Metabolic Networks

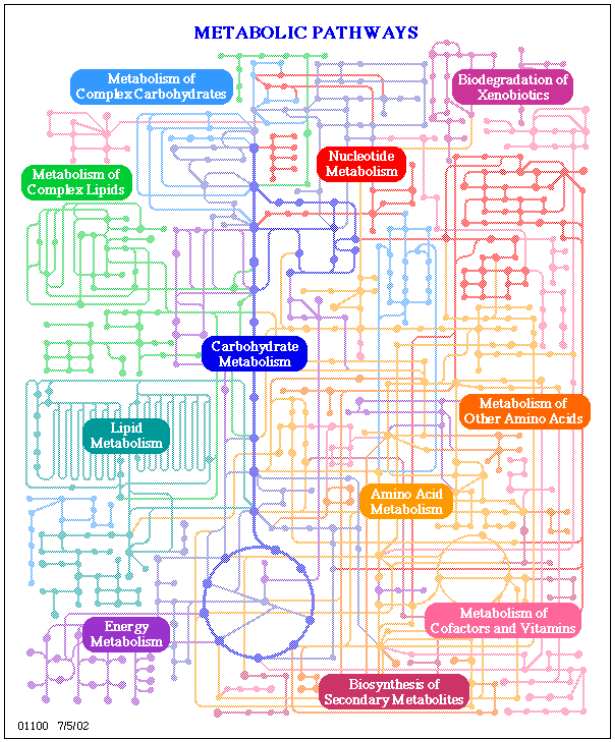
Ulf Leser and Michael Weidlich

This Lecture

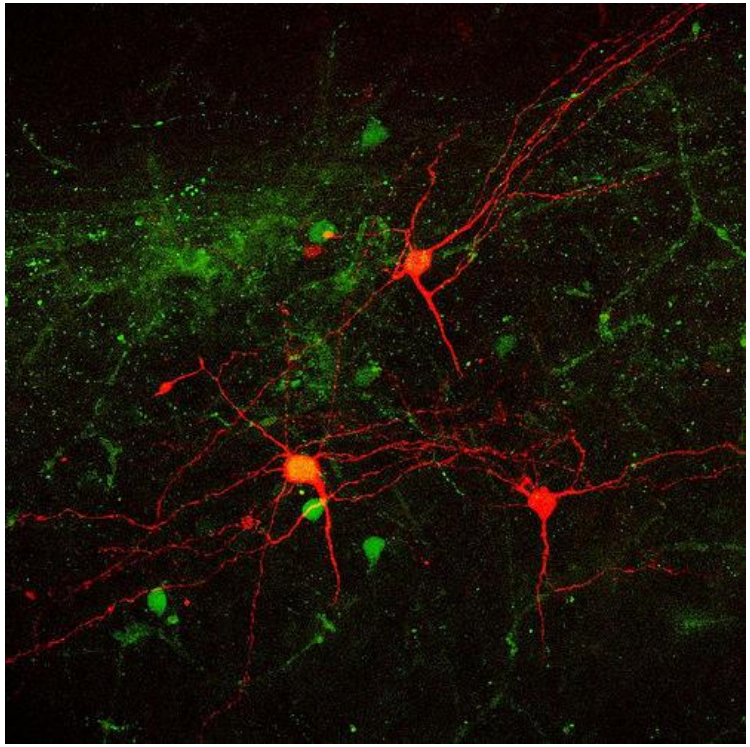
- Introduction
  - Systems biology & modelling
  - Metabolism & metabolic networks
- Network reconstruction
  - Strategy & workflow
- Mathematical representation
  - The stoichiometric matrix
  - Convex analysis & solution space
  - Elementary Modes & Extreme Pathways
- Constraint based network analysis
  - Flux Balance Analysis & Optimization
  - Tools

Motivation

Complex systems in biology



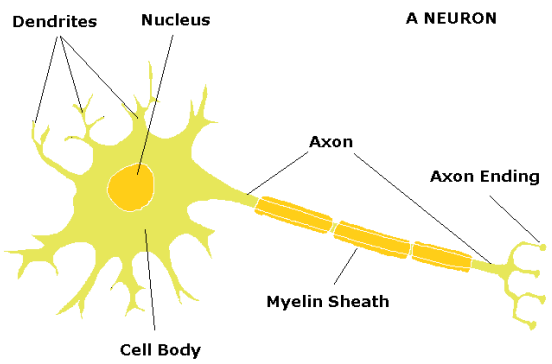
Metabolic network



Neural network

„More than the sum of its parts“

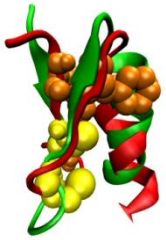
- Many biological systems have emergent properties (they are called irreducible since they cannot be fully comprehended when broken down into smaller segments)



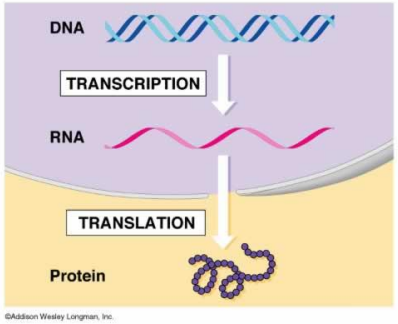
- Studying a neuron will not explain how information is stored in the brain? (fire patterns)

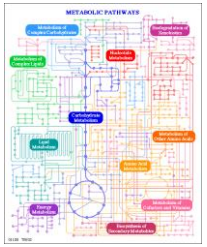
Systems biology

"Systems biology...is about putting together rather than taking apart, integration rather than reduction. It requires that we develop ways of thinking about integration that are as rigorous as our reductionist programmes, but different....It means changing our philosophy, in the full sense of the term" - Dennis Noble



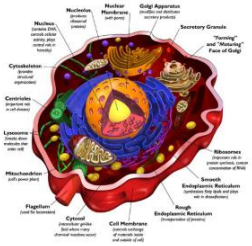
Protein structure



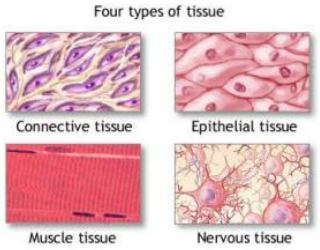


Metabolic networks

Systems biology - Bridging the scales

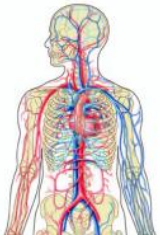


Cell



Tissue

Organ

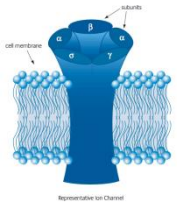


Whole body

20th century

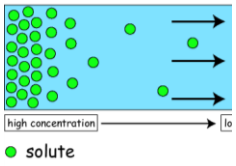
From component to systems analysis

21th century



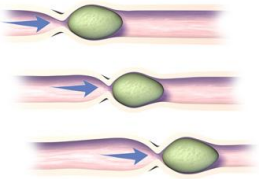
10^{-6} s

Ion channel gating



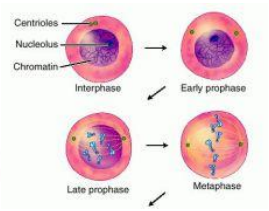
10^{-3} s

Diffusion



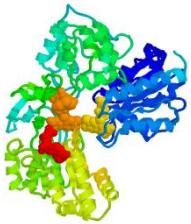
1s

Motility



10^{3} s

Mitosis



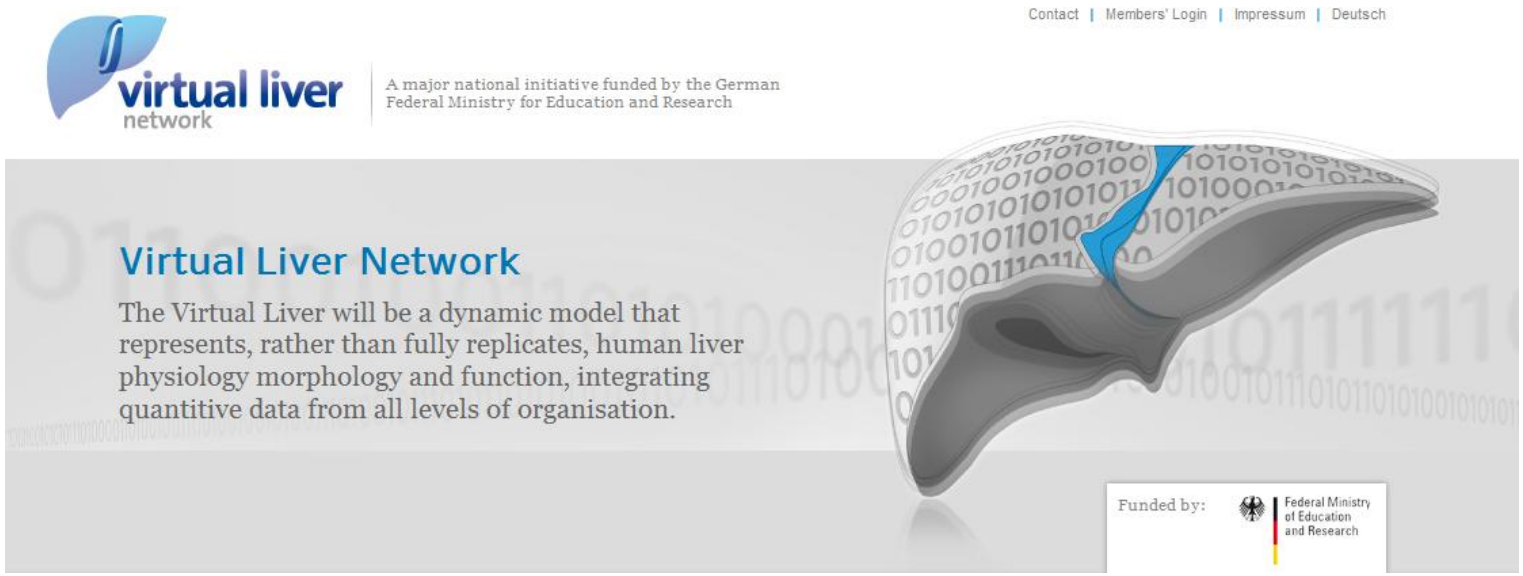
10^{6} s

Protein turnover

10^{9} s

Human lifetime

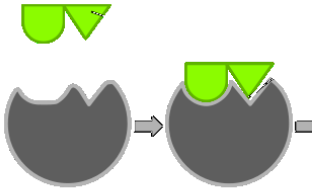
Excursion – Virtual Liver Network



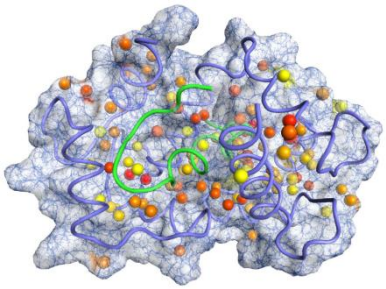
- Interdisciplinary competence network of experimental and theoretical research groups (since 2004)
- Liver - A most relevant research object for applications in medicine, pharma research and nutrition
- Example for a german systems biology initiative

What is a Model ?

vs



vs



A Model is a simplified image of reality.

