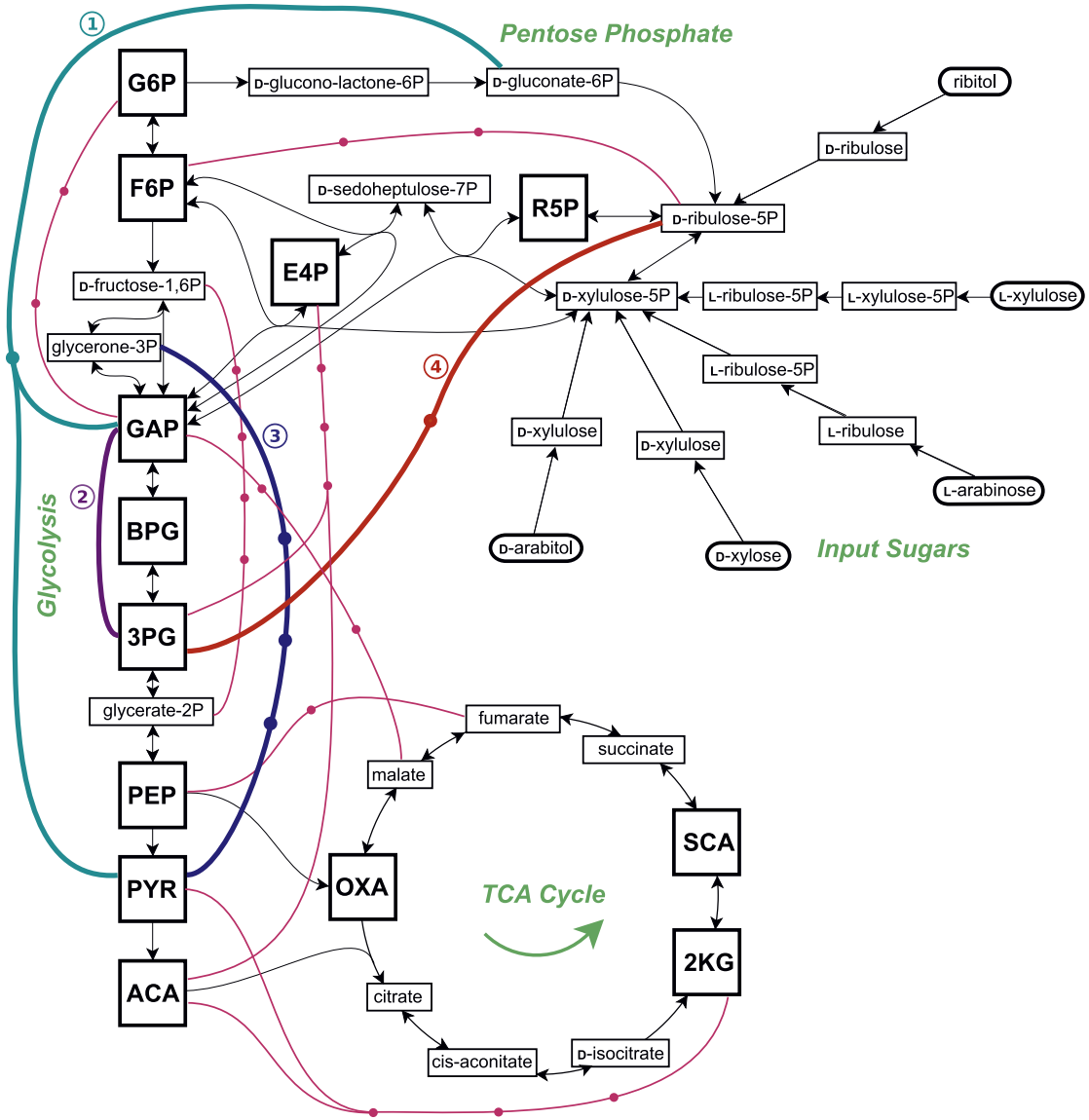
