Flux Balance Analysis and its applications

Balázs Papp

Balazs.Papp@manchester.ac.uk

Evangelos Simeonidis

v.simeonidis@manchester.ac.uk

Topics covered

- **Constraint-based approach to metabolic modelling**
- Principles of FBA and some biological applications
- Mathematics behind FBA: Optimisation
- Mathematical examples

Introduction

• A major goal of systems biology is to relate genome sequence to cell physiology

• This requires the identification of the components and their interactions in the system + mathematical modelling

• Small molecule metabolism is the best described molecular network in the cell and there are various computational tools to model its behaviour

Metabolic network reconstructions

Network reconstruction = delineation of the chemical and physical interactions between the components

Abbreviation	Glycolytic reactions	Genes
HEX1	[c]GLC + ATP → G6P + ADP + H	*glk*
PGI	[c]G6P ↔ F6P	*pgi*
PFK	[c]ATP + F6P → ADP + FDP + H	*pfkA, pfkB*
FBA	[c]FDP ↔ DHAP + G3P	*fbaA, fbaB*
TPI	[c]DHAP ↔ G3P	*tpiA*
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH	*gapA, gapC1, gapC2*
PGK	[c]13DPG + ADP ↔ 3PG + ATP	*pgk*
PGM	[c]3PG ↔ 2PG	*gpmA, gpmB*
ENO	[c]2PG ↔ H_2O + PEP	*eno*
PYK	[c]ADP + H + PEP → ATP + PYR	*pykA, pykF*

Reed JL, Famili I, Thiele I, Palsson BO. Towards multidimensional genome annotation.
Nat Rev Genet. 2006 Feb;7(2):130-41. Review.

PMID: 16418748

Metabolic network reconstructions

- Automated metabolic reconstructions for > 500 organisms based on genome sequence data (e.g. KEGG database)
- Automated reconstructions are usually not suitable for modelling
- Manual assembly gives higher quality networks and is based on genomic + biochemical + physiological data

Reed JL, Famili I, Thiele I, Palsson BO. Towards multidimensional genome annotation.
Nat Rev Genet. 2006 Feb;7(2):130-41. Review.
PMID: 16418748

High quality manual reconstructions

- Incorporate information on:
  - reaction reversibility
  - cofactor usage
  - transport reactions
  - cellular compartments (e.g. mitochondrion)
  - biomass composition
- Only available for well studied microbes (e.g. yeast, *E. coli* and ~10 other bacteria)
- Amenable to modelling

For examples, see:

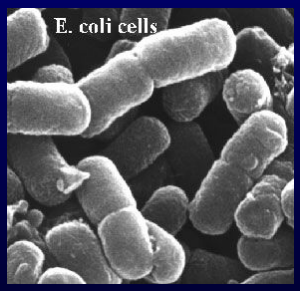
http://gcrg.ucsd.edu/organisms/index.html

High quality manual reconstructions

Example: *Escherichia coli* metabolic reconstruction*

- the best characterized network
- 931 reactions
- 625 different metabolites

But: 67 are dead end!

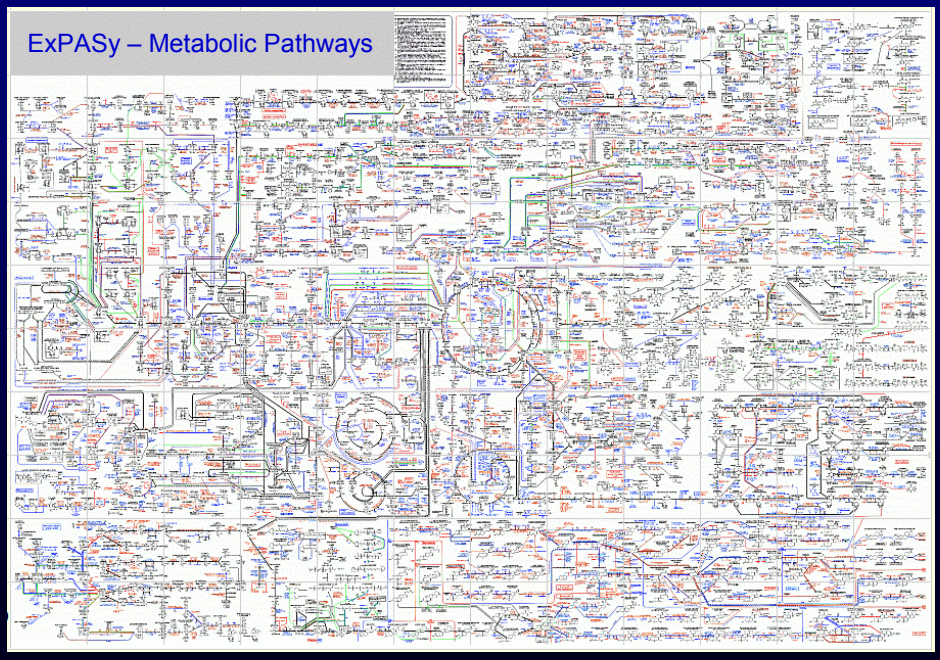


* Reed et al . (2003) Genome Biol 4: R54

Reed JL, Vo TD, Schilling CH, Palsson BO. An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR).
Genome Biol. 2003;4(9):R54. Epub 2003 Aug 28.

PMID: 12952533

How to analyse such a complex network?



http://www.expasy.ch/cgi-bin/search-biochem-index

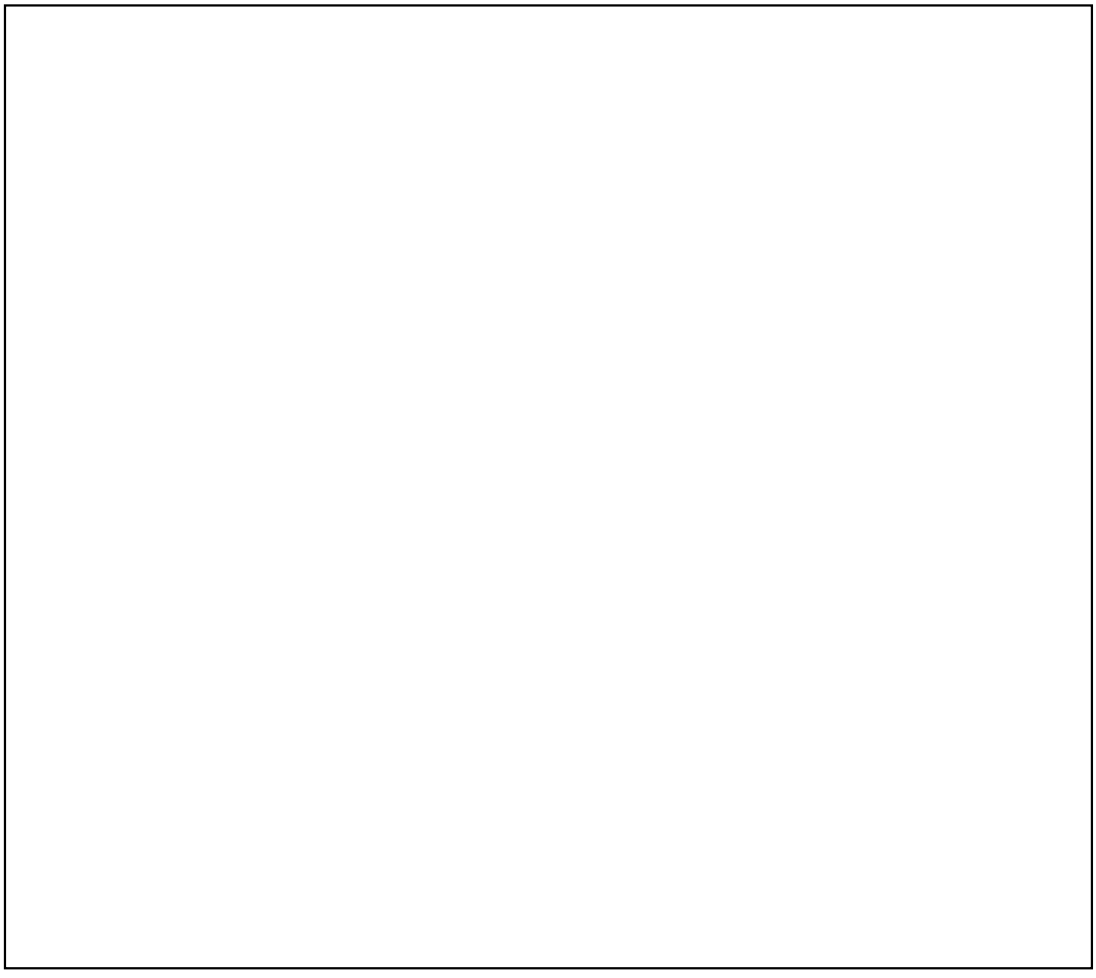
In silico analysis of metabolic networks

- **Topological analysis**

  → identify pathways, redundancies in the network, graph theoretical properties, etc.