- Models are able to reduce reality to certain aspects.
- Definition of these aspects is guided by our comprehension of which aspects are crucial (for functionality)
- Simplification leads to understanding of complex systems
- Models should generate testable predictions (purpose driven)

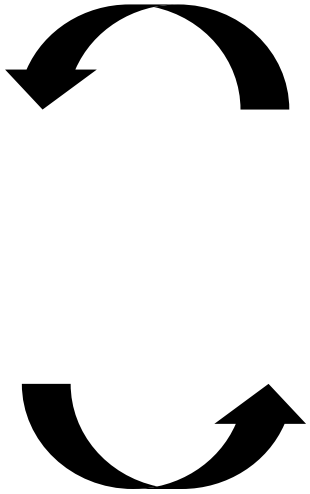
Process of modelling

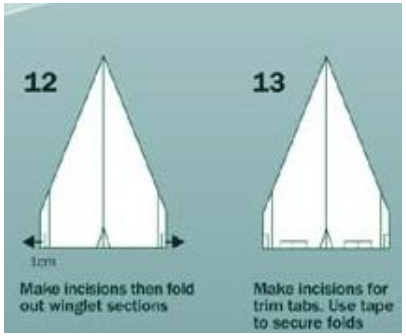
- The direct problem consists in generating data from models (i.e. doing simulations) → relativly easy
- The inverse problem consists in generating a model from data → generally difficult

- Modelling is an iterative process of both problems

Simulation

Experiment





Model

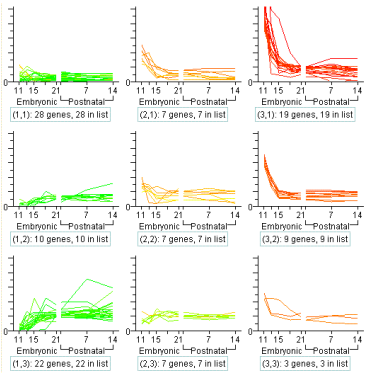
Process of modelling

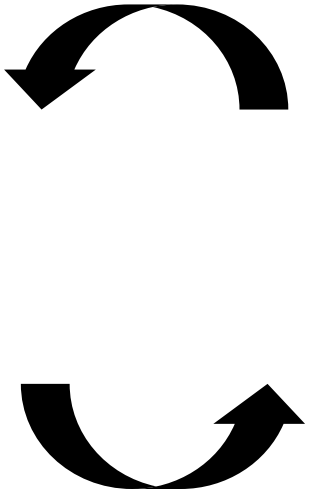
- The direct problem consists in generating data from models (i.e. doing simulations) → relativly easy
- The inverse problem consists in generating a model from data → generally difficult

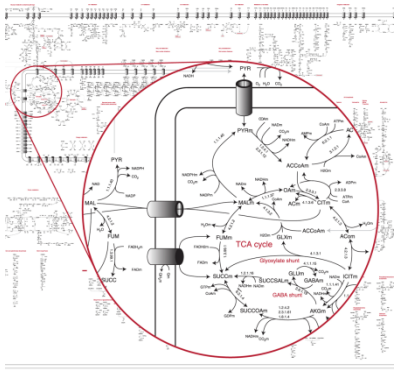
- Modelling is an iterative process of both problems

Simulation

Experiment

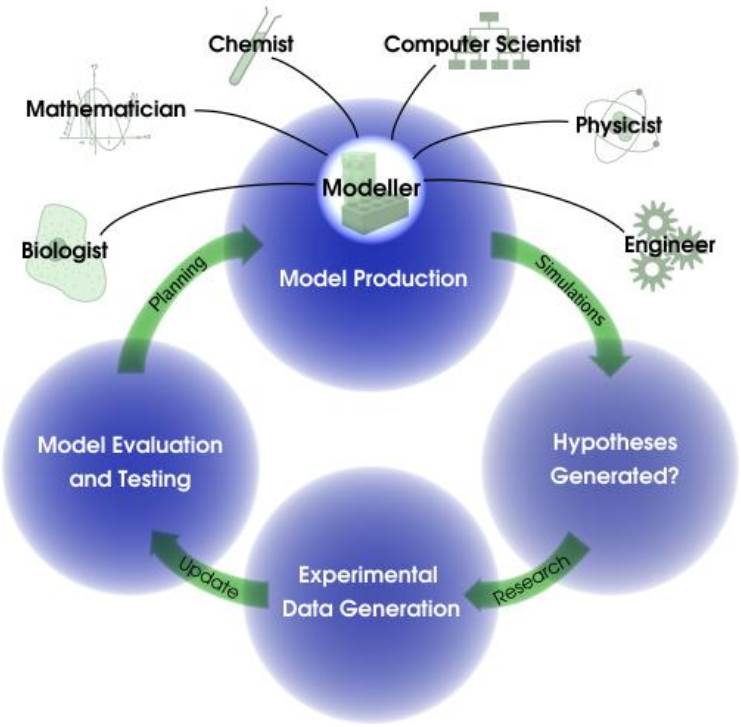






Model

Modelling - futher specified

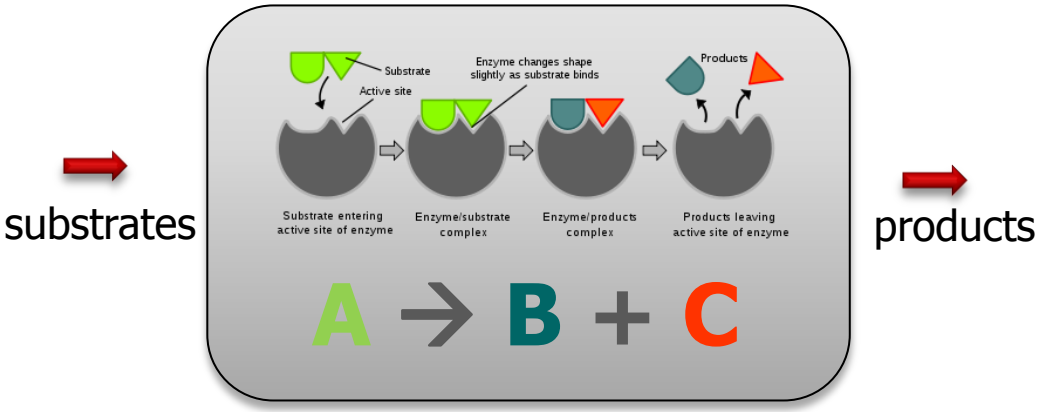


Okay i got it! What are we going to model ?

Metabolism

Metabolism

- is the set of chemical reactions that happen in living organisms to maintain life, simply spoken it is the uptake, transport and conversion of chemical entities in an organism
- These processes allow for growth and reproduction, maintenance of structures and responsiveness to environmental stress.

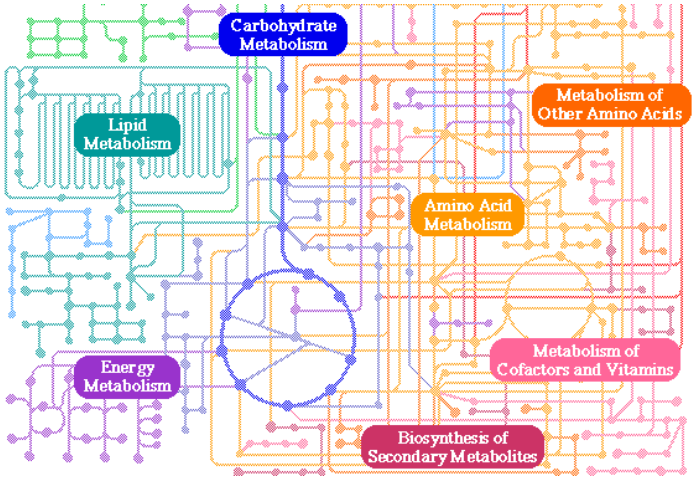


Metabolism and Disease

- Metabolism is in a balanced state which it tries to maintain
- any permanent perturbation of this balance can lead to deseases (i.e. diabetes, cancer, etc.)

Metabolism and Evolution

- Metabolism is subjected to evolutionary optimization which leads to robustness and modularity



- Redundancy - there is often more than one solution to a certain objective
- Pathways - TCA and glycolysis are shared among many different organisms

Key-players in metabolism

Chemical Compounds / Metabolites

- small molecules that are imported/exported and/or synthesized/degraded inside an organism.

Biochemical Reactions

- produce a set of metabolites (called products) from another set of compounds (called substrates)
- can be reversible or irreversible
- some are spontaneous, but most are enzymatic
- a set of reactions is often referred to as a pathway (still there is no consensus on the boundaries)

Key-players in metabolism

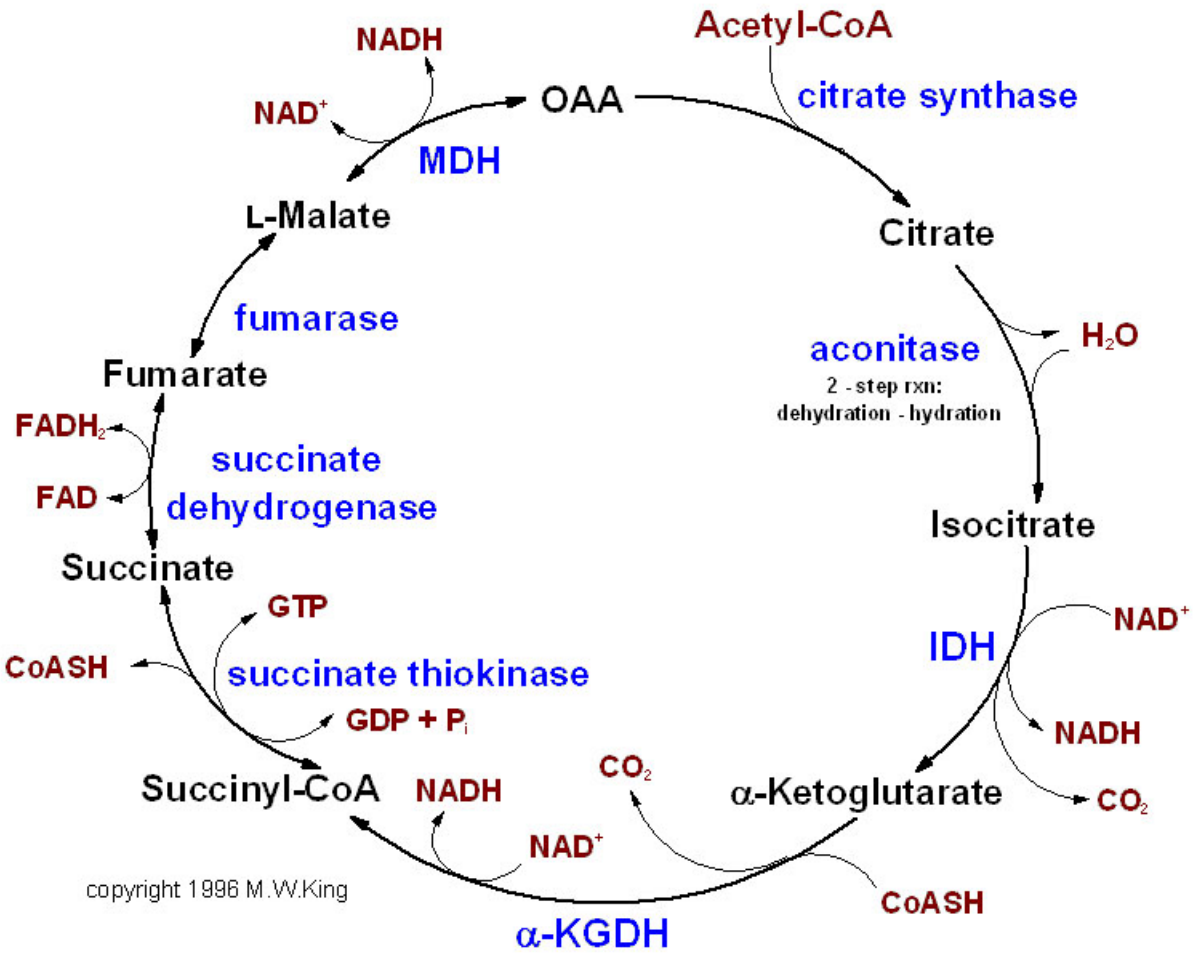
Enzyme

- a protein or a protein complex coded by one or several genes
- a single enzyme may accept distinct substrates and may catalyze several reactions
- a single reaction may be catalyzed by several enzymes
- the links between genes, proteins, and reactions (called GPR relationship) is nontrivial

