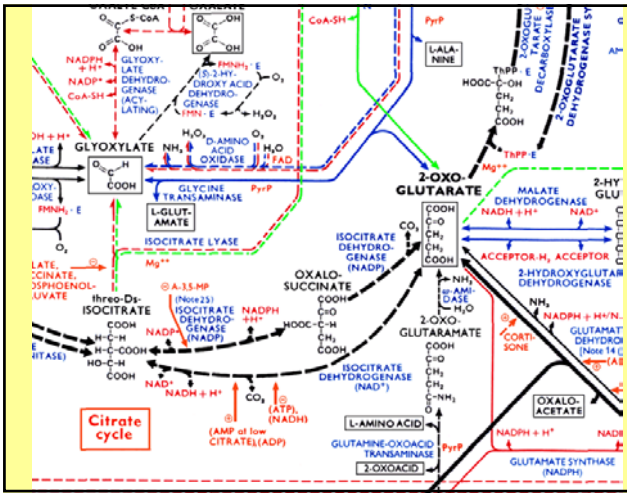
Metabolic Pathway Analysis. Fundamentals and Applications

Stefan Schuster
Friedrich Schiller University Jena
Dept. of Bioinformatics



Introduction

- Analysis of metabolic systems requires theoretical methods due to high complexity
- Major challenge: clarify relationship between structure and function in complex intracellular networks
- Study of robustness to enzyme deficiencies and knock-out mutations is of high medical and biotechnological relevance

Theoretical Methods

- Dynamic Simulation
- Stability and bifurcation analyses
- Metabolic Control Analysis (MCA)
- Metabolic Pathway Analysis
- Metabolic Flux Analysis (MFA)
- Optimization
- and others

Theoretical Methods

- Dynamic Simulation
- Stability and bifurcation analyses
- Metabolic Control Analysis (MCA)
- **Metabolic Pathway Analysis**
- Metabolic Flux Analysis (MFA)
- Optimization
- and others

Metabolic Pathway Analysis (or Metabolic Network Analysis)

- Decomposition of the network into the smallest functional entities (metabolic pathways)
- Does not require knowledge of kinetic parameters!!
- Uses stoichiometric coefficients and reversibility/irreversibility of reactions

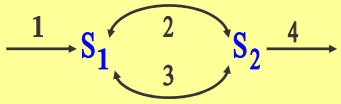
History of pathway analysis

- „Direct mechanisms“ in chemistry (Milner 1964, Happel & Sellers 1982)
- Clarke 1980 „extreme currents“
- Seressiotis & Bailey 1986 „biochemical pathways“
- Leiser & Blum 1987 „fundamental modes“
- Mavrovouniotis et al. 1990 „biochemical pathways“
- Fell (1990) „linearly independent basis vectors“
- Schuster & Hilgetag 1994 „elementary flux modes“
- Liao et al. 1996 „basic reaction modes“
- Schilling, Letscher and Palsson 2000 „extreme pathways“

Mathematical background

Stoichiometry matrix

- *Example*:



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

Steady-state condition

Balance equations for metabolites:

$$\frac{dS_i}{dt} = \sum_j n_{ij} v_j$$

$d\mathbf{S}/dt = \mathbf{NV(S)}$

At any stationary state, this simplifies to:

$$\mathbf{NV(S)} = \mathbf{0}$$

Kernel of N

Steady-state condition NV(S) = 0

If the kinetic parameters were known, this could be solved for **S**.

If not, one can try to solve it for **V**. The equation system is linear in **V**. However, usually there is a manifold of solutions.

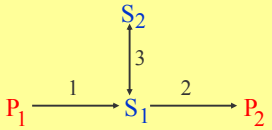
Mathematically: kernel (null-space) of **N**. Spanned by basis vectors. These are not unique.

Use of null-space

The basis vectors can be gathered in a matrix, **K**. They can be interpreted as biochemical routes across the system.

If some row in **K** is a null row, the corresponding reaction is at thermodynamic equilibrium in any steady state of the system.

Example:



$$\mathbf{K} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

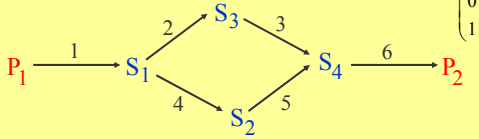
Use of null-space (2)

It allows one to determine „enzyme subsets“ = sets of enzymes that always operate together at steady, in fixed flux proportions.

The rows in **K** corresponding to the reactions of an enzyme subset are proportional to each other.