- **Modelling:** simulating the behaviour of metabolism

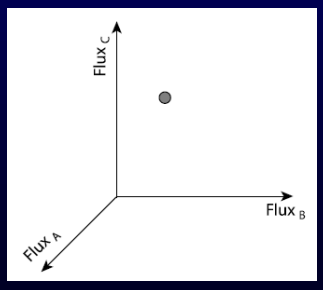
  → deduce phenotype from genotype + environment



Two modelling approaches

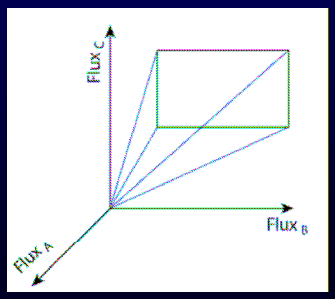
Mechanistic
(kinetic)

Find an exact solution



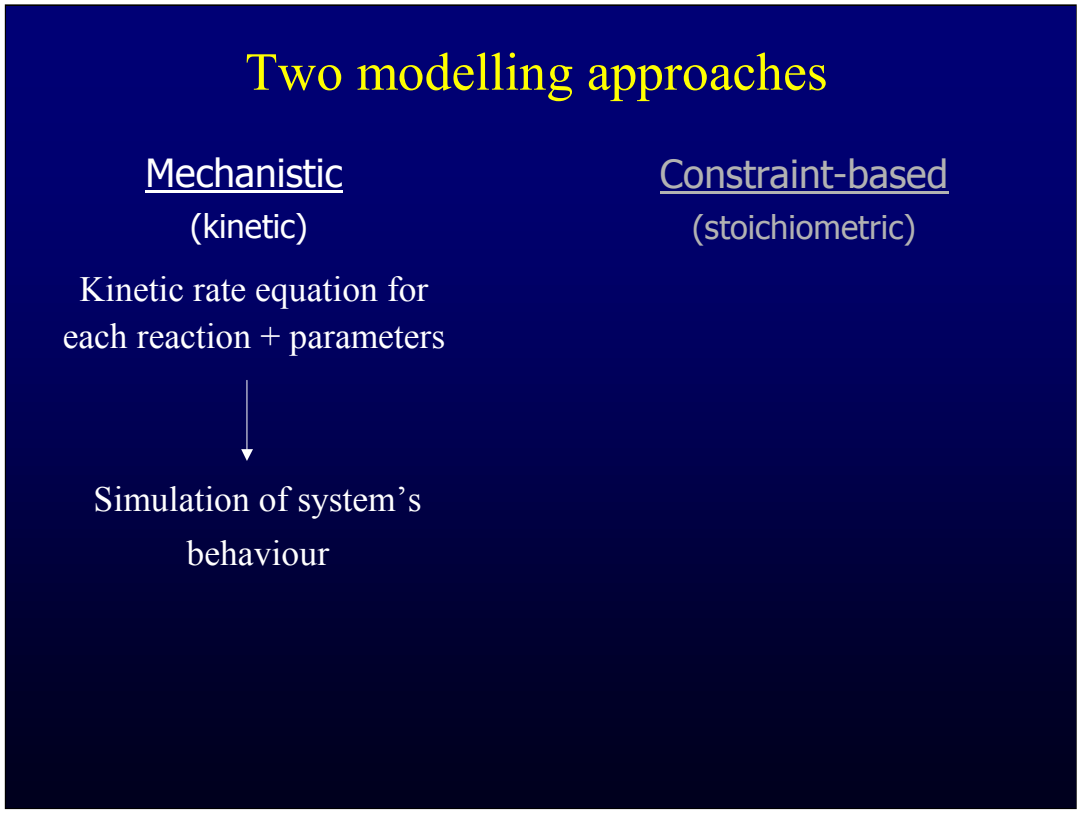
Constraint-based
(stoichiometric)

Find a range of allowable solutions



Covert MW, Famili I, Palsson BO. Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?
Biotechnol Bioeng. 2003 Dec 30;84(7):763-72. Review.
PMID: 14708117

Wiechert W. Modeling and simulation: tools for metabolic engineering.
J Biotechnol. 2002 Mar 14;94(1):37-63. Review.
PMID: 11792451



Covert MW, Famili I, Palsson BO. Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?
Biotechnol Bioeng. 2003 Dec 30;84(7):763-72. Review.
PMID: 14708117

Wiechert W. Modeling and simulation: tools for metabolic engineering.
J Biotechnol. 2002 Mar 14;94(1):37-63. Review.
PMID: 11792451

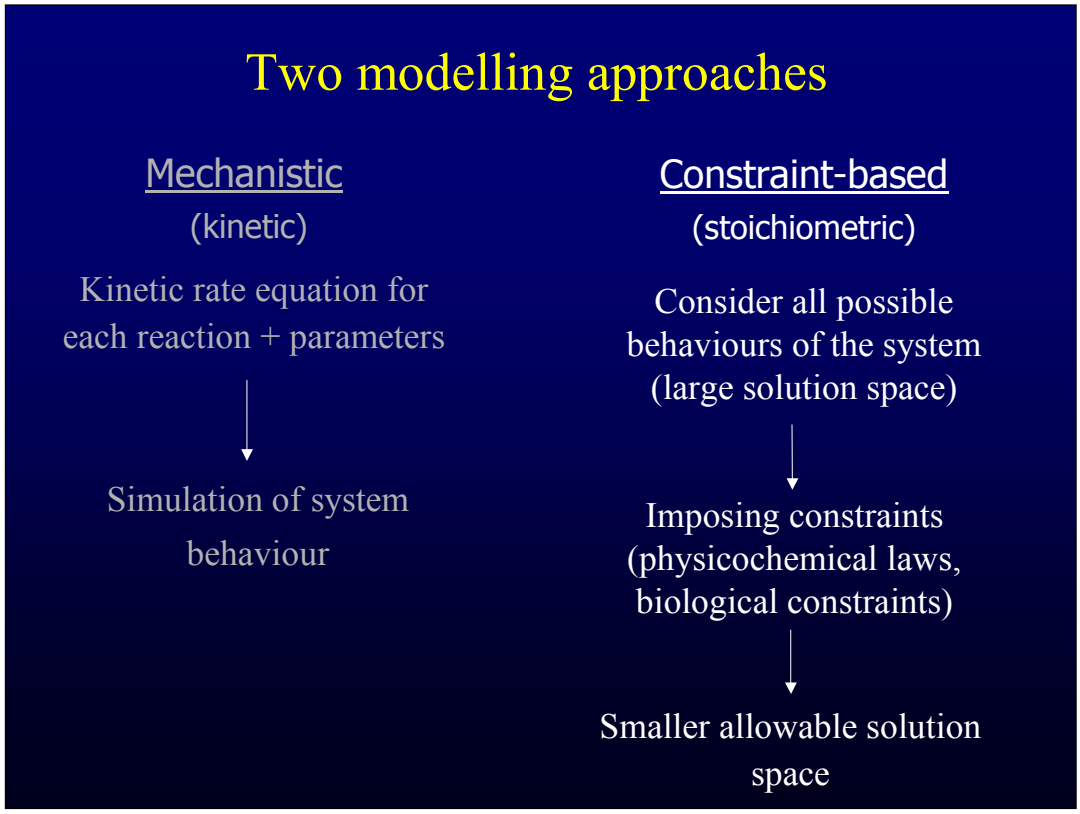
Problem with kinetic modelling

A lot of data is required to parameterize large-scale models, experimentally intractable at present.

The largest kinetic metabolic model available:

→ Human red blood cell (35 enzymes)

Jamshidi N, Edwards JS, Fahland T, Church GM, Palsson BO. Dynamic simulation of the human red blood cell metabolic network. Bioinformatics. 2001 Mar;17(3):286-7.
PMID: 11294796



Covert MW, Famili I, Palsson BO. Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?
Biotechnol Bioeng. 2003 Dec 30;84(7):763-72. Review.
PMID: 14708117

Types of constraints

- **Physico-chemical constraints**

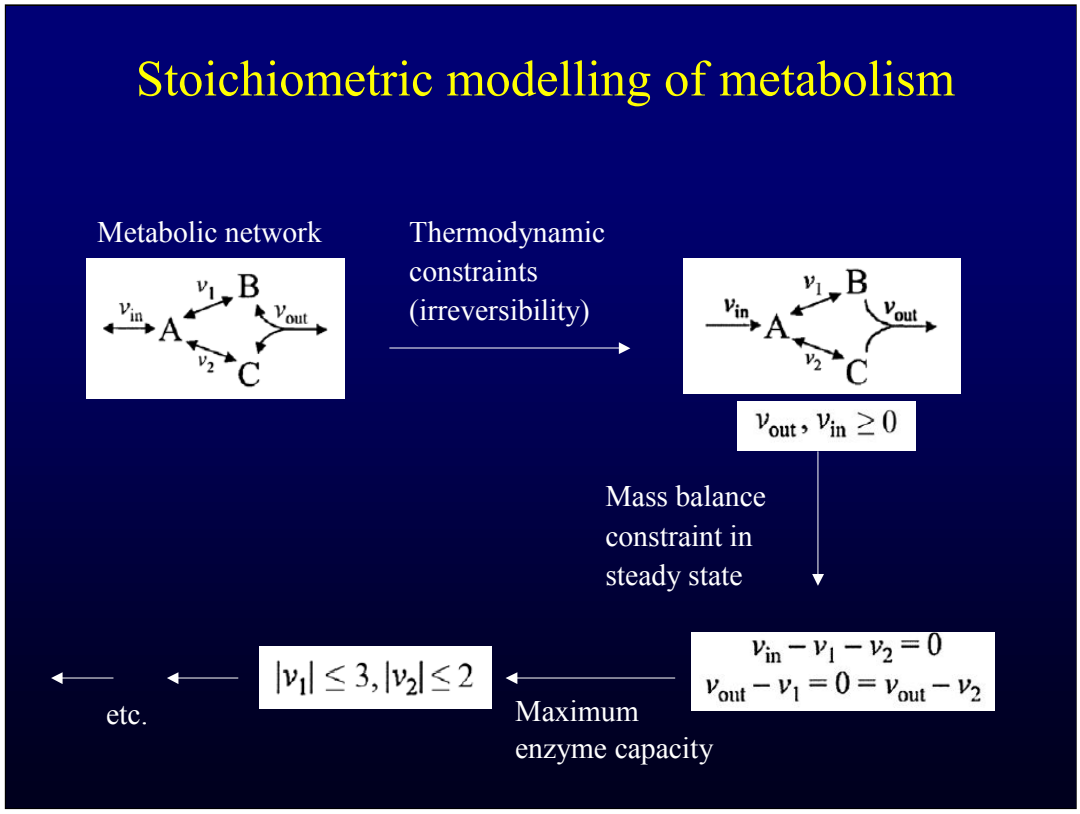
  → mass, charge and energy conservation, laws of thermodynamics

- **Biological constraints:**

  → external environment, regulatory constraints

Covert MW, Famili I, Palsson BO. Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?
Biotechnol Bioeng. 2003 Dec 30;84(7):763-72. Review.
PMID: 14708117

Price ND, Reed JL, Palsson BO. Genome-scale models of microbial cells: evaluating the consequences of constraints.
Nat Rev Microbiol. 2004 Nov;2(11):886-97. Review.
PMID: 15494745



Covert MW, Famili I, Palsson BO. Identifying constraints that govern cell behavior: a key to converting conceptual to computational models in biology?
Biotechnol Bioeng. 2003 Dec 30;84(7):763-72. Review.
PMID: 14708117

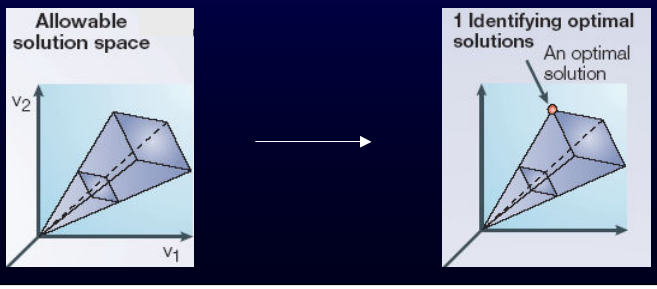
Price ND, Reed JL, Palsson BO. Genome-scale models of microbial cells: evaluating the consequences of constraints.
Nat Rev Microbiol. 2004 Nov;2(11):886-97. Review.
PMID: 15494745

Problem:

Addition of constraints reduces the allowable solution space, but usually not to a single point (underdetermined system).