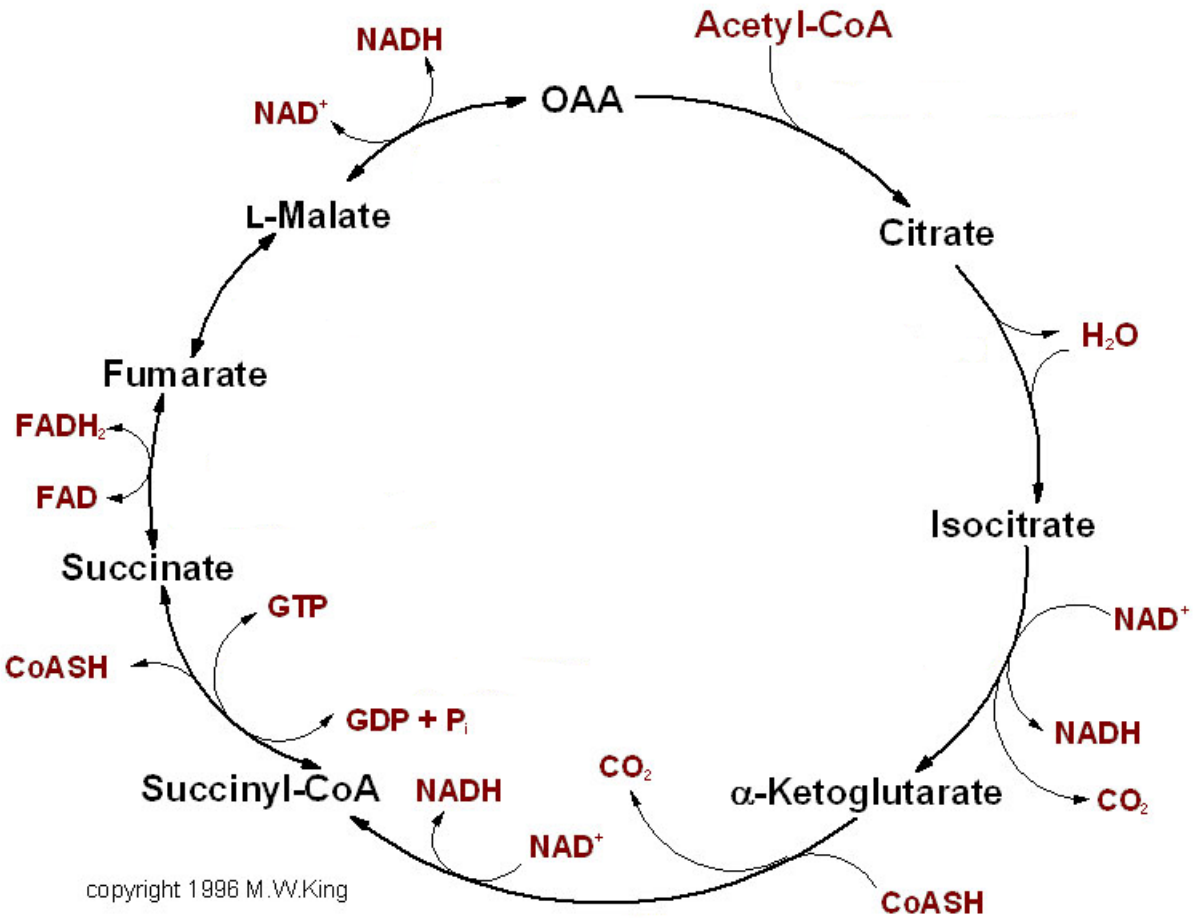
Cofactors

- small molecules
- essential to allow the catalysis of a reaction by an enzyme

Example - tricarboxylic acid cycle (TCA)



Components of a metabolic network

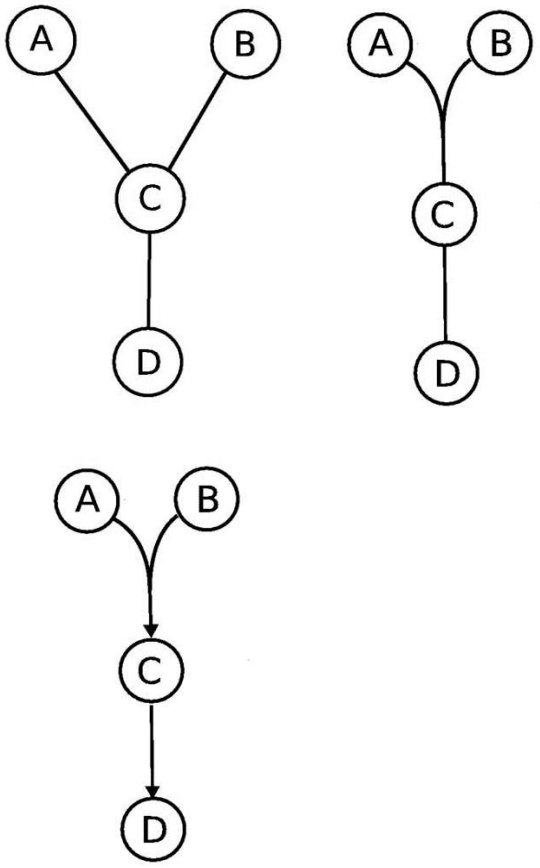


Metabolic networks

Metabolic networks can be represented as graphs

A + B → C

C → D



Possibility A

Definition of a graph: *G = (V,E)*

- *V is the set of nodes (metabolites)*
- *E is the set of edges (reactions)*

directed graph:

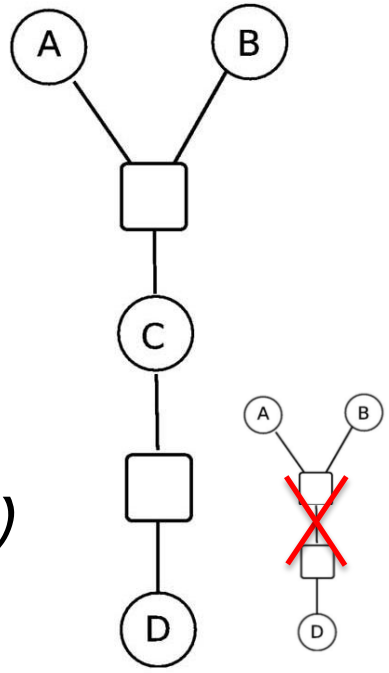
- Directed edge $e = (v_1, v_2)$ with v_1 (*start node)* und v_2 (*end node)* of e.

Metabolic networks

Metabolic networks can be represented as graphs

A + B → C

C → D



Possibility B

Bipartite Graph *G = (U,V,E)*

- *U is one type of nodes (metabolites)*
- *V is the other type of nodes (reactions)*
- *E is the set of edges with e = (u,v) or e=(v,u)*
- *directed* bipartite graphs are very common (not shown here)

Which graph to choose depends on the type of questions asked.

Graph analysis

- Topological network properties like clustering, centrality, degree, density, distance and neighborhood etc. can be calculated

- But as these measures have already been covered in the lecture about protein interaction networks they will be skipped in favor of flux balance analysis which yields more interesting results for metabolic networks

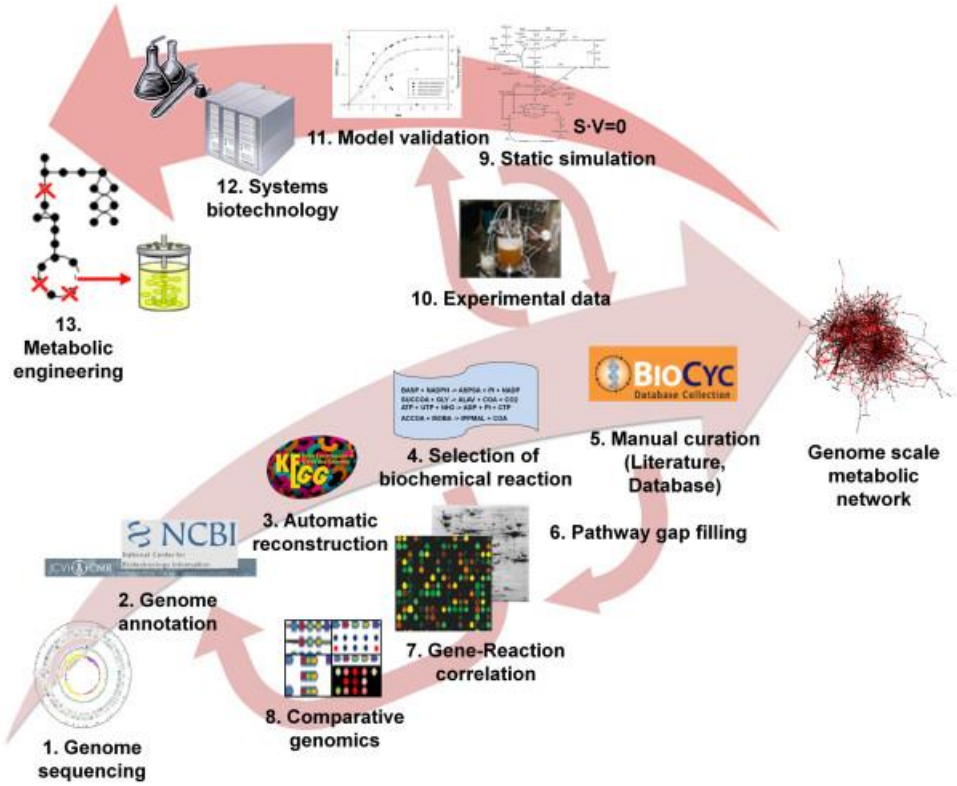
This Lecture

- Introduction
  - Systems biology & modelling
  - Metabolism & metabolic networks
- Network reconstruction
  - Strategy & workflow
- Mathematical representation
  - The stoichiometric matrix
  - Convex analysis & solution space
  - Elementary Modes & Extreme Pathways
- Constraint based network analysis
  - Flux Balance Analysis & Optimization
  - Tools

(Genome Scale) Network Reconstruction

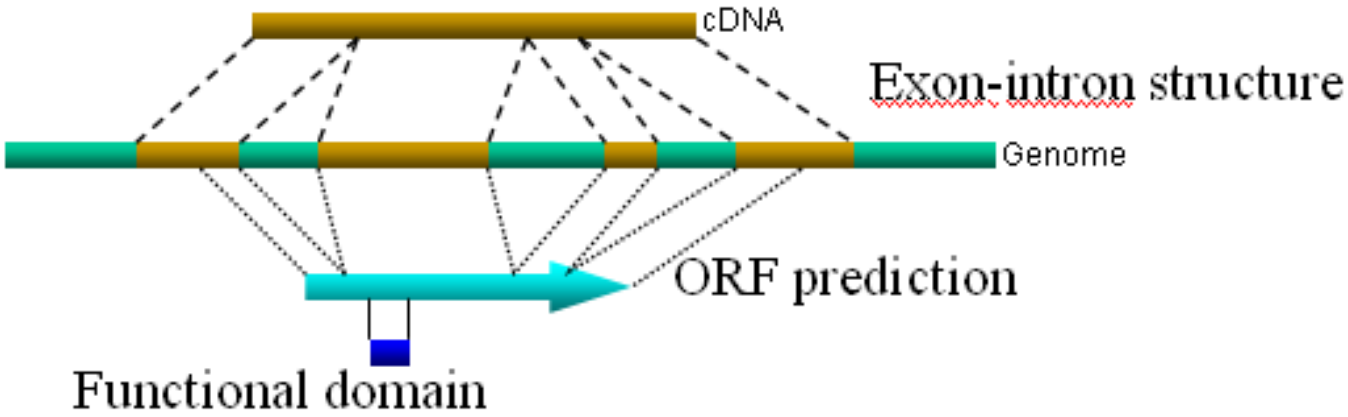
Three stages

- Use gene annotation for enzyme activity
- Link enzyme activity with reactions
- Compile the reactionlist



Genome Annotation

- In silico annotation methods typically lead to a functional assignment of 40-70% of identified ORFs on a freshly sequenced microbial genome