Example:
Enzyme subsets: {1,6}, {2,3}, {4,5}

$$\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$



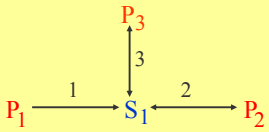
Pfeiffer et al., *Bioinformatics* 15 (1999) 251-257.

Drawbacks of null-space

The basis vectors are not necessarily the simplest possible.

They do not necessarily comply with the directionality of irreversible reactions.

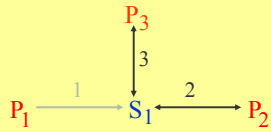
They do not always properly describe knock-outs.



$$\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

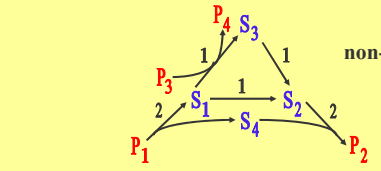
Drawbacks of null-space

They do not always properly describe knock-outs.

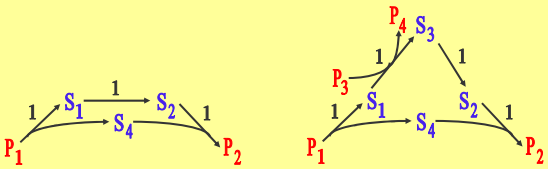


$$\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

After knock-out of enzyme 1, the route {-2, 3} remains!



non-elementary flux mode



elementary flux modes

S. Schuster und C. Hilgetag: *J. Biol. Syst.* 2 (1994) 165-182

An elementary mode is a minimal set of enzymes that can operate at steady state with all irreversible reactions used in the appropriate direction

The enzymes are weighted by the relative flux they carry.

The elementary modes are unique up to scaling.

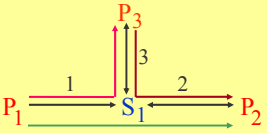
All flux distributions in the living cell are non-negative linear combinations of elementary modes

Non-Decomposability property:

For any elementary mode, there is no other flux vector that uses only a proper subset of the enzymes used by the elementary mode.

For example, {HK, PGI, PFK, FBPase} is not elementary if {HK, PGI, PFK} is an admissible flux distribution.

Simple example:



Elementary modes: $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$

They describe knock-outs properly.

Mathematical background (cont.)

Steady-state condition $\mathbf{NV = 0}$
Sign restriction for irreversible fluxes: $\mathbf{V^{irr} \geqq 0}$

This represents a linear equation/inequality system.

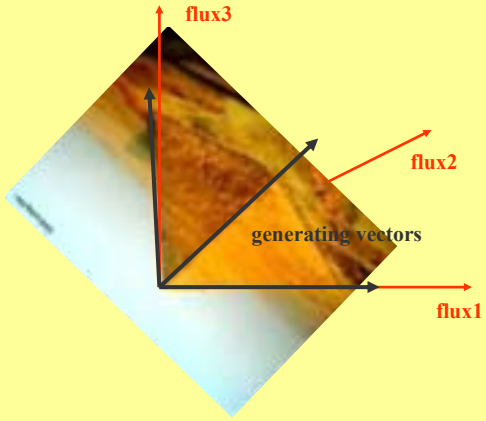
Solution is a convex region.

All edges correspond to elementary modes.

In addition, there may be elementary modes in the interior.

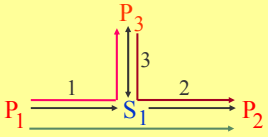
Geometrical interpretation

Elementary modes correspond to generating vectors (edges) of a convex polyhedral cone (= pyramid) in flux space (if all modes are irreversible)

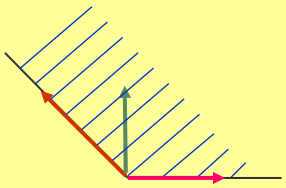


If the system involves reversible reactions, there may be elementary modes in the interior of the cone.