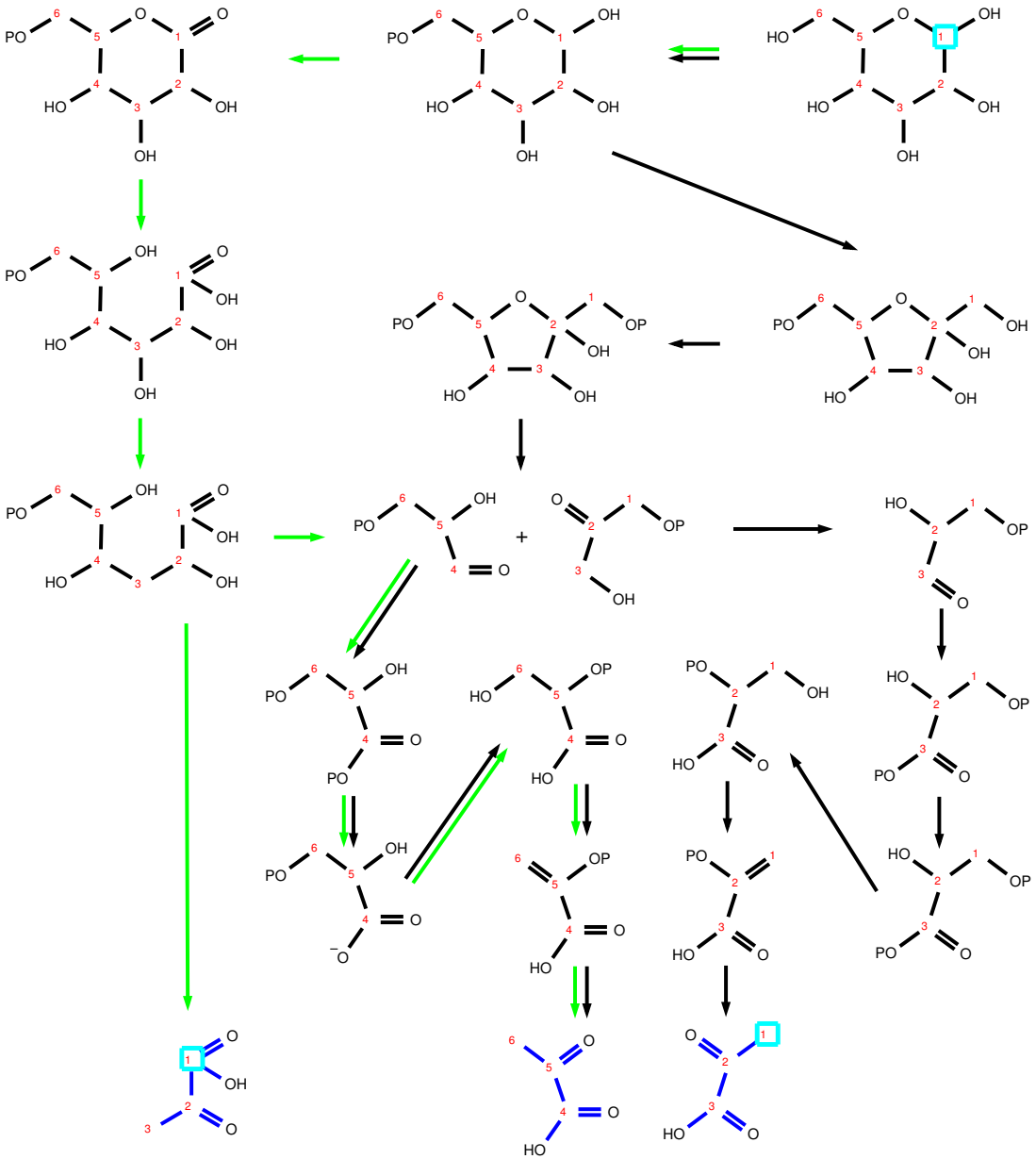