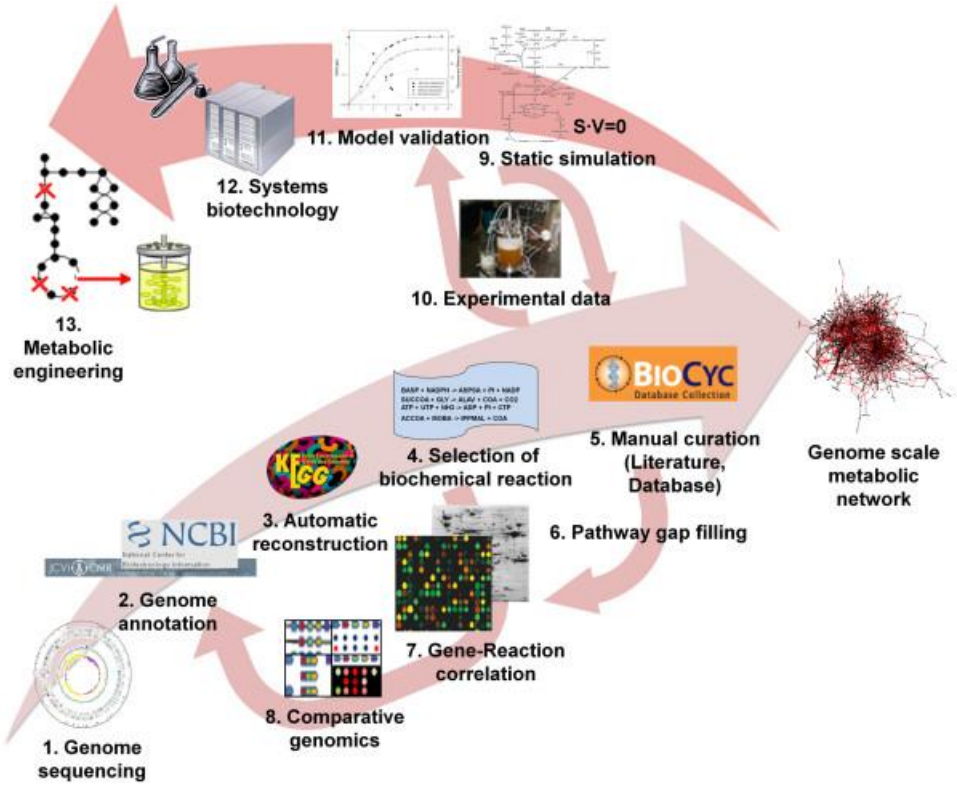


- Automatic network reconstruction from genome annotation has been attempted but needs intensive manual verification of the components and links of a network to produce high quality reconstructions.

(Genome Scale) Network Reconstruction

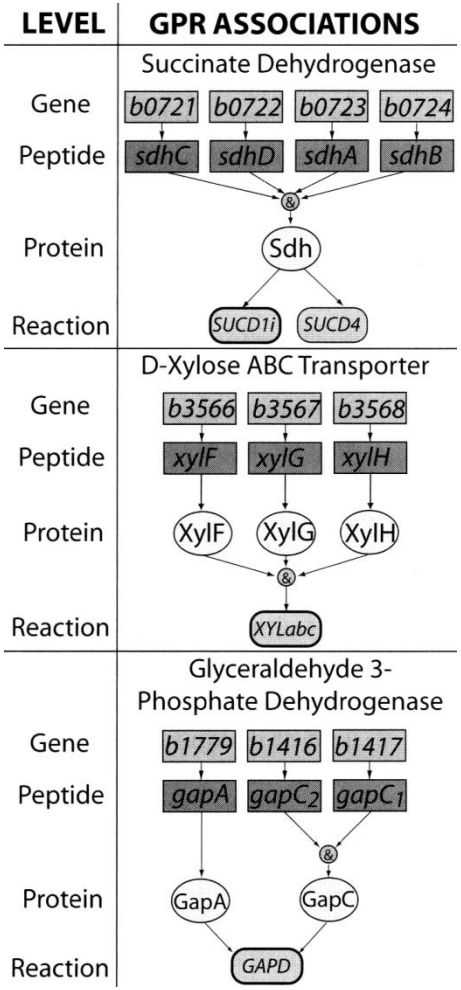
Three stages

- Use gene annotation for enzyme activity
- Link enzyme activity with reactions
- Compile the reactionlist



GPR association

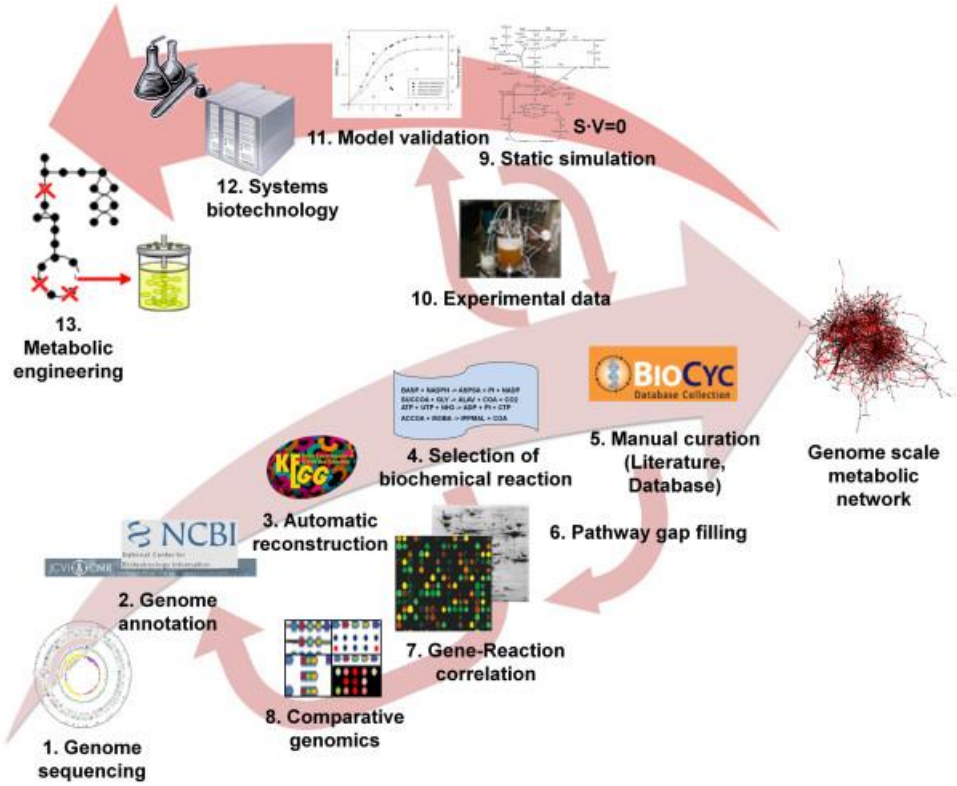
- Not all genes have a one-to-one relationship with their corresponding enzymes or metabolic reactions
  - Promiscous enzymes catalyzing more than one reaction
  - enzyme complexes consisting of different protein subunits which all need to be present for the reaction to occur
  - Iso enzymes /Isozymes catalyzing the same reaction



(Genome Scale) Network Reconstruction

Three stages

- Use gene annotation for enzyme activity
- Link enzyme activity with reactions
- Compile the reactionlist



Compiling the reaction list

Based on Evidence

- Biochemistry : direct enzyme assay
- Genomics: functional assignments to ORFs
- Physiology: indirect information
- In silico modelling: inferred reactions

Better evidence

- Make extensive use of existing knowledge (databases)

All in all lots of manual refinement & curation

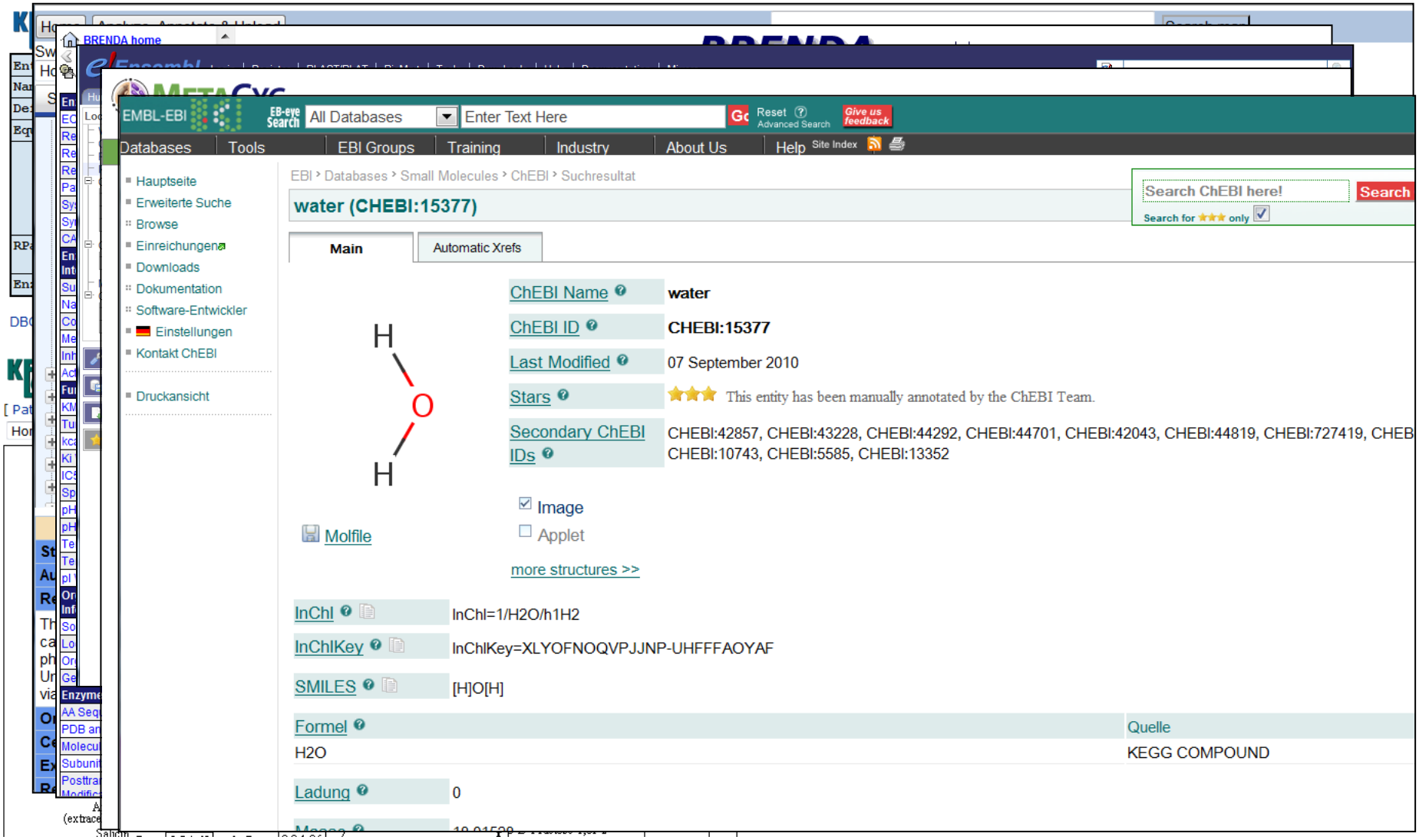
Databases

- Genomes
  ENSEMBL
- Pathway
  KEGG, REACTOME, METACYC
- Enzyme
  BRENDA
- Compounds
  KEGG, CHEBI, PUBCHEM
- Reconstruction Tools
  PATIKA, PathwayTools

Table 3.3: Publicly available genome databases. Prepared by Ines Thiele.