How to find a particular solution?

We can look for a solution which optimises a particular network function (e.g. production of ATP or biomass) → **FBA**

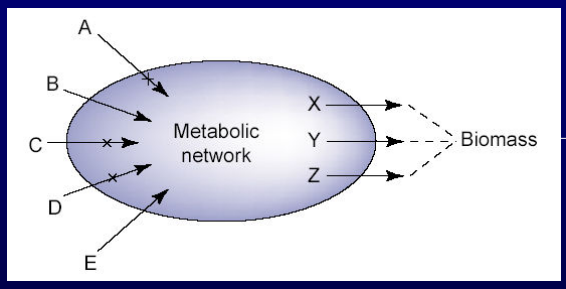


Price ND, Reed JL, Palsson BO. Genome-scale models of microbial cells: evaluating the consequences of constraints.
Nat Rev Microbiol. 2004 Nov;2(11):886-97. Review.
PMID: 15494745

Topics covered

- Constraint-based approach to metabolic modelling
- **Principles of FBA and some biological applications**
- Mathematics behind FBA: Optimisation
- Mathematical examples

Flux Balance Analysis (FBA) with growth optimisation



growth rate (fitness)

1) Start from a reconstructed network (transport processes, direction and stoichiometry of reactions, biomass components (X,Y,Z) important for cell growth)
2) Specify the nutrients available in the environment (B,E) and impose constraints (mass balance, etc.)
3) Calculate optimal growth rate: maximize biomass yield

Reviews:

Kauffman KJ, Prakash P, Edwards JS. Advances in flux balance analysis.
Curr Opin Biotechnol. 2003 Oct;14(5):491-6.
PMID: 14580578

Bonarius HPJ, Schmid G, Tramper J (1997) Flux analysis of underdetermined metabolic networks: The quest for the missing constraints.
Trends Biotech **15**: 308–314.

Price ND, Reed JL, Palsson BO. Genome-scale models of microbial cells: evaluating the consequences of constraints.
Nat Rev Microbiol. 2004 Nov;2(11):886-97. Review.
PMID: 15494745

Biomass components in yeast

(based on cellular composition)

1) Amino acids
2) Nucleotides
3) Carbohydrates
4) Lipids, sterols and fatty acids

Table 9. Cellular components of *S. cerevisiae*

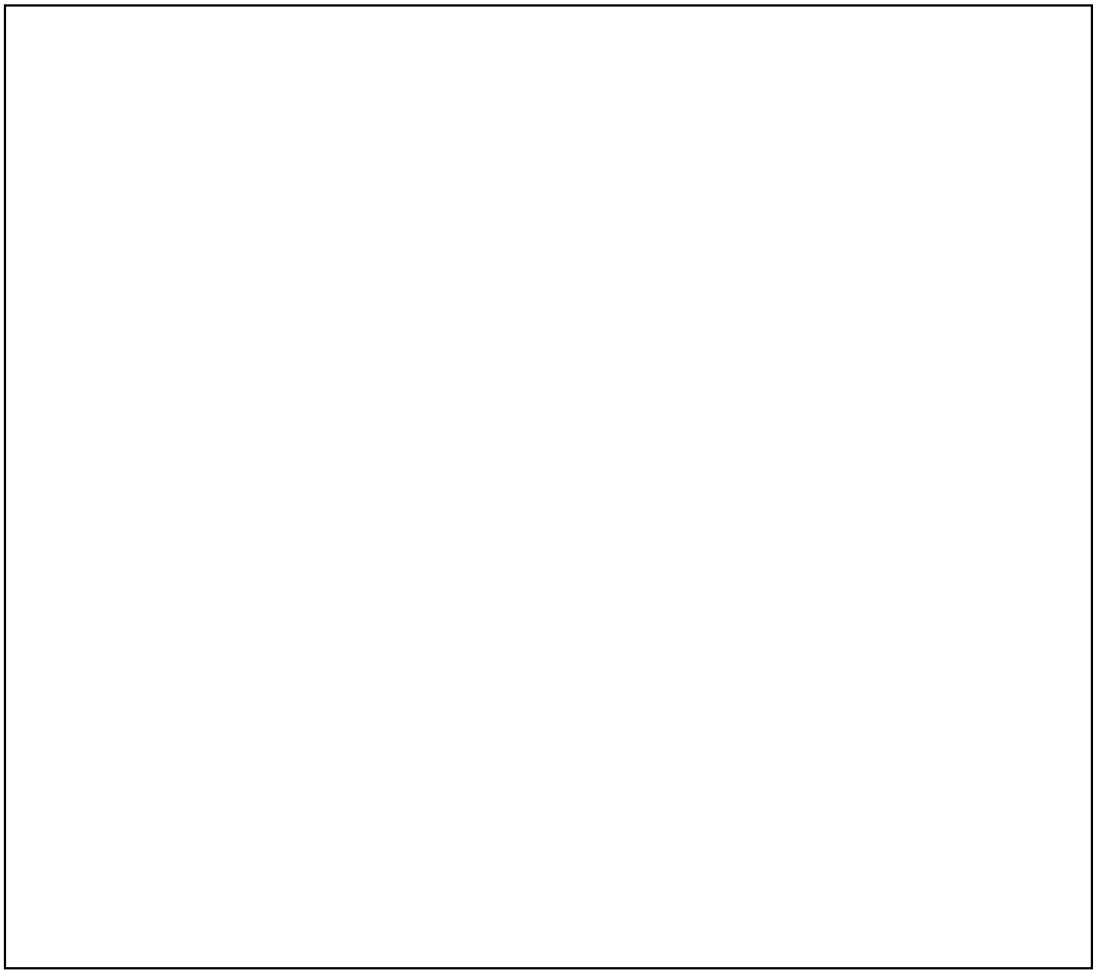
ALA	0.459	CMP	0.05
ARG	0.161	dAMP	0.0036
ASN	0.102	dCMP	0-0024
ASP	0.297	dGMP	0.0024
CYS	0.007	DTMP	0.0036
GLU	0.302	TAGLY	0.007
GLN	0.105	ERGOST	0.0007
GLY	0.290	ZYMST	0.015
HIS	0.066	PA	0.0006
ILE	0.193	PINS	0.005
LEU	0.296	PS	0.002
LYS	0.286	PE	0.005
MET	0.051	PC	0.006
PHE	0.134	GLYCOGEN	0.519
PRO	0.165	TRE	0.023
SER	0.185	Mannan	0.809
THR	0.191	13GLUCAN	1.136
TRP	0.028	SLF	0.02
TYR	0.102	ATP	23.9166
VAL	0.265	ADP	23.9166
AMP	0.051	PI	23.9456
GMP	0.051	Biomass	1
UMP	0.067		

Shulze 1995, Forster et al. 2003

What is Flux Balance Analysis good for?

Large systems can be analyzed (hundreds of reactions):

1) Prediction of optimal steady-state flux distributions in the network (it's not necessarily the *in vivo* flux distribution!)
2) Simulate different environments
3) Simulate different genotypes (perturbations to network structure)



Major assumptions of FBA

Physiological: all metabolites are in steady state

→ quasi steady state might be a good assumption (fast reactions and high turnover of reactants)

Evolutionary: the cell has adapted to maximize the efficiency of biomass production (optimality)

→ could be valid for certain microbes only (but not for multicellulars!)