Figure from Noor et al (2010), Central Carbon Metabolism as a Minimal Biochemical Walk between Precursors for Biomass and Energy, J Mol Cell 39:809-820 | doi:10.1016/j.molcel.2010.08.031

Carbon atom trace of Glycolysis



`Glucose` ⇒ `Pyruvate`

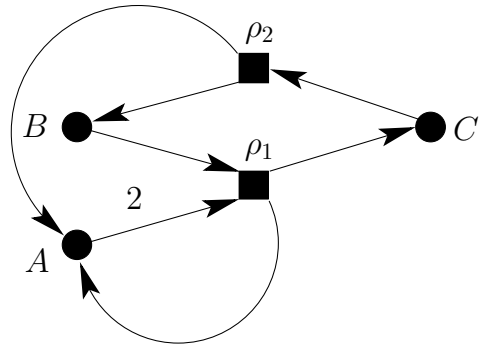
(Bio)Chemical *reaction databases* usually list only products, educts, and (sometimes) the type of transformation, but not the atom map itself.

Rule composition can be used to list all possible atom traces, here for the glycolysis (EMP) and the Entner-Doudoroff (ED) pathway from central carbon metabolism.

Andersen JL et al (2014), 50 Shades of Rule Composition: From Chemical Reactions to Higher Levels of Abstraction, *LNCS* **8738**:117-135 | doi:10.1007/978-3-319-10398-3_9

Connection between stoichiometric Matrix and ODEs

$$2A + B \rightarrow C + A$$
$$C \rightarrow A + B$$



$$\frac{\mathrm{d}}{\mathrm{dt}}[\vec{X}] = S \cdot \vec{v} = \frac{\mathrm{d}}{\mathrm{dt}} \begin{pmatrix} [A] \\ [B] \\ [C] \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_{\rho_1} \\ v_{\rho_2} \end{pmatrix}$$

Rate Model

mass action kinetics

$$v_{\rho_1} = k_{\rho_1} \cdot [A]^2 \cdot [B]$$
$$v_{\rho_2} = k_{\rho_2} \cdot [C]$$

$$[\dot{A}] = k_{\rho_2} \cdot [C] - k_{\rho_1} \cdot [A]^2 \cdot [B]$$
$$[\dot{B}] = k_{\rho_2} \cdot [C] - k_{\rho_1} \cdot [A]^2 \cdot [B]$$
$$[\dot{C}] = k_{\rho_1} \cdot [A]^2 \cdot [B] - k_{\rho_2} \cdot [C]$$

Mass action kinetics is grounded in the theory of collision processes; **alternative rate models** to describe the reaction velocities for enzyme systems are called **Michaelis-Menten** kinetics or **Hill** equations.

Describing the Dynamics on the Network

Basic Idea:

1. Characterize the state of a system by a vector of state variables (e.g. concentrations of chemical species).
2. Formulate equations describing the change of state variables.

Solution:
Use ordrinary differantial equations (ODEs)

$$\frac{\mathrm{d}X_i}{\mathrm{d}t} = F_i(X_1, X_2, \ldots, X_n; p_1, p_2, \ldots, p_m), i = 1, 2, \ldots, n$$

X_i	...	state variables e.g. concentrations
p_i	...	system parameters e.g. kinetic constants
t	...	time

Solving ODEs

Possibilities:

1. **analytic** solution.
2. **numeric** solution.

Analytic solutions can be found for *linear systems*.

Advantage:

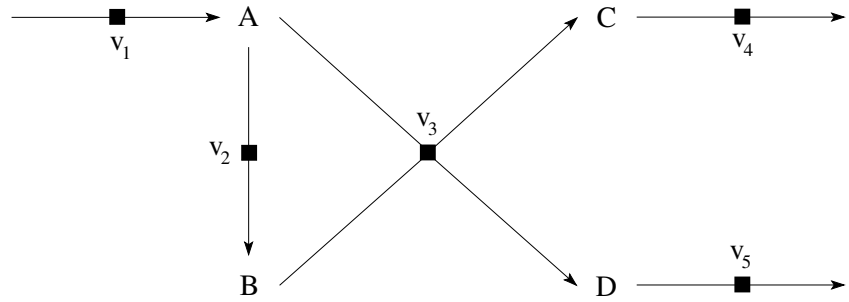
- solution is **valid** for all initial conditions.
- parameter dependencys can easily be discovered.

Nonlinear systems (typical case for (bio)chemical and genetic systems) **exhibit a wide range of dynamic behaviors** and do not typically admit analytic solutions.