Microbial genomes and annotation	
DDBJ	http://www.ddbj.nig.ac.jp/
EBI	http://www.ebi.ac.uk/
EMBL	http://www.ebi.ac.uk/embl/
GenBank (NCBI)	http://www.ncbi.nlm.nih.gov/Genbank/
TIGR annotation software	http://www.tigr.org/software/
Comparative genomics	
ERGO	http://ergo.integratedgenomics.com/ERGO/
The SEED	http://theseed.uchicago.edu/FIG/index.cgi
GenDB	http://www.cebitec.uni-bielefeld.de/groups/brf/software/gendb_info/index.html
GeneQuiz	http://jura.ebi.ac.uk:8765/ext-genequiz/
MBGD	http://mbgd.genome.ad.jp/
Pedant	http://pedant.gsf.de/
Prolinks	http://128.97.39.94/cgi-bin/functionator/pronav
String	http://string.embl.de/
PUMA2	http://compbio.mcs.anl.gov/puma2/cgi-bin/index.cgi
Pathway/ Reconstruction tools	
INSILICO discovery	http://www.insilico-biotechnology.com/f_products.html
MetaFluxNet	http://mbel.kaist.ac.kr/mfn
MFAML (Metabolic Flux Analysis Markup Language)	http://mbel.kaist.ac.kr/mfaml
SimPheny	http://www.genomatica.com/solutions_simpheny.shtml
Pathfinder	http://bibiserv.techfak.uni-bielefeld.de/pathfinder/
PATIKA	http://www.patika.org/
Pathway databases	
BioSilico	http://biosilico.kaist.ac.kr or http://biosilico.org
KEGG	http://kegg.com/
MetaCyc	http://metacyc.org/
MRAD	http://capb.dbi.udel.edu/whisler/
Phylosopher	http://www.genedata.com/phylosopher.php
PUMA2	http://compbio.mcs.anl.gov/puma2/cgi-bin/index.cgi
EMP	http://www.empproject.com/
Enzymes	
Brenda	http://www.brenda.uni-koeln.de/
KEGG	http://www.kegg.com/
IntEnz	http://www.ebi.ac.uk/intenz/
Proteins	
HAMAP project	http://www.expasy.org/sprot/hamap/
InterPro	http://www.ebi.ac.uk/interpro/

Database Screenshots



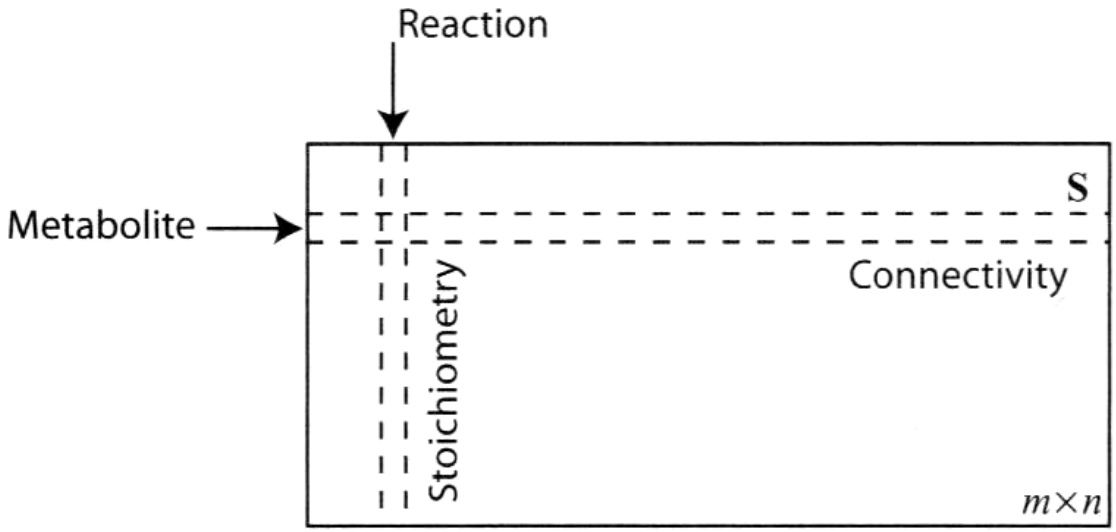
This Lecture

- Introduction
  - Systems biology & modelling
  - Metabolism & metabolic networks
- Network reconstruction
  - Strategy & workflow
- Mathematical representation
  - The stoichiometric matrix
  - Convex analysis & solution space
  - Elementary Modes & Extreme Pathways
- Constraint based network analysis
  - Flux Balance Analysis & Optimization
  - Tools

Stoichiometry

- The set of chemical reactions that comprise a network can be represented as a set of chemical equations
- Embedded in these chemical equations is information about reaction stoichiometry (the quantitative relationships of the reaction's reactants and products)
- Stoichiometry is invariant between organisms for the same reactions and does not change with pressure, temperature, or other conditions. $2\ H_2 + O_2 \longrightarrow 2\ H_2O$
- All this stoichiometric information can be represented in a matrix form; the stoichiometric matrix, denoted by **S**.

Stoichiometric matrix



m rows	m metabolites
n columns	n reactions
S_{ij}	number of molecules of the i^{th} metabolite produced in j^{th} reaction.

- Consumption is understood as negative production.

Example 1 - Reversible Conversion

$$\mathrm{CP} \rightleftharpoons \mathrm{PC}$$

Simple chemical rearrangement of the molecule without any change in its elemental composition. Isomerase catalyze such reaction.

$$\mathbf{S} = \begin{array}{c} \\ \mathrm{CP} \\ \mathrm{PC} \end{array} \begin{array}{c} \begin{array}{cc} \text{Forward} & \text{Backward} \end{array} \\ \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}$$

Example 2 - Bimolecular association

$$\mathrm{C} + \mathrm{P} \rightleftharpoons \mathrm{CP}$$

Combination of two moieties, C and P, to form a new compound, e.g., the dimerization of two protein molecules, or the initial binding of a substrate to an active site on an enzyme molecules.

$$\mathbf{S} = \begin{array}{c} \\ \mathrm{C} \\ \mathrm{P} \\ \mathrm{CP} \end{array} \begin{array}{c} \begin{array}{cc} \text{Forward} & \text{Backward} \end{array} \\ \begin{pmatrix} -1 & 1 \\ -1 & 1 \\ 1 & -1 \end{pmatrix} \end{array}$$

Example 3 - Cofactor-Coupled Reaction

$$\mathrm{C} + \mathrm{AP} \rightleftharpoons \mathrm{CP} + \mathrm{A}$$

One compound AP donates a moiety P to another compound C. Such reaction can be decomposed into two bimolecular assoication reactions.

$$\mathbf{S} = \begin{array}{c} \\ \mathrm{C} \\ \mathrm{AP} \\ \mathrm{CP} \\ \mathrm{A} \end{array} \begin{array}{c} \begin{array}{cc} \text{Forward} & \text{Backward} \end{array} \\ \begin{pmatrix} -1 & 1 \\ -1 & 1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix} \end{array}$$

Mathematical Interpretation

Mathematically, the stoichiometric matrix S is a linear transformation of the flux vector

$$\mathbf{v} = (v_1, v_2, \ldots, v_n)$$

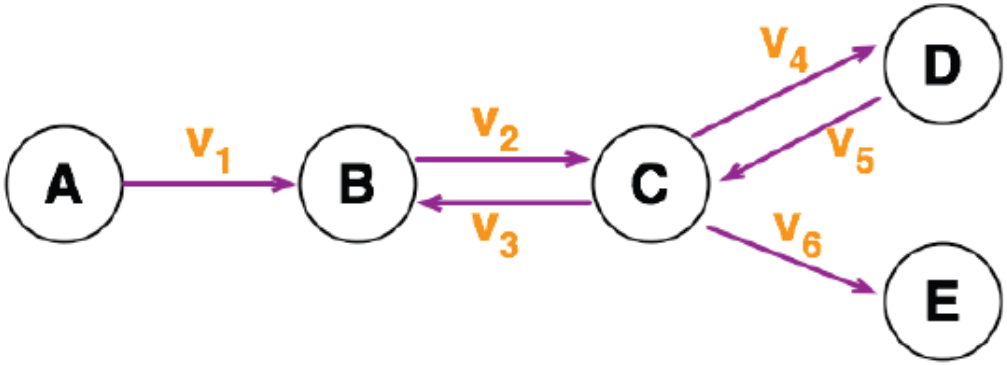
to a vector of time derivatives of the concentration

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = (\frac{\mathrm{d}x_1}{\mathrm{d}t}, \frac{\mathrm{d}x_2}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}x_m}{\mathrm{d}t})$$

or

$$\mathbf{S} \cdot \mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$$

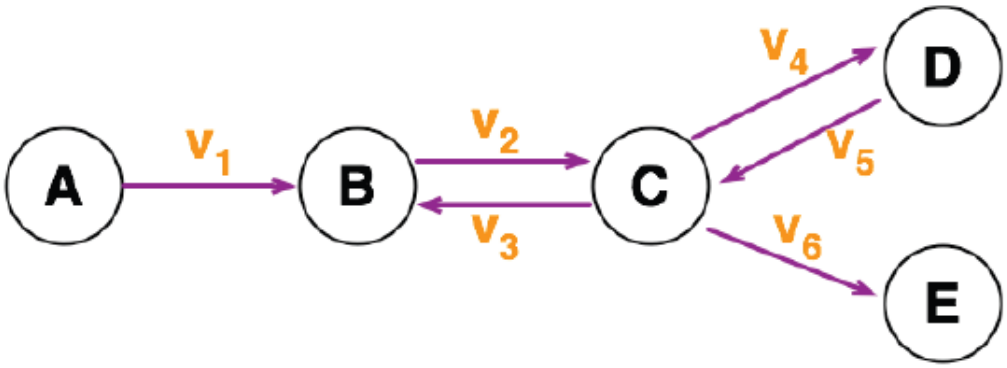
Example 4 – system of 5 metabolites



$$\mathbf{S} \cdot \mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$$

$$\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = \begin{pmatrix} \dot{\mathrm{A}} \\ \dot{\mathrm{B}} \\ \dot{\mathrm{C}} \\ \dot{\mathrm{D}} \\ \dot{\mathrm{E}} \end{pmatrix}$$

Example 4 – viewed from a row



$$\frac{dC}{dt} = 0v_1 + 1v_2 - 1v_3 - 1v_4 + 1v_5 - 1v_6$$

Fluxes that form C

Fluxes that degrade C

First constraint

- If we think about metabolism, we remember that it tends to maintain a balanced state.
- So a good assumption might be that under normal environmental conditions a constant flow through the system would be optimal.

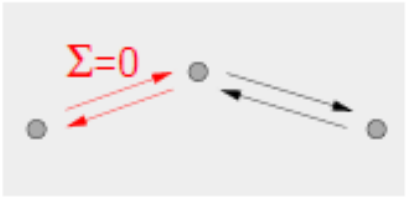
- This implies that the relative concentration of metabolites would remain constant over time, which in fact is the central constraint for flux balance analysis

$$Sv = 0$$

Steady state

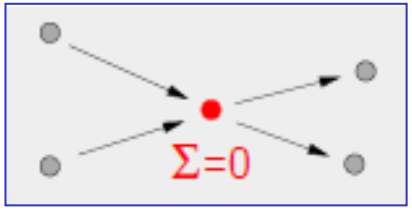
Balance

- A state in which no macroscopic events take place
- For every reaction the forward and reverse direction cancel each other out
- Analogy: „calm lake"



Steady state

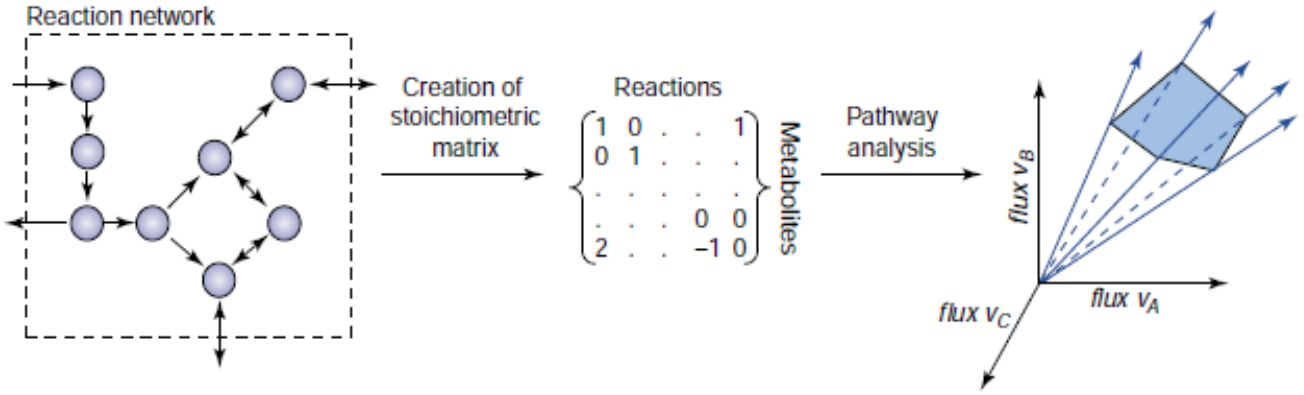
- A state in which every variable is temporally constant
- For every substance generating and consuming reactions cancel each other out → such a system can „live"
- Analogy: „calm river"



$$Sv = 0$$

Convex analysis

- Equalities (in this case Sv = 0) and inequalities (in this case, $0 \leq v_i \leq v_{i,max}$) lead to convex analysis.
- It leads to the definition of a set of nonnegative generating vectors that span the solution space, the so called Extreme Pathways.
- All possible flux distributions of a metabolic network lie within the cone circumscribed by these pathways.