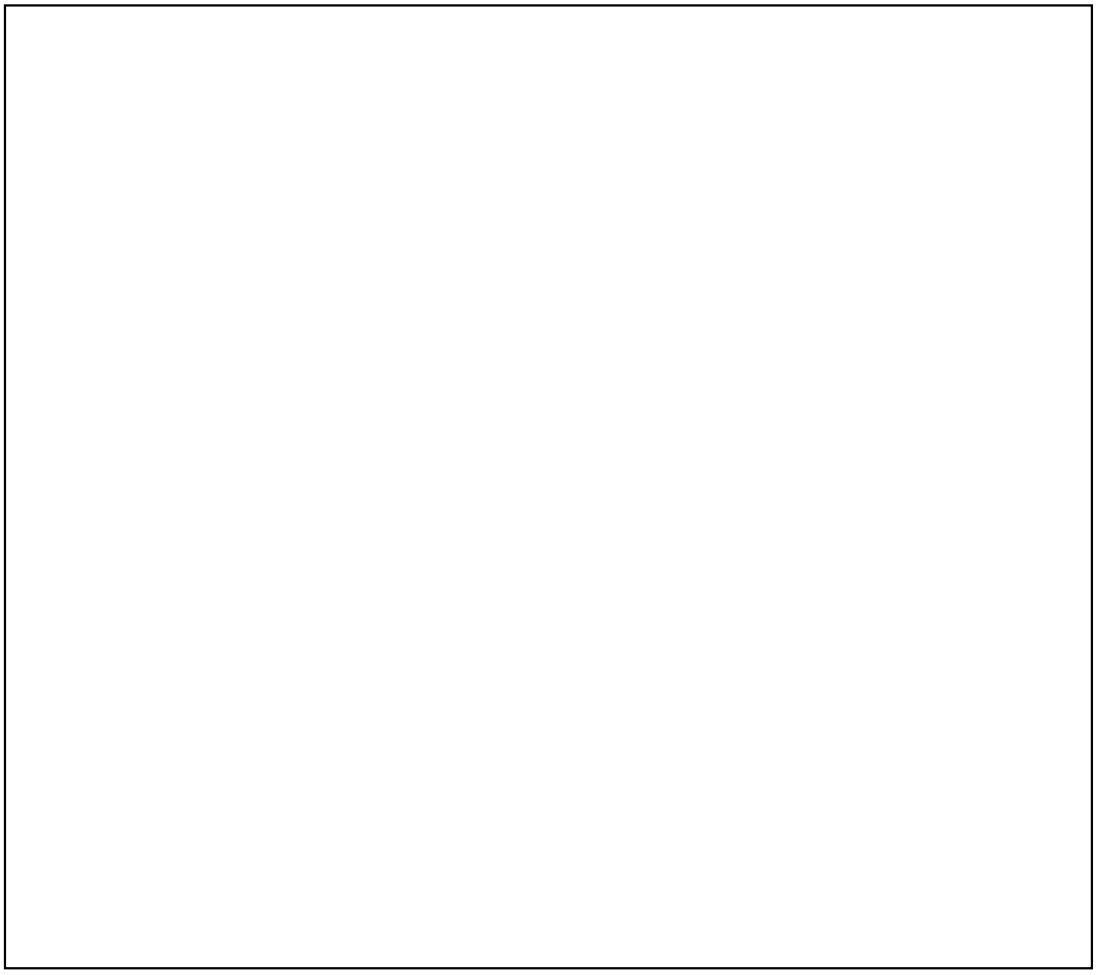
Limitations of FBA

1) Cannot track the dynamics of the system or determine the metabolite concentrations

2) Data not incorporated:

- enzyme concentrations
- mechanistic details on enzyme regulation (but gene regulation can be incorporated as further constraints)

3) The assumption of optimality (but subtopimal phenotypic states can also be investigated)



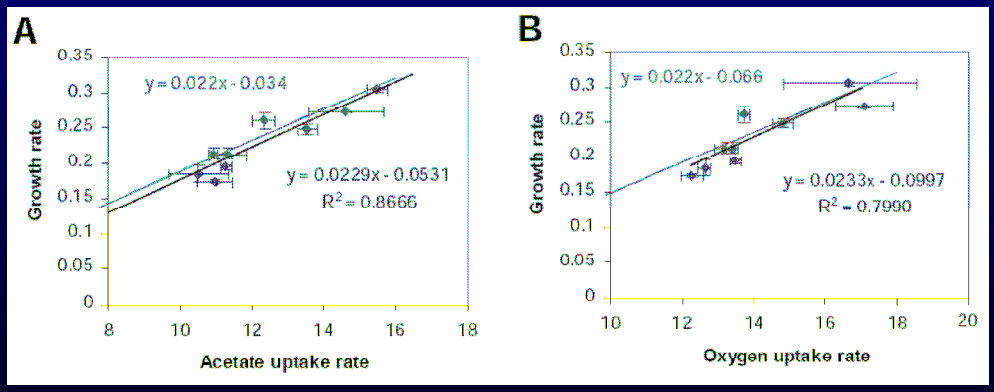
So, how good is this modelling framework?

→ need to compare FBA predictions with experimental data…

Testing predictions of FBA I.

Growth properties

Example: acetate and oxygen uptake rates in *E. coli* *



red line: prediction, black line: linear regression through the experimental data points

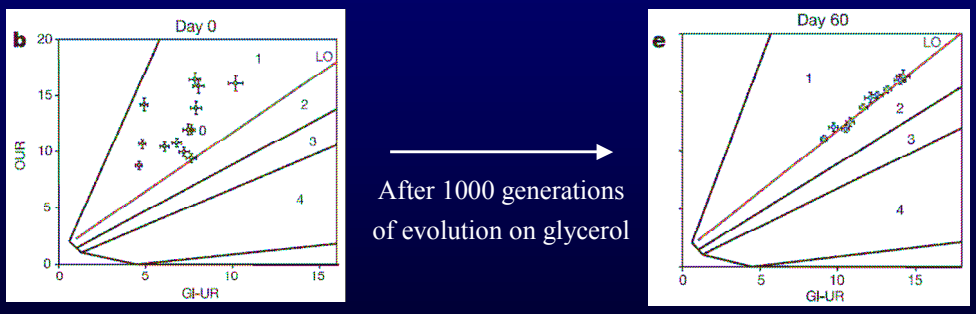
* Edwards et al . (2001) Nat Biotech 19: 125

Edwards JS, Ibarra RU, Palsson BO. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data.
Nat Biotechnol. 2001 Feb;19(2):125-30.
PMID: 11175725

Testing predictions of FBA I.

Growth properties

BUT: growth of *E. coli* is not optimal on glycerol*



Validity of the optimality assumption depends on the evolutionary history of the strain!

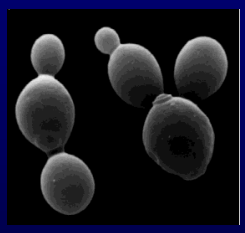
* Ibarra et al . (2002) Nature 420: 186

Ibarra RU, Edwards JS, Palsson BO. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. Nature. 2002 Nov 14;420(6912):186-9.
PMID: 12432395

Testing predictions of FBA II.

Gene essentiality

Single gene deletions were investigated both *in silico* and *in vivo* in yeast[1,2]:



→ FBA model of yeast metabolism predicts gene deletion phenotypes (viable / lethal) with 81 – 89% accuracy[2]

→ Assumption of optimal growth in mutants is problematic. Therefore other optimality criteria have been suggested[3].

1) Forster et al. (2003) OMICS 7:193
2) Kuepfer et al. (2005) Genome Res 15: 1421
3) Segre et al. (2002) PNAS 99: 15112

Forster J, Famili I, Palsson BO, Nielsen J. Large-scale evaluation of in silico gene deletions in Saccharomyces cerevisiae.
OMICS. 2003 Summer;7(2):193-202.
PMID: 14506848

Kuepfer L, Sauer U, Blank LM. Metabolic functions of duplicate genes in Saccharomyces cerevisiae.
Genome Res. 2005 Oct;15(10):1421-30.
PMID: 16204195

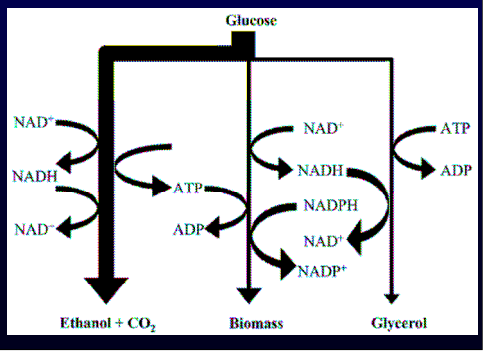
Segre D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks.
Proc Natl Acad Sci U S A. 2002 Nov 12;99(23):15112-7.
PMID: 12415116

Application of FBA I. Bioengineering

Q: How to engineer the metabolism of a microbe to improve the production of certain compounds?

Improving ethanol production in yeast:

- Ethanol is the largest fermentation product (10^9 $ annual sale)
- produced by anaerobic fermentations with *S. cerevisiae*
- production could be increased by redirecting carbon flow to glycerol towards ethanol



Related applications in biotechnology:

Burgard AP, Maranas CD. Probing the performance limits of the Escherichia coli metabolic network subject to gene additions or deletions.
Biotechnol Bioeng. 2001 Sep 5;74(5):364-75.
PMID: 11427938

Burgard AP, Pharkya P, Maranas CD. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization.
Biotechnol Bioeng. 2003 Dec 20;84(6):647-57.
PMID: 14595777

Fong SS, Burgard AP, Herring CD, Knight EM, Blattner FR, Maranas CD, Palsson BO. In silico design and adaptive evolution of Escherichia coli for production of lactic acid.
Biotechnol Bioeng. 2005 Sep 5;91(5):643-8.
PMID: 15962337

Application of FBA I. Bioengineering

There can be alternative strategies to redirect glycerol flux by engineering redox metabolism.

But which one is the most efficient?

A computational approach to find the best strategy:

1) Take a database of reactions not found in yeast (~3800 reactions from different species)

2) Use yeast FBA model to assess the effect of inserting these reactions one at a time

3) Identify reactions that improve both growth and ethanol production

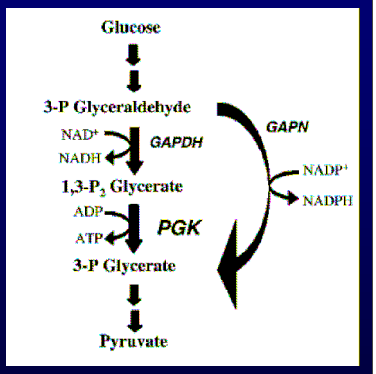
Bro et al . (2006) Metab Engineering 8: 102

Bro C, Regenberg B, Forster J, Nielsen J. In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production.
Metab Eng. 2006 Mar;8(2):102-11.
PMID: 16289778

Application of FBA I. Bioengineering

One of the best strategies according to the model: inserting the gene for GAPN, which substitutes production of glycerol with production of ethanol

→ Theoretically, this could increase ethanol production by 10%



Experimental test: *gapN* gene from *Streptococcus mutans* was expressed in yeast

→ 3% increase in ethanol production *in vivo*

Bro et al . (2006) Metab Engineering 8: 102

Bro C, Regenberg B, Forster J, Nielsen J. In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production.
Metab Eng. 2006 Mar;8(2):102-11.
PMID: 16289778

GAPN: non-phosphorylating, $NADP^+$-dependent glyceraldehyde-3-P dehydrogenase

Applications of FBA II. Network evolution

Q: How new enzymes are added to the network during evolution?