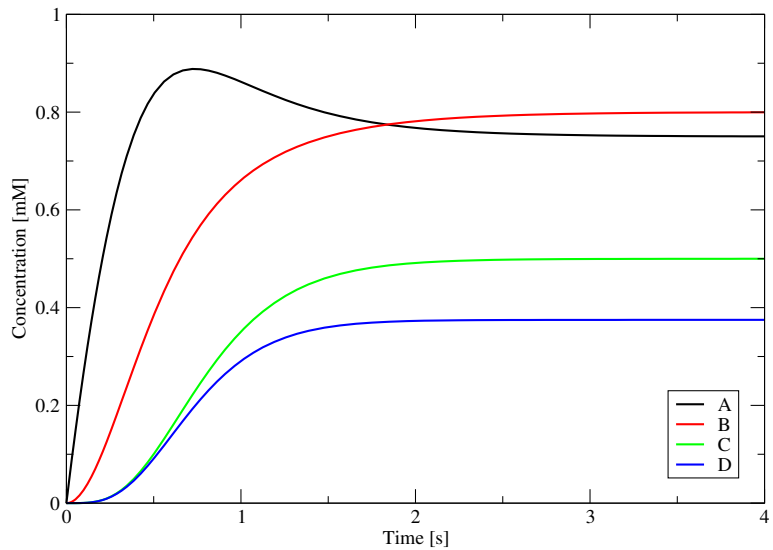
The **accuracy** of the numeric solution **strongly depends on** the used **integration algorithm** (e.g. Euler method).

Example: Nonlinear (open) Reaction Sysem

$$
\begin{aligned}
\frac{dA}{dt} &= v_1 - v_2 - v_3 \\
\frac{dB}{dt} &= v_2 - v_3 \\
\frac{dC}{dt} &= v_3 - v_4 \\
\frac{dD}{dt} &= v_3 - v_5
\end{aligned}
$$



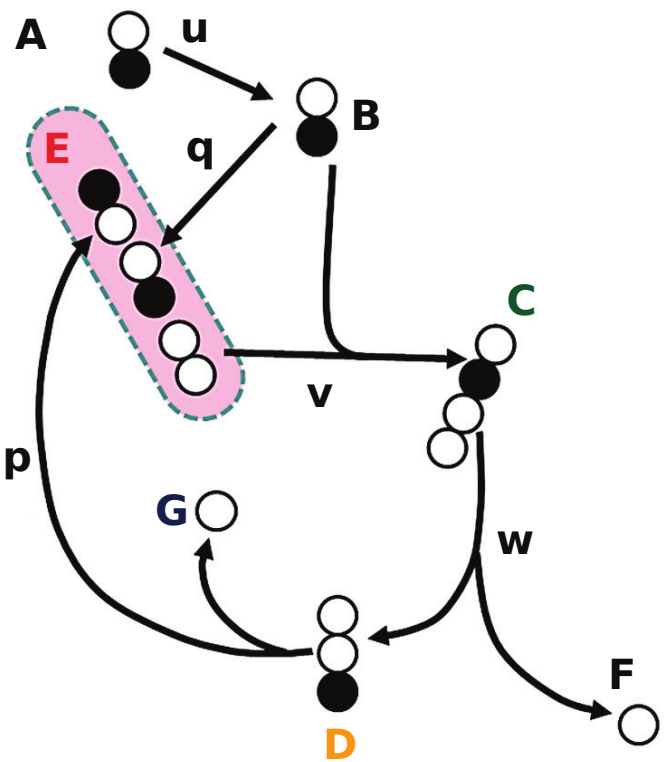
$$
\begin{aligned}
v_1 &= k_1 \\
v_2 &= k_2 \cdot [A] \\
v_3 &= k_3 \cdot [A] \cdot [B] \\
v_4 &= k_4 \cdot [C] \\
v_5 &= k_5 \cdot [D]
\end{aligned}
$$