Elementary Flux Modes

Three Conditions for S_{EFM}

1. Steady state $\mathbf{S} \cdot \mathbf{e} = 0$
2. Thermodynamic feasibility $e_i \geq 0$ if reaction i irreversible
3. Non-decomposability no nontrivial vector v fulfilling the two conditions above and such that $P(\mathbf{v}) \subset P(\mathbf{e})$

Where P(v) denotes the set of reaction with non-zero flux in the mode v.

- A EFM is a set of enzymes that operate together at steady state and a mode is elementary when the removal of one enzyme causes to fail.

Extreme Pathways

Two additional Conditions for S_{EP}

4. Network reconfiguration: reactions must be classified either as exchange or internal. All reversible internal reactions are splitted into two separate, irreversible reactions with opposite direction. Internal flux can only admit non-negative value. Exchange reactions can be reversible.
5. Systemic independance: the set of EPs in a network is the minimal set of EFMs such that all feasible steady-state fluxes are non-negative linear combinations of these extremal pathways.

Extreme Pathways (Definition)

Comparison of network-based pathway analysis methods

Jason A. Papin[1], Joerg Stelling[2], Nathan D. Price[1], Steffen Klamt[2], Stefan Schuster[3] and Bernhard O. Palsson[1]

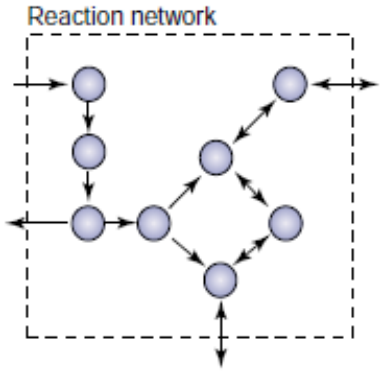
Extreme pathways

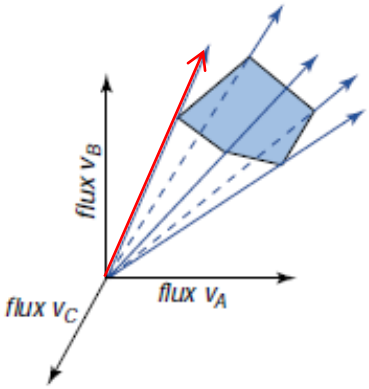
Extreme pathways ($\mathbf{p}_i$) are a set of convex basis vectors derived from the stoichiometric matrix [5]. They have the following properties:

(I) There is a unique set of extreme pathways for a given network.

(II) Each extreme pathway consists of the minimum number of reactions that it needs to exist as a functional unit.

(III) The extreme pathways are the systemically independent subset of elementary modes; that is, no extreme pathway can be represented as a nonnegative linear combination of any other extreme pathways.





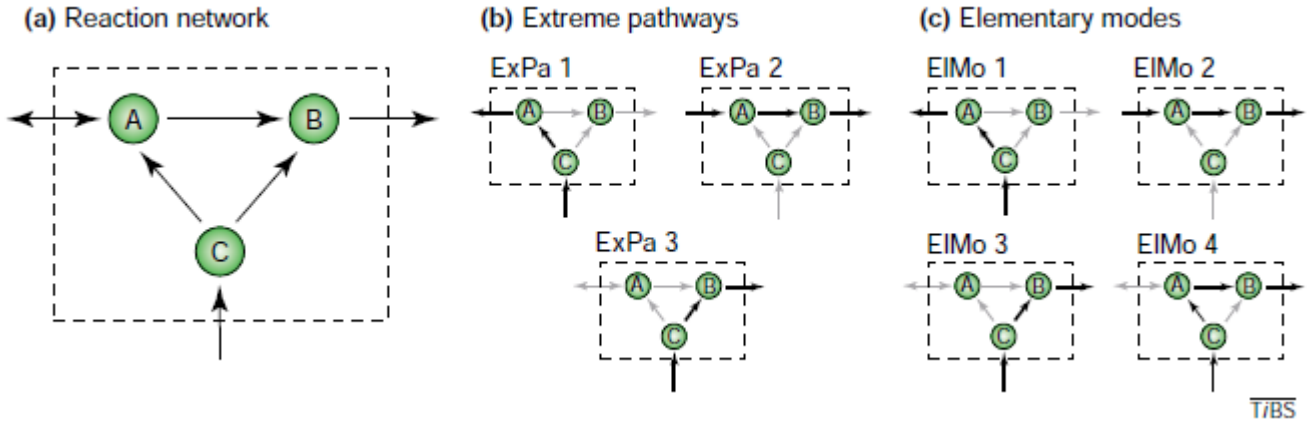
Papin 2004 Comparison of network-based pathway analysis methods - PMID: 15283984

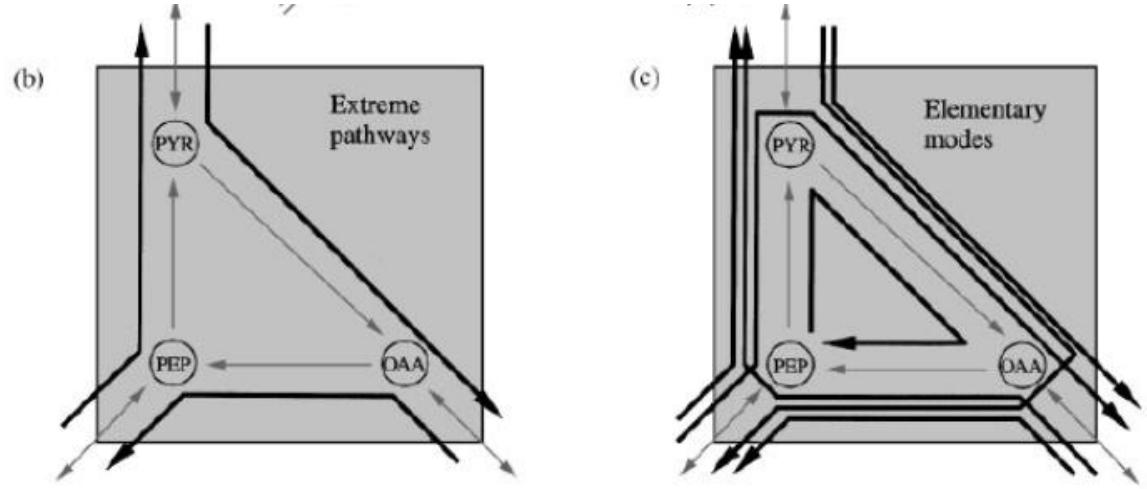
EP vs EFM

- EP is the minimal set of EFM in the sense that no element in EP is the convex combination of other elements.

- EP = EFM in the case where all exchange reactions (reactions connecting a metabolite with the outside of a metabolic system) are irreversible (e.g. glucose can only go into and ammonia can only go out of the system).

Examples (EP & EFM)

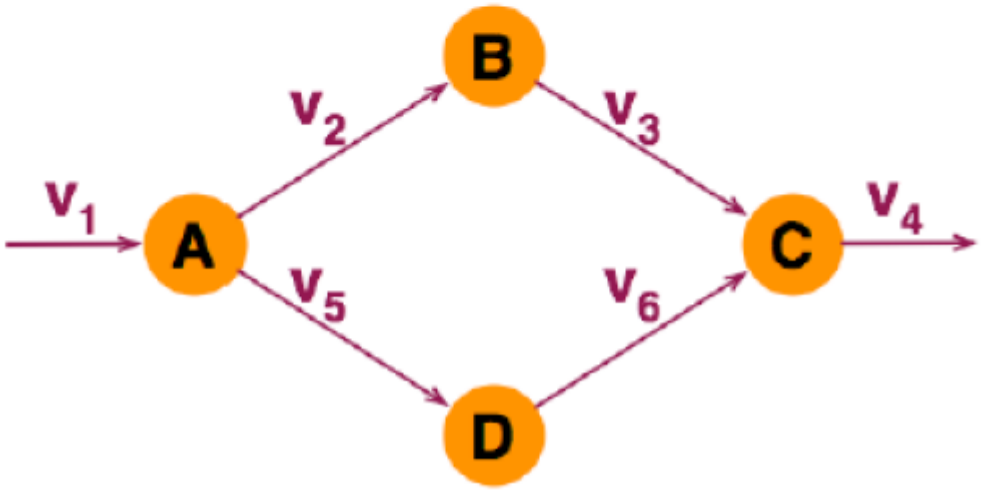




How to compute Extreme Pathways (1)

Recall the stoichiometric matrix

$$N = \begin{bmatrix} 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$



How to compute Extreme Pathways (2)

Rewrite N^T by grouping internal reactions and external reactions.

$$N^T = \begin{array}{c} \\ v_2 \\ v_3 \\ v_5 \\ v_6 \\ v_1 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \end{array}$$

How to compute Extreme Pathways (3)

Use extended indentity matrix, left panel record the history of reaction usage, right panel record the current net change of metabolic species.

$$
\begin{array}{c} \\ v_2 \\ v_3 \\ v_5 \\ v_6 \\ v_1 \\ v_4 \end{array}
\begin{array}{c}
\begin{array}{cccccc|cccc} v_2 & v_3 & v_5 & v_6 & v_1 & v_4 & A & B & C & D \end{array} \\
\left(\begin{array}{cccccc|cccc}
1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0
\end{array}\right)
\end{array}
$$

How to compute Extreme Pathways (4)

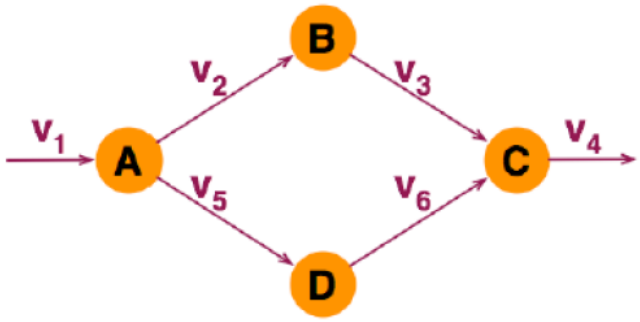
Identify metabolites that are not involved in external reactions. In this example, B and D. Consider B first, copy rows that correspond to internal reactions which do not involve B (in this case v_5 and v_6). Combine reaction pairs which run opposite on B (in this case v_2 and v_3). External reactions remain the same.

$$
\begin{array}{c}
 \\
v_2 + v_3 \\
v_5 \\
v_6 \\
v_1 \\
v_4
\end{array}
\begin{array}{c}
\begin{array}{cccccccccc} v_2 & v_3 & v_5 & v_6 & v_1 & v_4 & A & B & C & D \end{array} \\
\left(\begin{array}{cccccccccc}
1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0
\end{array} \right)
\end{array}
$$

How to compute Extreme Pathways (5)

Repeat the same for D

$$
\begin{array}{c}
 & \begin{array}{cccccccccc} v_2 & v_3 & v_5 & v_6 & v_1 & v_4 & A & B & C & D \end{array} \\
\begin{array}{l} v_2 + v_3 \\ v_5 + v_6 \\ v_1 \\ v_4 \end{array} & \left(\begin{array}{cccccccccc}
1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0
\end{array} \right)
\end{array}
$$



How to compute Extreme Pathways (6)

Use external reactions to balance internal net changes of all species remained. add v_1 onto $v_2 + v_3$ and $v_5 + v_6$ and remove v_1 (after balance A)

$$\begin{array}{c} \\ v_1 + v_2 + v_3 \\ v_1 + v_5 + v_6 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccccccccc} v_2 & v_3 & v_5 & v_6 & v_1 & v_4 & A & B & C & D \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \end{pmatrix} \end{array}$$