1) Gene duplication
2) Horizontal gene transfer (HGT): acquisition of genes from other species

   → Can be frequent among bacteria, but rare in eukaryotes

What is the advantage of acquiring enzymes via HGT?

Which enzymes are most prone to HGT?

Related applications in evolutionary genetics:

Papp B, Pal C, Hurst LD. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast.
Nature. 2004 Jun 10;429(6992):661-4.
PMID: 15190353

Blank LM, Kuepfer L, Sauer U. Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast.
Genome Biol. 2005;6(6):R49.
PMID: 15960801

Segre D, Deluna A, Church GM, Kishony R. Modular epistasis in yeast metabolism.
Nat Genet. 2005 Jan;37(1):77-83.
PMID: 15592468

Pal C, Papp B, Lercher MJ, Csermely P, Oliver SG, Hurst LD. Chance and necessity in the evolution of minimal metabolic networks.
Nature. 2006 Mar 30;440(7084):667-70.
PMID: 16572170

Applications of FBA II. Network evolution

Hypothesis: enzymes acquired via HGT enable adaptation to new environments (i.e. not housekeeping)

Prediction: HGT enzymes should have environment-specific growth contributions

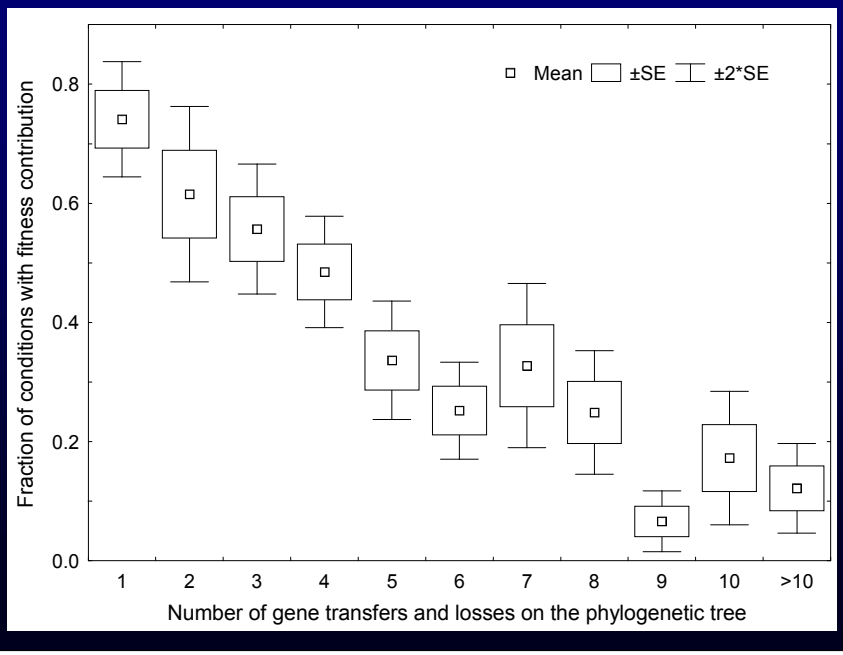
Test: use FBA to simulate the growth effects of gene deletions in *E. coli* under a large number of environments

Pál et al . (2005) Nat Genet 37: 1372

Pal C, Papp B, Lercher MJ. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer.
Nat Genet. 2005 Dec;37(12):1372-5.

PMID: 16311593

Horizontally transferred enzymes have growth contributions only under a small number of environments



Pal C, Papp B, Lercher MJ. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer.
Nat Genet. 2005 Dec;37(12):1372-5.

PMID: 16311593

Topics covered

- Constraint-based approach to metabolic modelling
- Principles of FBA and some biological applications
- **Mathematics behind FBA: Optimisation**
- Mathematical examples

Optimisation and Mathematical Programming

- optimisation problem or mathematical programming problem: a formulation in which a function is minimised by systematically choosing the values of variables from within an allowed set

> Given a function f: $A \rightarrow \mathbf{R}$ (*e.g.* min x^2+1)
>
> Find an element x_0 in A such that $f(x_0) \leq f(x)$ for all x in A

- The domain A of f is called the ***search space***, while the elements of A are called ***feasible solutions***
- A is specified by a set of ***constraints*** (equalities or inequalities)
- function f is called an ***objective function***
- A feasible solution that minimizes the objective function is called an ***optimal solution***

Subfields

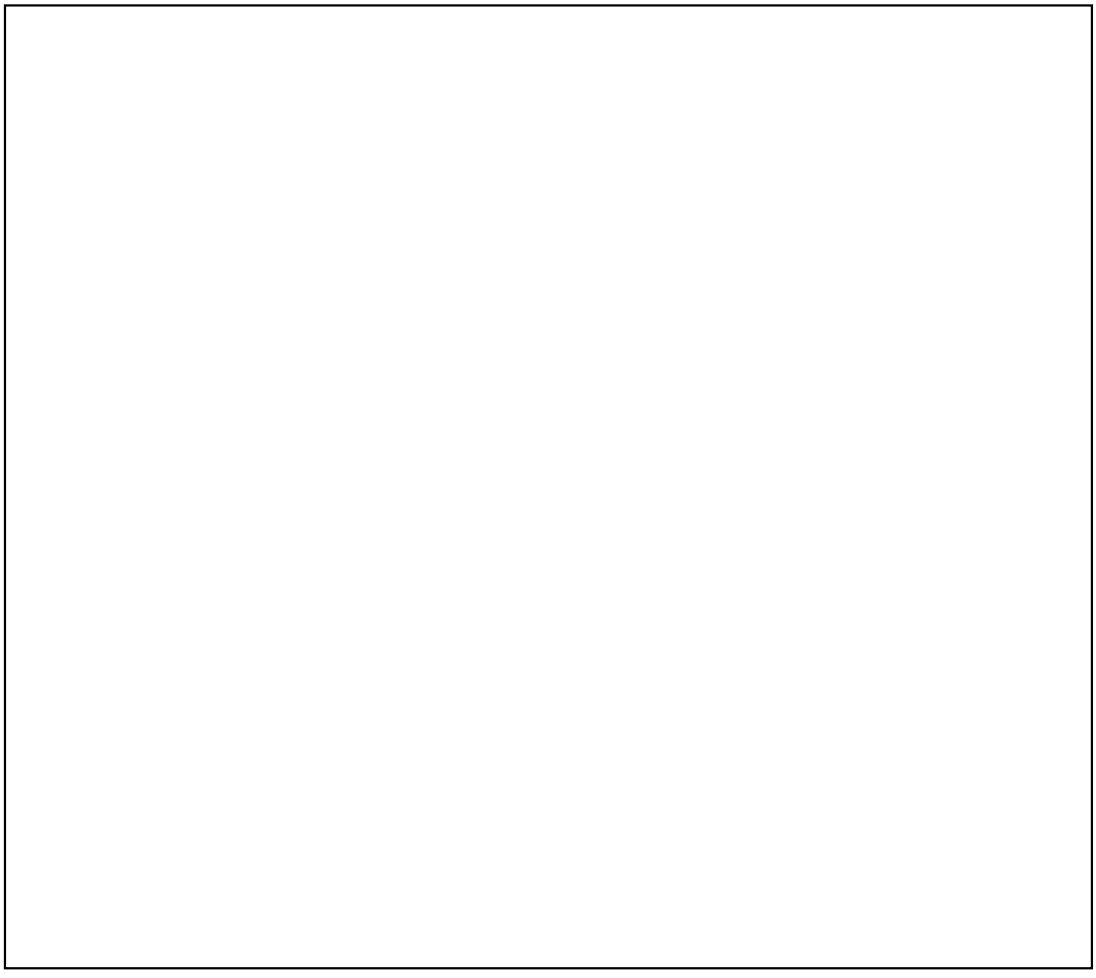
- *Linear programming* studies the case in which the objective function f is linear and the constraints linear equalities and inequalities
- *Integer linear programming* studies linear programs in which some or all variables take on integer values
- *Nonlinear programming* studies the general case in which the objective function or the constraints or both are nonlinear

Techniques for solving mathematical programming problems

- There exist robust, fast numerical techniques for optimising mathematical programming problems
  - Gradient descent (steepest descent)
  - Nelder-Mead method
  - Simplex method
  - Ellipsoid method
  - Newton's method
  - Quasi-Newton methods
  - Interior point methods
  - Conjugate gradient method

Alternatives for optimisation

- Mathematical programming and its techniques for solving of optimisation problems are powerful tools, but are not the only solutions available. Other approaches (that usually apply numerical analysis approximations) are:
  - Hill climbing
  - Simulated annealing
  - Quantum annealing
  - Tabu search
  - Beam search
  - Genetic algorithms
  - Ant colony optimization
  - Evolution strategy
  - Stochastic tunneling
  - Particle swarm optimization
  - Differential evolution



Linear Programming (LP)