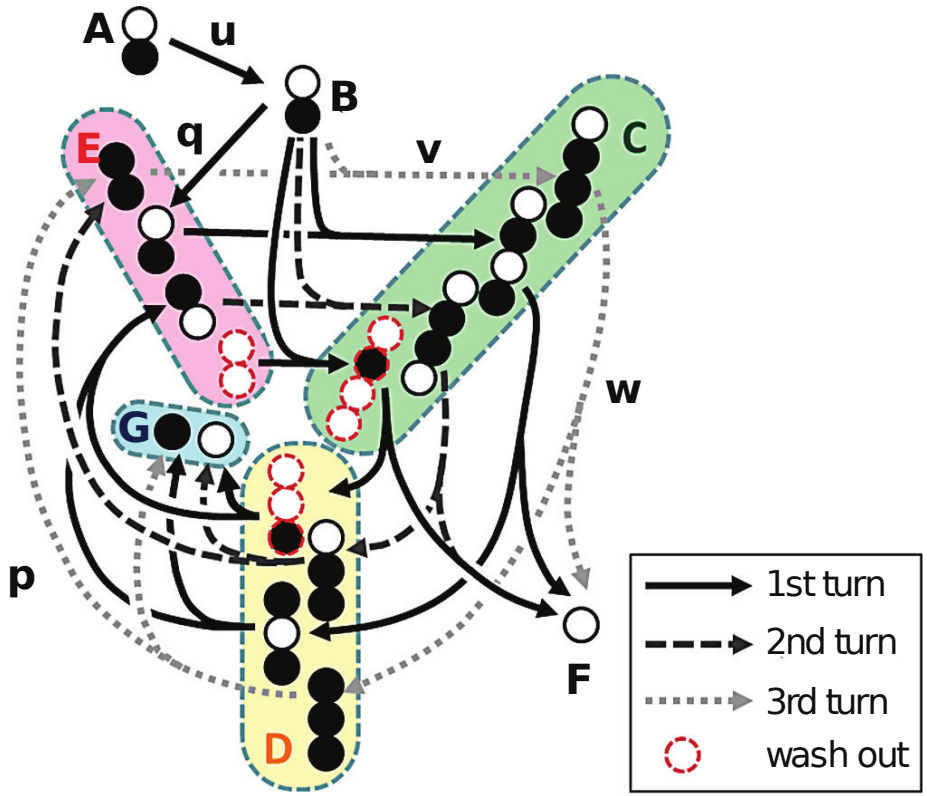


Note: because rate v_3 depends on the product of $[A]$ and $[B]$, this system of equations is nonlinear.

k_1=3 mM/s, k_2=2/s, k_3=2.5/mM/s, k_4=3/s and k_5=4/s

Dynamics of Isotopomere Patterns





With each turn of the cycle further isotopomeres emerge.
Isotopomeres marked in dashed red are "washed out" over time.

Isotopomeres shown in unbroken black reside in the stationary distribution.

Further Reading

Gagneur J, Klamt S.

Computation of elementary modes: a unifying framework and the new binary approach.
BMC Bioinf **5**:175 2004 | doi:10.1186/1471-2105-5-175

Wagner C, Urbanczik R.

The Geometry of the Flux Cone of a Metabolic Network.
Biophys J **89**:38373845, 2005 | doi:10.1529/biophysj.104.055129

Price ND, Reed JL, Palsson BO.

Genome-scale models of microbial cells: Evaluating the consequences of constraints.
Nature Rev Nicrobiol **2**:886-897, 2004 | doi:10.1038/nrmicro1023

Orth JD, Ines Thiele I, Palsson BO

What is flux balance analysis?
Nature Biotech **28**:245-248, 2010 | doi:10.1038/nbt.1614

Antoniewicz MR, Kelleher JK, Stephanopoulos G.

Elementary metabolic units (EMU): A novel framework for modeling isotopic distributions.
Metab Eng **9**:68-86, 2007 | doi:10.1016/j.ymben.2006.09.001

Sauer U.

Metabolic networks in motion: ^{13}C flux analysis.
Mol Sys Biol **2**:62, 2006 | doi:10.1038/msb4100109

Wiechert W.

^{13}C Metabolic Flux Analysis.
Metab Eng **3**(3):195–206, 2001 | doi:10.1006/mben.2001.0187

Chen WL, Chen DZ, Taylor KT.

Automated reaction mapping and reaction center detection.
WIREs Comput Mol Sci 2013 | doi:10.1002/wcms.1140