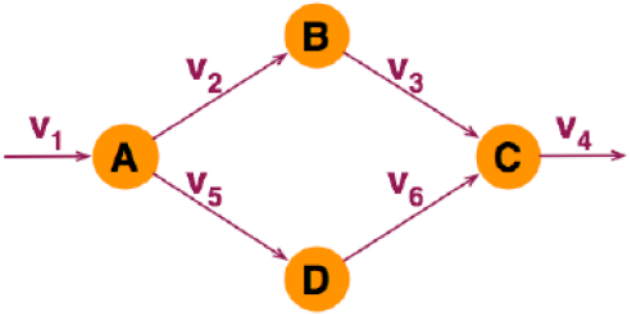
add v_4 onto $v_1 + v_2 + v_3$ and $v_1 + v_5 + v_6$ and remove v_4 (after balance C

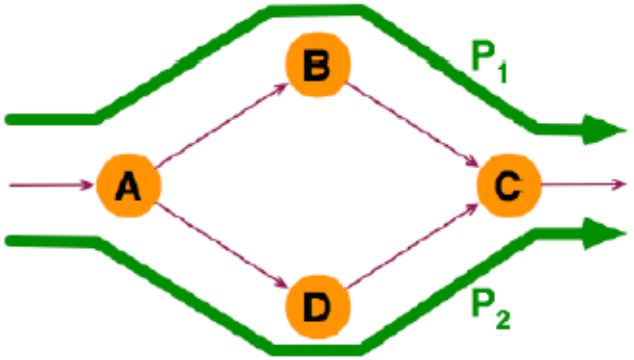
$$\begin{array}{c} \\ v_1 + v_2 + v_3 + v_4 \\ v_1 + v_4 + v_5 + v_6 \end{array} \begin{array}{c} \begin{array}{cccccccccc} v_2 & v_3 & v_5 & v_6 & v_1 & v_4 & A & B & C & D \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

How to compute Extreme Pathways (7)

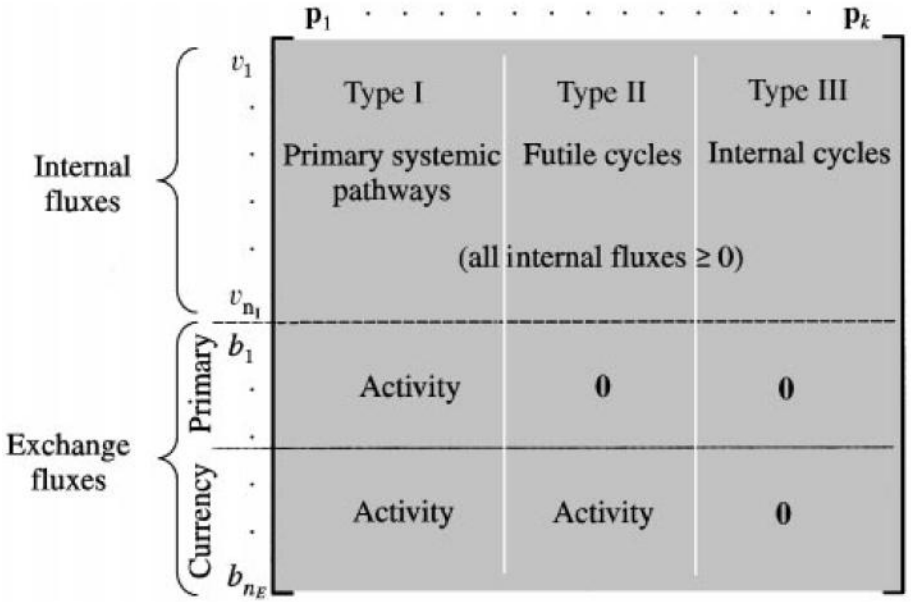
- Now we balanced all the species (net change is zero) and the flux combination recorded on the left panel corresponds to the extremal pathways

$$\begin{matrix} & v_2 & v_3 & v_5 & v_6 & v_1 & v_4 \end{matrix}$$
$$EP_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$
$$EP_2 = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$



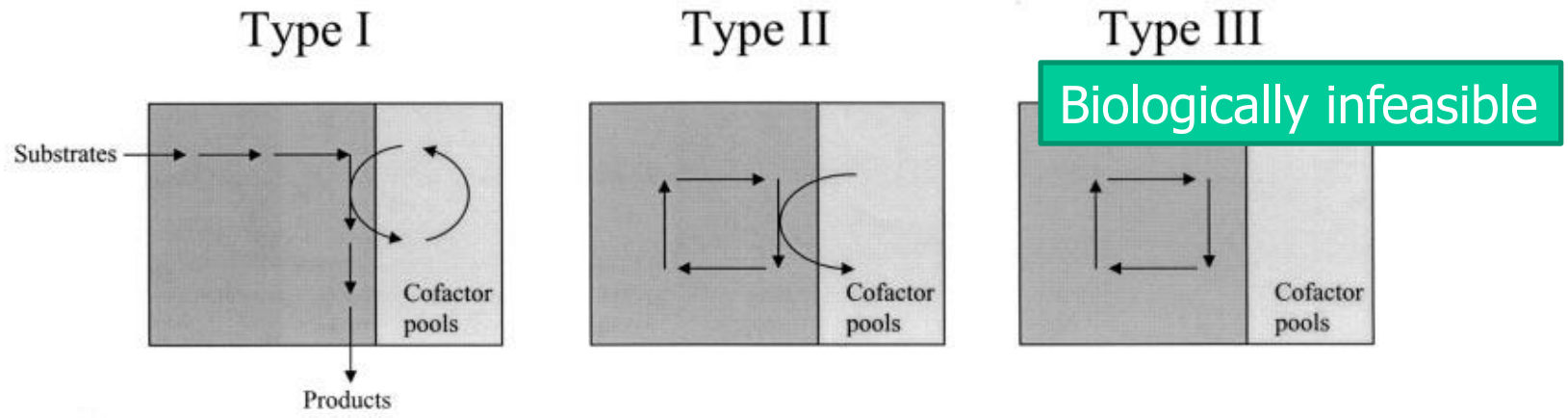


Extreme Pathway Classification



$$EP_1 = \begin{matrix} v_2 & v_3 & v_5 & v_6 & v_1 & v_4 \end{matrix} \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$EP_2 = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

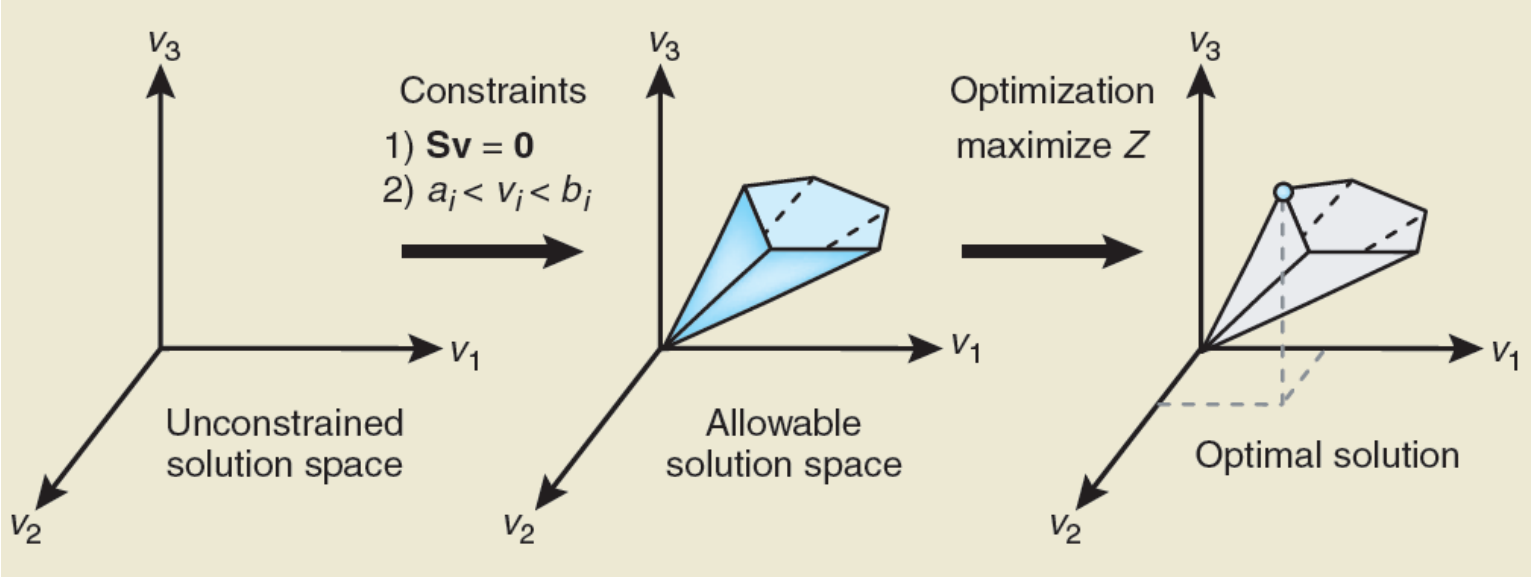


This Lecture

- Introduction
  - Systems biology & modelling
  - Metabolism & metabolic networks
- Network reconstruction
  - Strategy & workflow
- Mathematical representation
  - The stoichiometric matrix
  - Convex analysis & solution space
  - Elementary Modes & Extreme Pathways
- Constraint based network analysis
  - Flux Balance Analysis & Optimization
  - Tools

Optimization based methods (FBA)

- Identify an objective function Z to be minimized or maximized depending on the task at hand
- Set upper and lower bounds on the flux variables $v_{i,min} \leq v_i \leq v_{i,max}$ for i={1,…,n}
- Solve Linear Programming Problem (i.e. Simplex Algorithm)



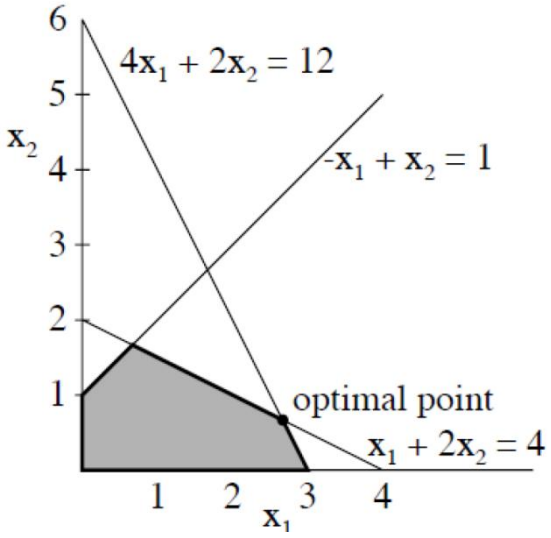
Linear Programming

Linear Programming (LP)

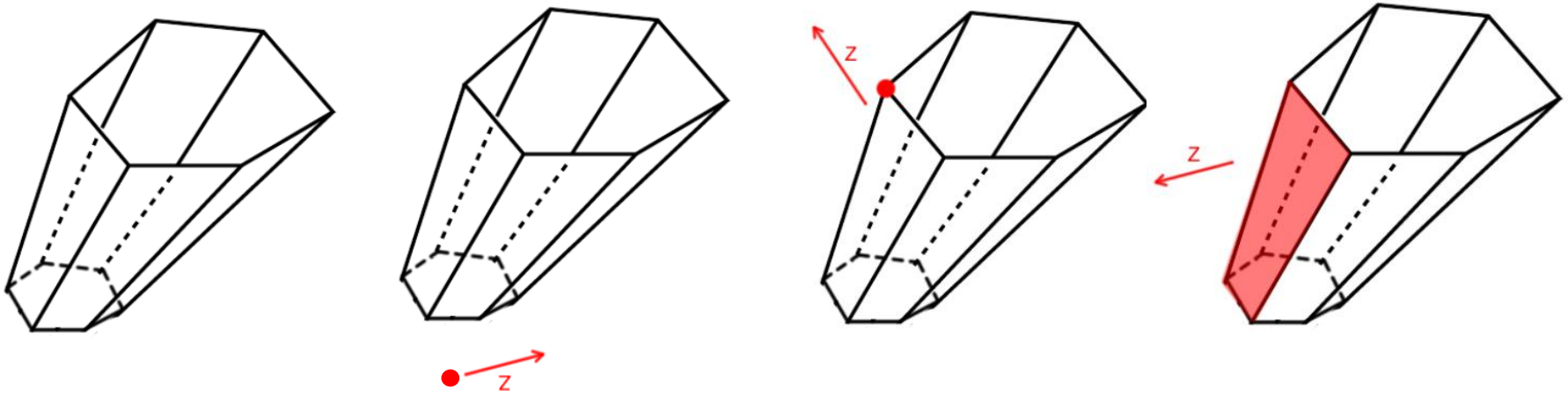
A linear programming problem is the problem of *maximizing* or *minimizing* a linear function subject to *linear constraints*. The constraints may be equalities or inequalities.