- LP model → objective function + linear constraints
- extensively used optimisation technique
- allocation of limited resources to competing activities in the optimal way
- examples of application: graphs, network flows, plant management, economics, business management
- most prominent method for solving: simplex method
- Prominent solver: CPLEX

minimise $c_1 \cdot x_1 + c_2 \cdot x_2 + \ldots + c_n \cdot x_n$

subject to:

linear constraints

$a_{i1} \cdot x_1 + a_{i2} \cdot x_2 + \ldots + a_{in} \cdot x_n = a_{i0}$

nonlinear constraints

$b_{i1} \cdot x_1 + b_{i2} \cdot x_2 + \ldots + b_{in} \cdot x_n \leq b_{i0}$

or in matrix form:

$$\min \quad c^T \cdot x$$

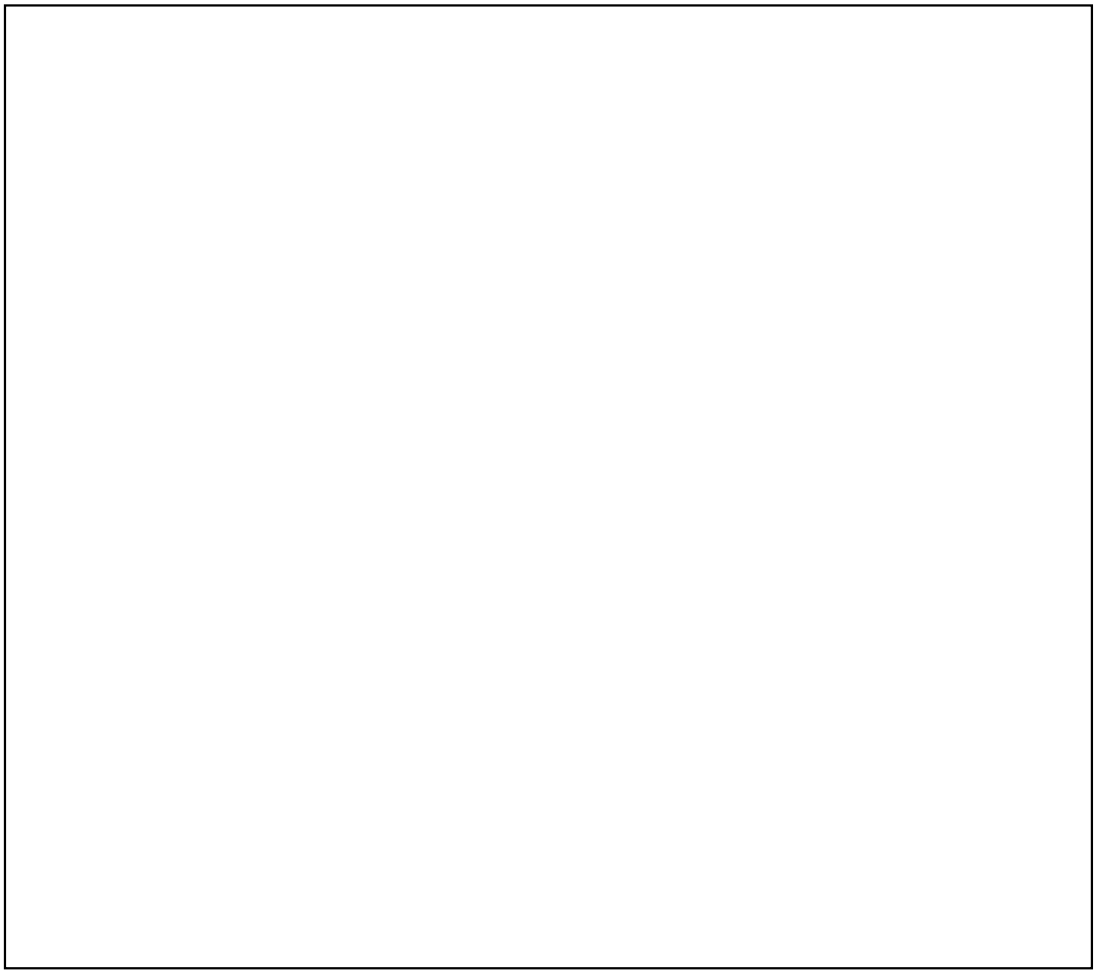
subject to:

$$A \cdot x = a$$

$$B \cdot x \leq b$$

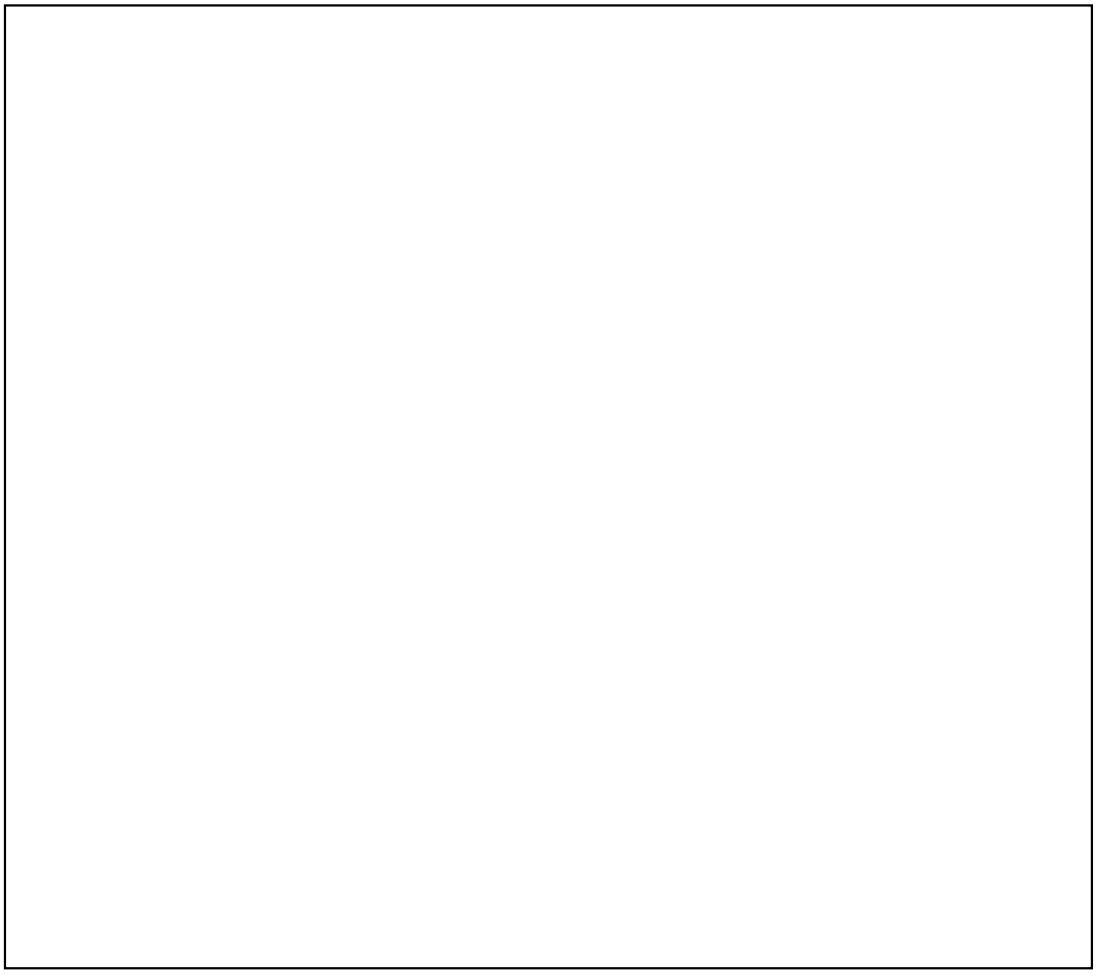
Integer Programming

- If variables are required to be integer, then the problem is an integer programming (IP) or mixed integer programming (MIP) problem
- In contrast to linear programming, which can be solved efficiently in the worst case, integer programming problems are in the worst case undecidable, and in many practical situations NP-hard
- MIP problems are solved using advanced algorithms such as branch and bound or branch and cut
- LP and MILP solvers are in widespread use for optimization of various problems in industry, such as optimization of flow in transportation networks
  - CPLEX
  - MINTO
  - AIMMS
  - SYMPHONY
  - Xpress-MP
  - GNU Linear Programming Kit
  - Qoca
  - Cassowary constraint solver



Software Applications

- General Algebraic Modelling System (GAMS)
  - consists of a language compiler and a number of integrated high-performance solvers for mathematical programming models
  - tailored for complex, large-scale modelling applications
- CPLEX solver (standalone)
- Simpheny (http://genomatica.com/solutions_simpheny.shtml)



Optimisation in FBA

• optimisation is used to predict metabolic flux distributions at steady state based on the assumption of maximised growth performance along evolution

• only stoichiometric data and cellular composition required

• valuable for identifying flux distribution boundaries for the metabolic function of cellular systems

• Linear Programming may be used to study the stoichiometric constraints on metabolic networks

Application

- FBA involves carrying out a steady state analysis, using the stoichiometric matrix (*S*) for the system in question
- The system is assumed to be optimised with respect to objectives such as maximisation of biomass production or minimisation of nutrient utilisation
- At steady state: $$\frac{dx}{dt} = S \cdot v = 0$$
- The required flux distribution is the null space of *S*. Since the number of fluxes typically exceeds the number of metabolites, the system is under-determined and may be solved by selecting an optimisation criterion, following which, the system translates into an LP problem

Mathematical Model

$$\max \sum_j c_j \cdot v_j$$

$$\text{s.t.} \quad \sum_j S_{ij} \cdot v_j = 0 \qquad \forall i$$

$$L_j \le v_j \le U_j \qquad \forall j_{reversible}$$

$$0 \le v_j \le U_j \qquad \forall j_{irreversible}$$

i: metabolites
j: reactions
S_{ij}: stoichiometric matrix
v_j: reaction fluxes (*mmol / gDW hr*)
c_j: weight
L_j, U_j: bounds

Small example for *S* matrix construction

$A \leftrightarrow B$, which is equivalent to:

$A \rightarrow B$ and $B \rightarrow A$

$A + B \rightarrow C$

$B + C \rightarrow 2\,A$

$$\begin{array}{c} \begin{matrix} j_1 & j_2 & j_3 & j_4 \end{matrix} \\ \begin{bmatrix} -1 & 1 & -1 & 2 \\ 1 & -1 & -1 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{matrix} A \\ B \\ C \end{matrix} \end{array}$$

- Maximise the sum of a particular reaction flux, or some combination of fluxes in order to accomplish a goal