For example, we want to find numbers x_1 and x_2 that maximize the sum $x_1 + x_2$ subject to the constraints $x_1 \geq 0$, $x_2 \geq 0$ and ,

$$x_1 + 2x_2 \leq 4$$
$$4x_1 + 2x_2 \leq 12$$
$$-x_1 + x_2 \leq 1$$



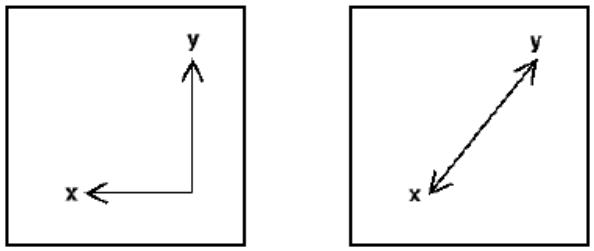
Possible Solutions for LP (visual)



$v_{i,min} \leq v_i \leq v_{i,max}$ empty unique degenerated

Minimization - Objective Functions

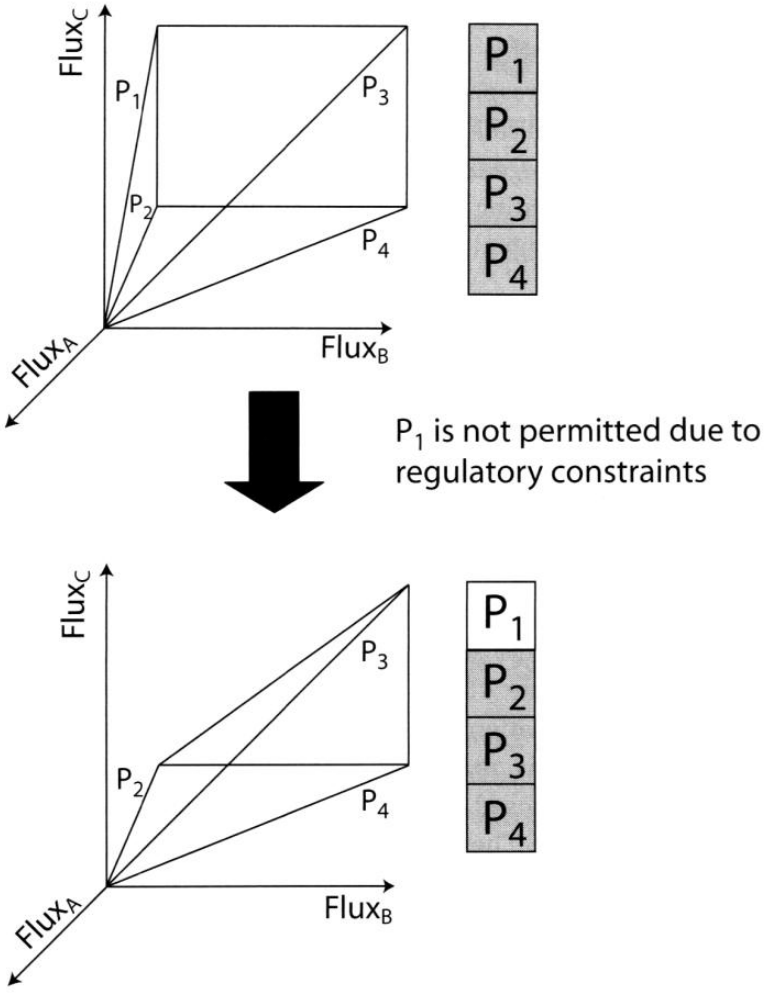
- Minimize ATP production : most energy-efficient state
- Minimize nutrient uptake : the fittest state under nutrient-shortage
- Minimize the Manhattan distance or Euclidian distance of the flux vector : minimize the overall flux



Maximization - Objective Functions

- Maximize metabolite production : the biochemical production capabilities of certain desirable metabolite such as lysine, phenylalanine etc.
- Maximize biomass formation : maximal growth rate
- Maximize biomass as well es metabolite production : the trade-off between cell-growth and forced metabolite production

Additional constraints restraining the solution space



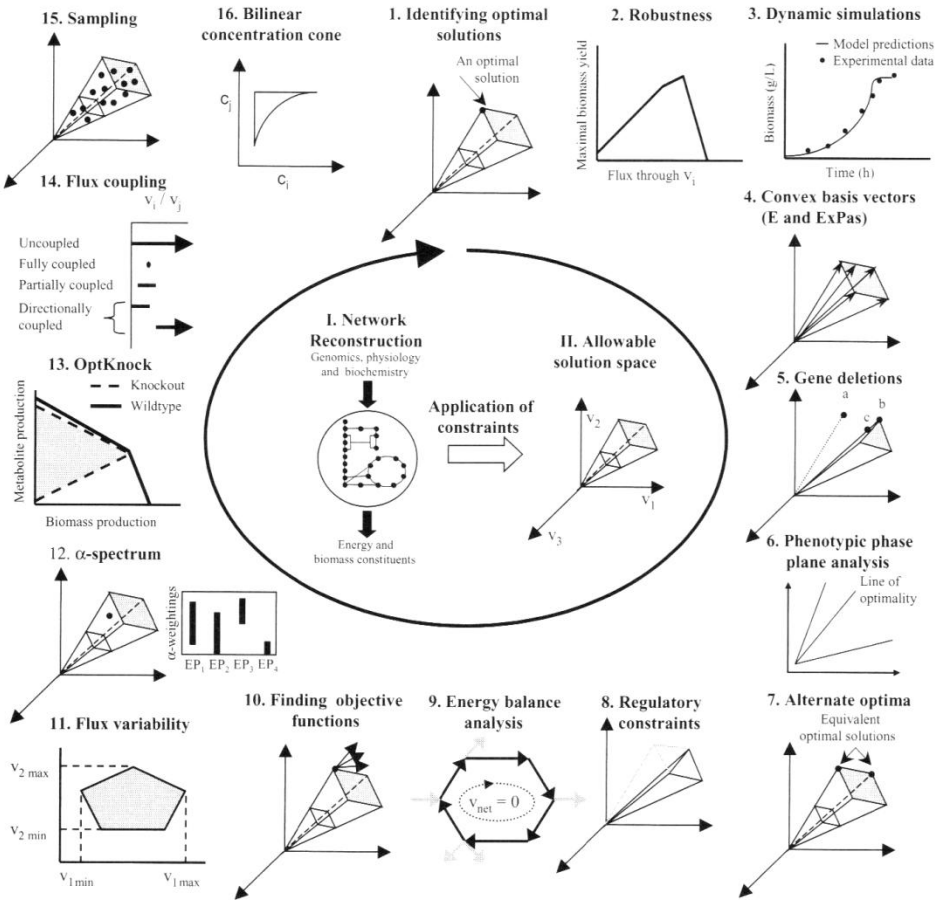


Figure 16.1: An overview of constraint-based methods used to analyze properties of reconstructed networks. Modified from [176].

Summary Flux Balance Analysis

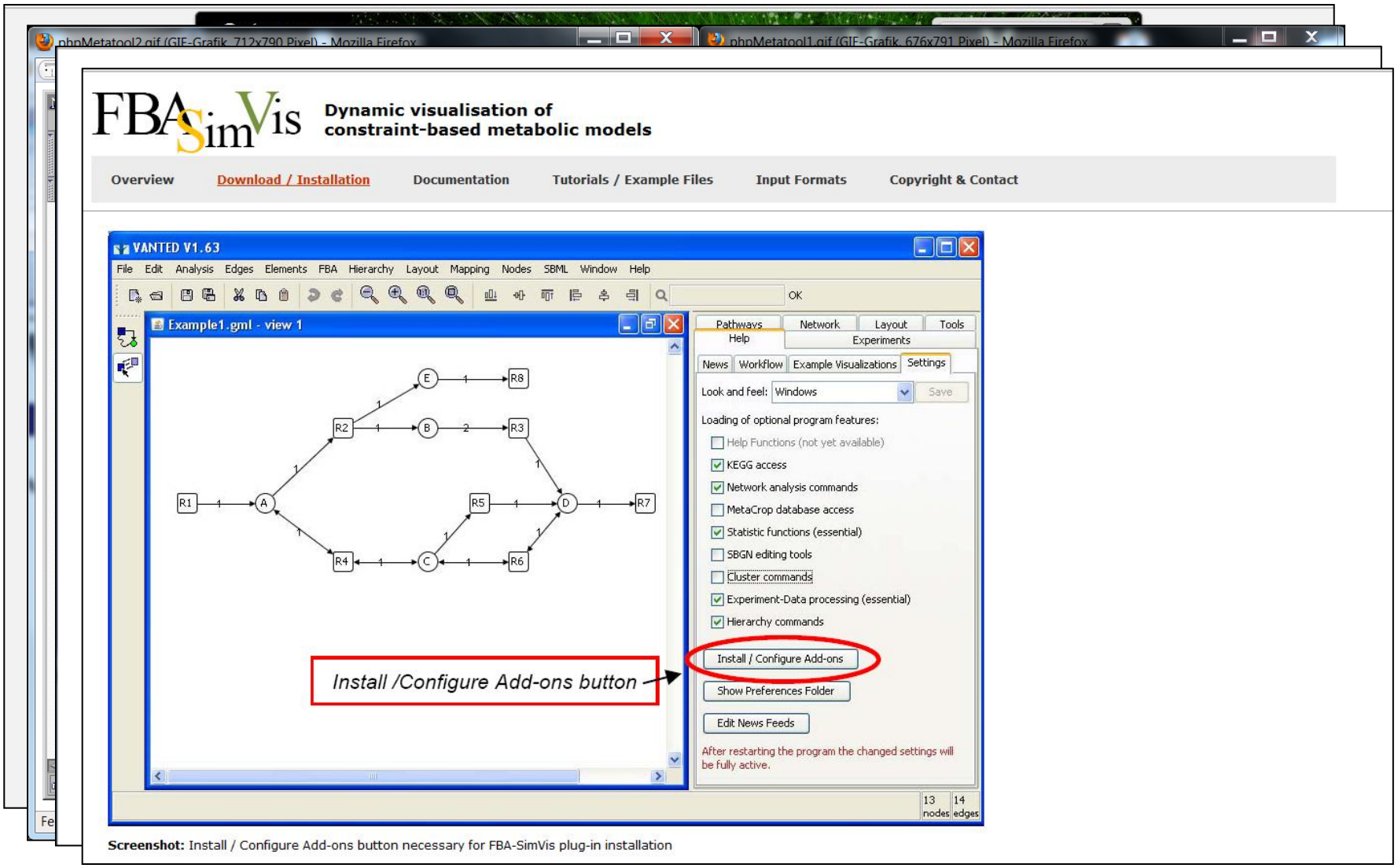
Advantages

- For an objective function we get a 'fast' answer what the optimal flux distribution is.
- Successfully employed for several microorganisms.

Drawbacks

- Sensitivity to the definition of objective function.
- Optimal flux distribution might not be unique.
- Microorganisms could use different optimization criteria depending on their environment.
- The exploration of all suitable objective function is a difficult task.

Tools for FBA



Screenshot: Install / Configure Add-ons button necessary for FBA-SimVis plug-in installation

Further Reading

- Llaneras 2008 Stoichiometric Modelling of Cell Metabolism PMID: 18295713
- Papin 2004 Comparison of network-based pathway analysis methods PMID: 15283984
- Harald Marx 2007 „Ausarbeitung Extreme Pathways“
- More kinetic analysis in cellular systems visit the lectures (Systemsbiology, Modelle zellulärer Prozesse) of Prof. Dr. Edda Klipp (Head of Theoretical Biophysics)