Topics covered

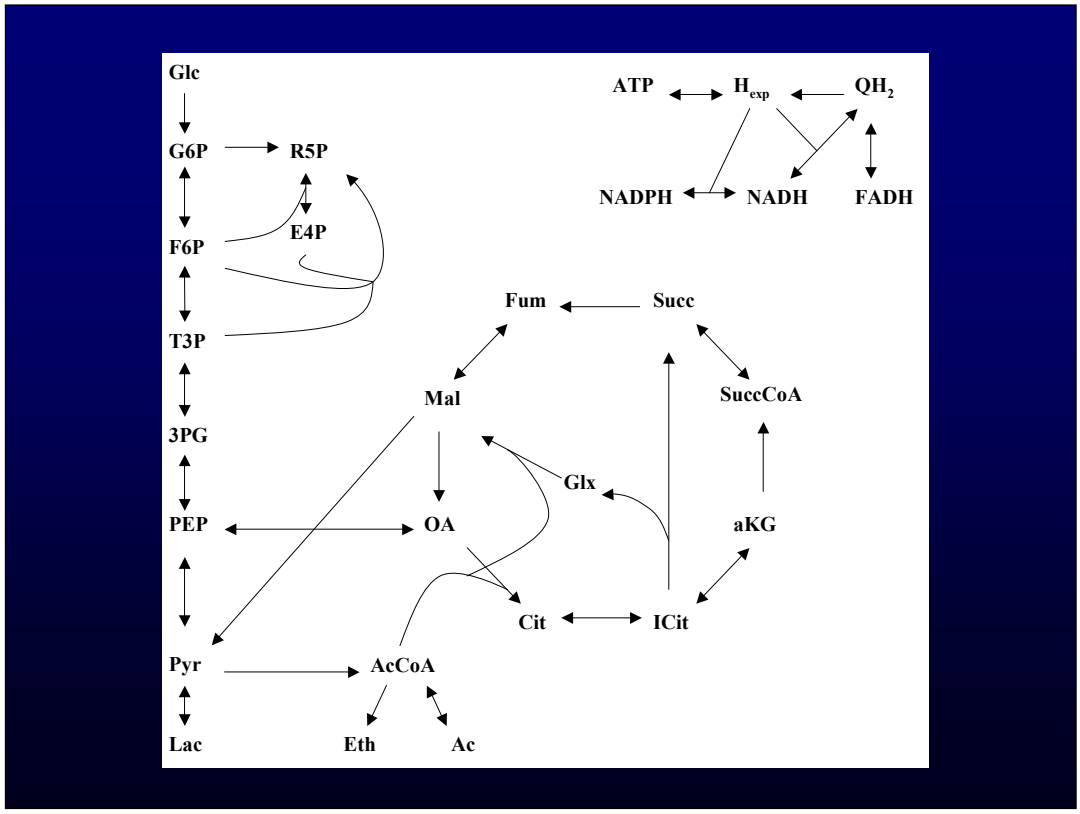
- Constraint-based approach to metabolic modelling
- Principles of FBA and some biological applications
- Mathematics behind FBA: optimisation
- **Mathematical examples**

Model Network

- Simplified model of *E. coli* metabolic network

  Varma and Palsson, J.Theor.Biol. (1993) 165, 477

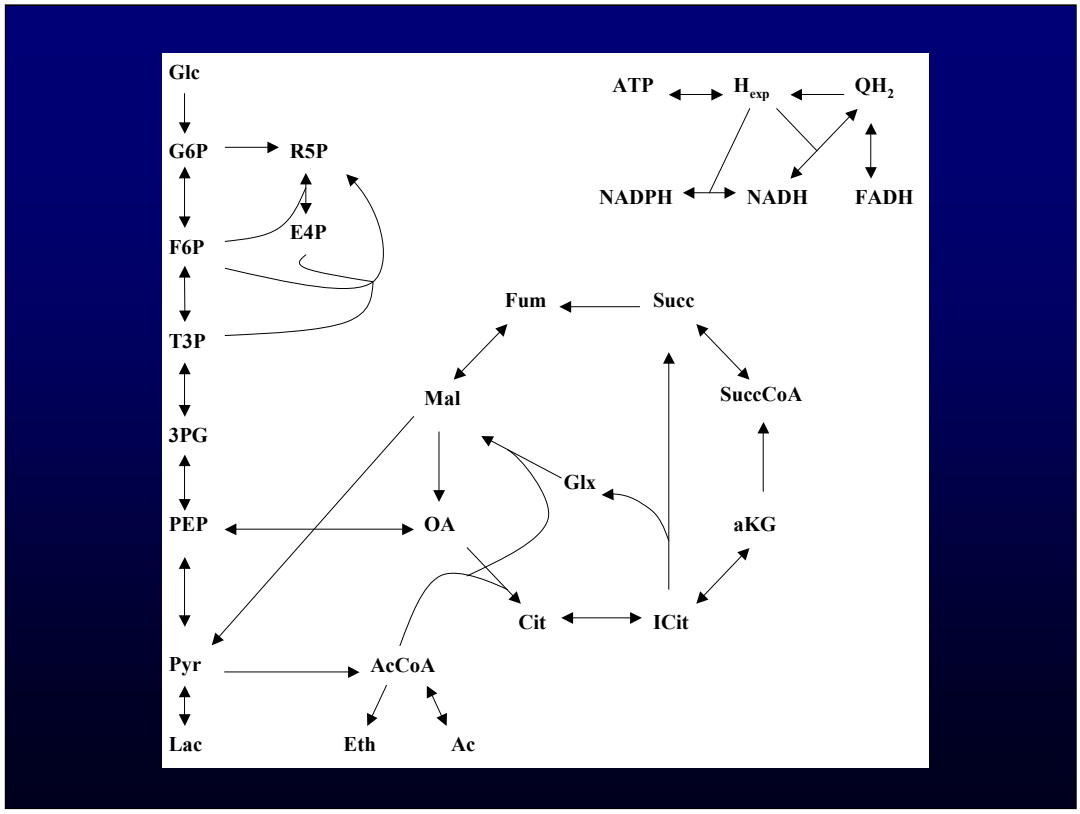
- Reactions: 35
  - 17 reversible
  - 18 irreversible
- Metabolites: 30


Glc
G6P
R5P
E4P
F6P
T3P
3PG
PEP
Pyr
Lac
AcCoA
Eth
Ac
Mal
OA
Fum
Succ
SuccCoA
Glx
aKG
Cit
ICit
ATP
H_{exp}
QH_2
NADPH
NADH
FADH

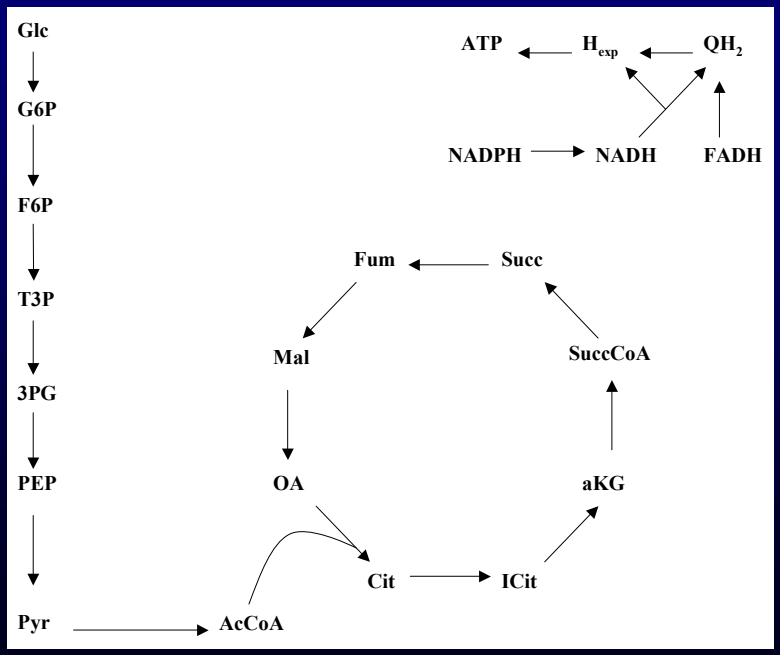
Single metabolite production maximisation

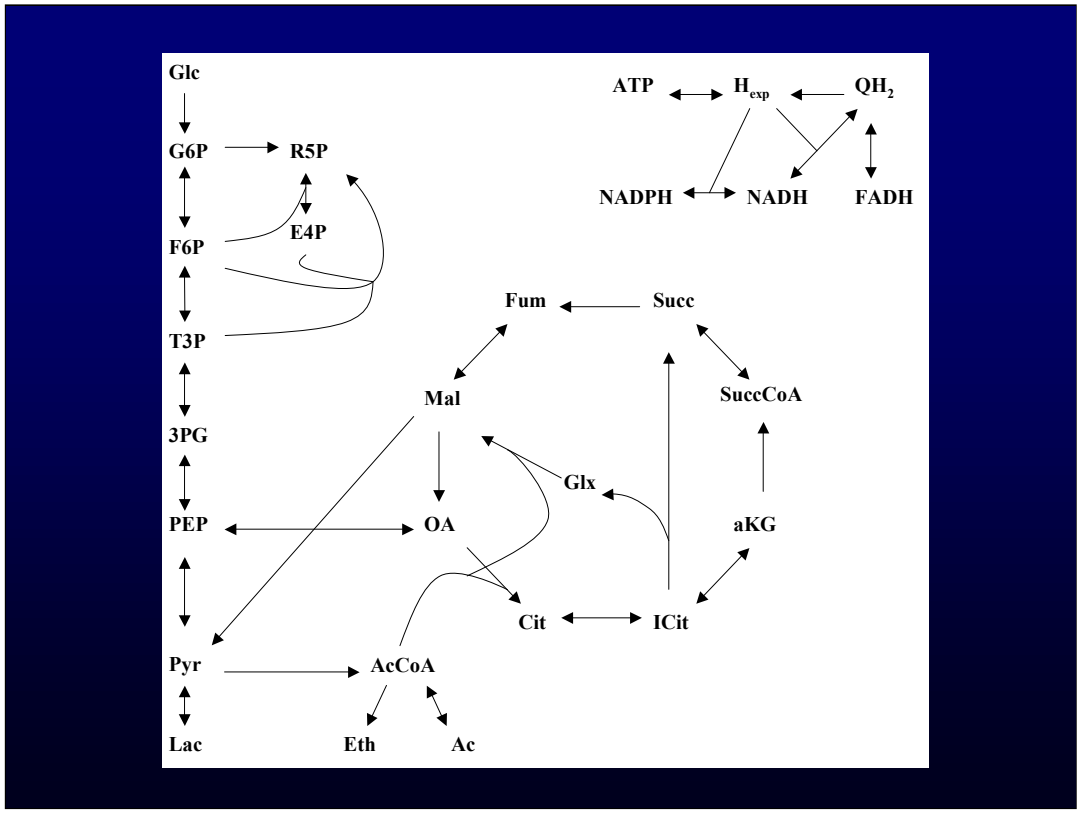
Maximum stoichiometric yields (*mmol/mmol* Glc)

ATP	18.67	**F6P**	0.91
NADH	11.57	**R5P**	1.08
NADPH	11.00	**E4P**	1.33
3PG	2.00	**T3P**	1.74
PEP	2.00	**AcCoA**	2.00
Pyr	2.00	**aKG**	1.00
OA	1.50	**SuccCoA**	1.00
G6P	0.91		

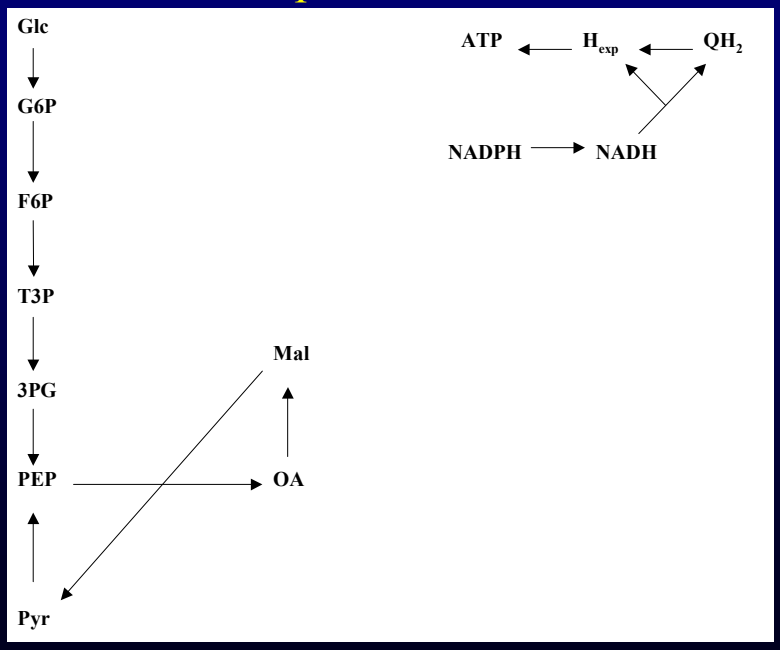

Glc
G6P
R5P
E4P
F6P
T3P
3PG
PEP
Pyr
Lac
AcCoA
Eth
Ac
Mal
OA
Fum
Succ
SuccCoA
Glx
aKG
Cit
ICit
ATP
H_{exp}
QH_2
NADPH
NADH
FADH

Flux distribution map for maximal ATP production




Glc
G6P
R5P
E4P
F6P
T3P
3PG
PEP
Pyr
Lac
Eth
Ac
AcCoA
Mal
OA
Fum
Succ
SuccCoA
Glx
aKG
Cit
ICit
ATP
H_{exp}
QH_2
NADPH
NADH
FADH

Flux distribution map for maximal PEP production



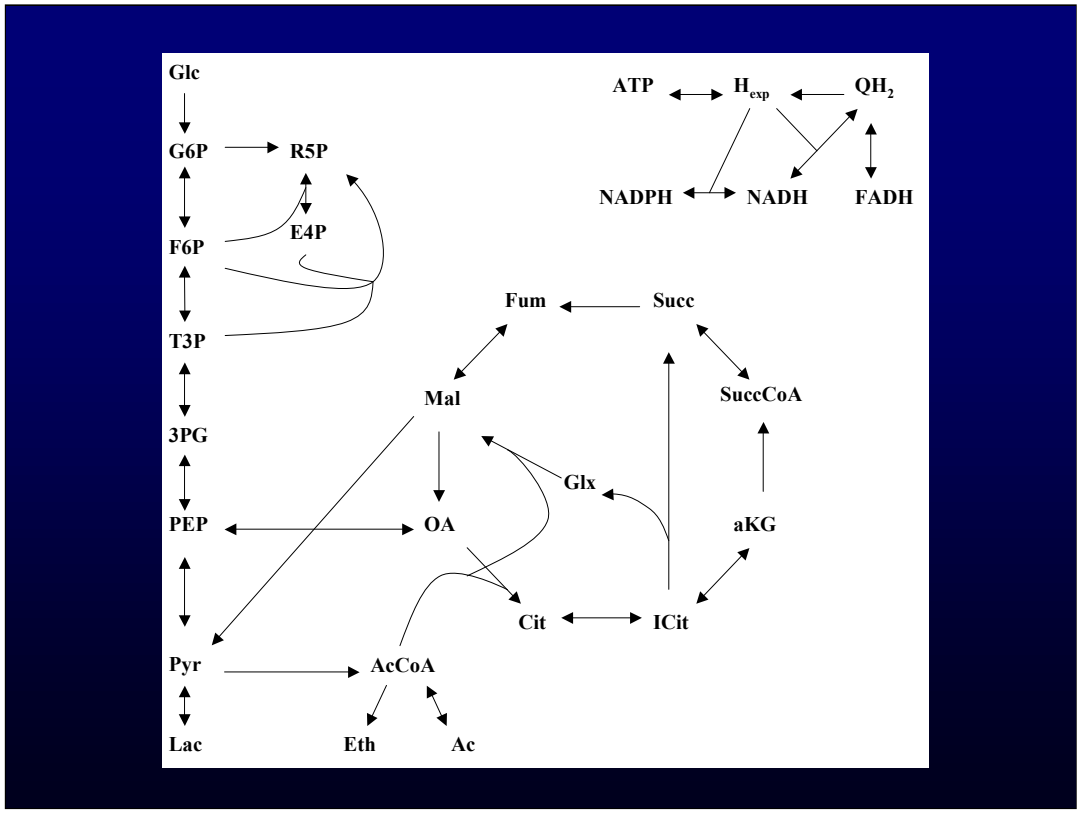
Biomass production maximisation

Metabolic Demands for 1g of biomass yield (*mmol*)

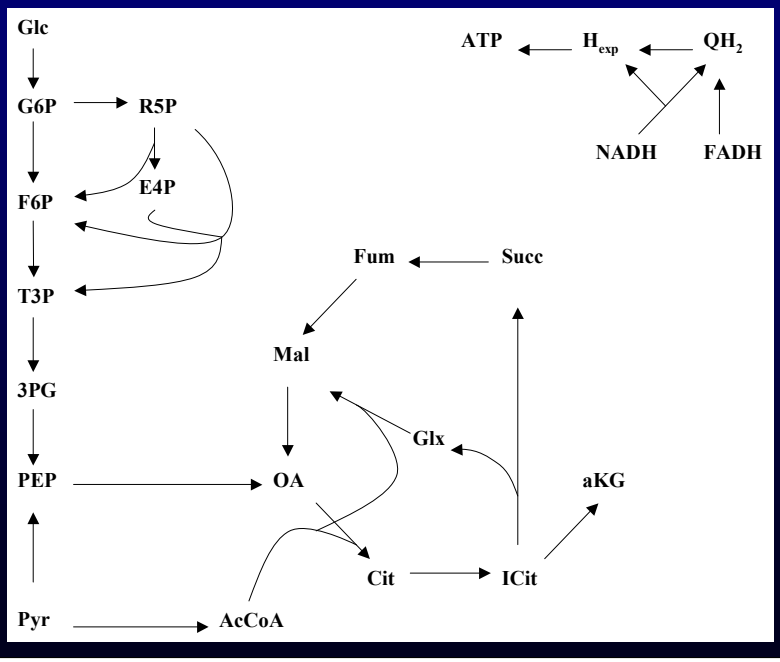
ATP	41.2570	**T3P**	0.1290
NADH	-3.5470	**3PG**	1.4960
NADPH	18.2250	**PEP**	0.5191
G6P	0.2050	**Pyr**	2.8328
F6P	0.0709	**AcCoA**	3.7478
R5P	0.8977	**OA**	1.7867
E4P	0.3610	**aKgG**	1.0789

Maximum biomass yield:

0.589 *g DW / g* Glc


Glc
G6P
R5P
E4P
F6P
T3P
3PG
PEP
Pyr
Lac
AcCoA
Eth
Ac
Mal
OA
Fum
Succ
SuccCoA
Glx
aKG
Cit
ICit
ATP
H_{exp}
QH_2
NADPH
NADH
FADH

Optimal flux distribution for aerobic growth on glucose



Finding minimal reaction sets

gene

↓

enzyme

↓

reaction

1. What is the smallest gene set capable of maximising biomass?
2. What is the maximum number of gene deletions that still maintains biomass production above certain levels?

- MILP problem

Mathematical model

$$\min \sum_k y_k$$

$$\text{s.t.} \quad \sum_j S_{ij} \cdot v_j = 0 \qquad \forall i$$

$$L_j \cdot \sum_k a_{jk} \cdot y_k \le v_j \le U_j \cdot \sum_k a_{jk} \cdot y_k \qquad \forall j$$

$$v_{biomass} \ge v_{target}$$

a_{jk}: 1 if gene k codes for an enzyme catalysing reaction j, 0 otherwise
k: genes
y_k: 1 if gene k is present and functional, 0 otherwise
L_j, U_j: bounds

Results

% max biomass	removals
100 %	12
90 %	19
70 %	20
50 %	21
30 %	22
20 %	24
10 %	24
1 %	24

- 23 out of 35 reactions are required to sustain optimal growth
- Small tolerance, albeit lessened biomass demands
- 25 removals render the network incapable of biomass formation

FBA

- Simple, no kinetic information needed
- Can be applied to large networks
- In accordance with experimental results
- Can be used for defining wider limits of metabolic behaviour

Further reading

Bernhard O. Palsson: Systems Biology: Properties of reconstructed networks, Cambridge University Press, 2006

Metabolic network reconstruction:
Nature Reviews Genetics (2006) **7**:130-141.

Metabolic modelling approaches:
Journal of Biotechnology (2002) **94**: 37-63.
Biotechnology and Bioengineering (2003) **84:** 763-772

Flux Balance Analysis:
Current Opinion in Biotechnology (2003) **14**: 491-496.
Nature Reviews Microbiology (2004) **2**: 886-897

Mathematical programming:
Paul Williams: Model Building in Mathematical Programming