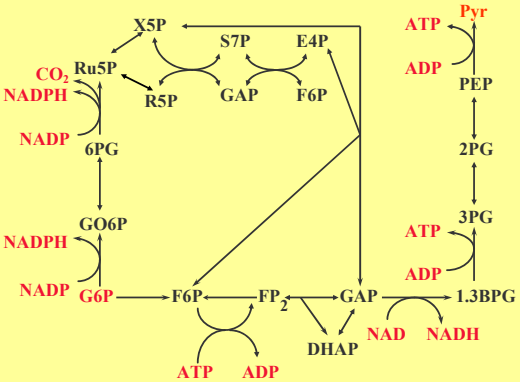
Example:



Flux cone:

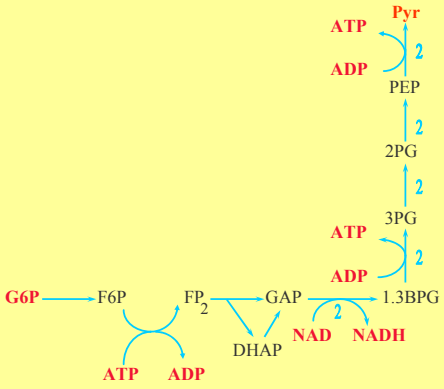


There are elementary modes in the interior of the cone.

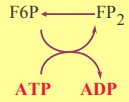


Part of monosaccharide metabolism

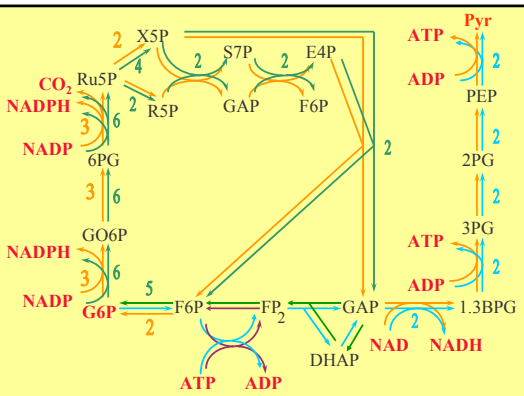
Red: external metabolites



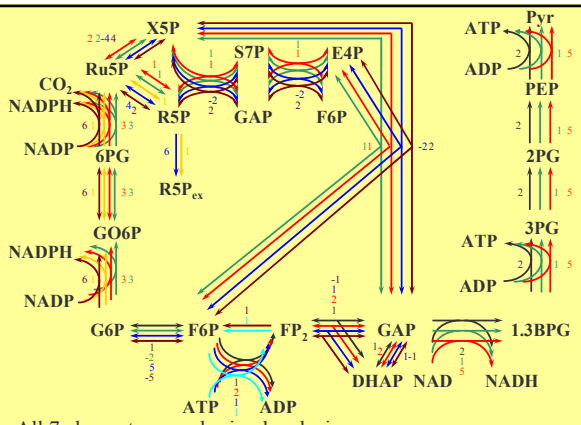
1st elementary mode: glycolysis



2nd elementary mode: fructose-bisphosphate cycle



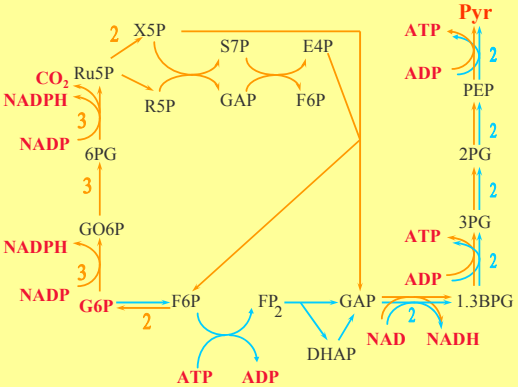
4 out of 7 elementary modes in glycolysis-pentose-phosphate system



All 7 elementary modes in glycolysis-pentose-phosphate system

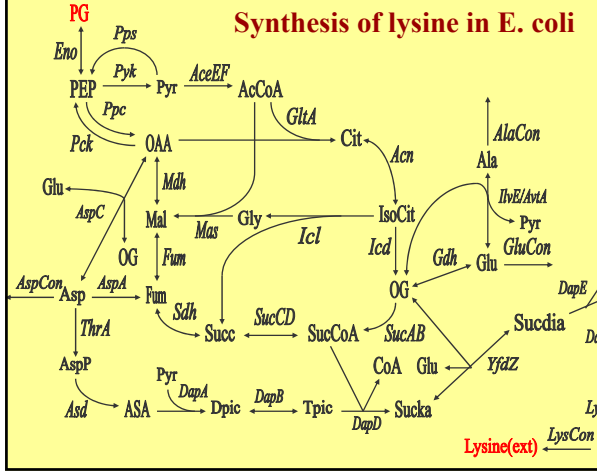
S. Schuster, D.A. Fell, T. Dandekar: *Nature Biotechnol.* 18 (2000) 326-332

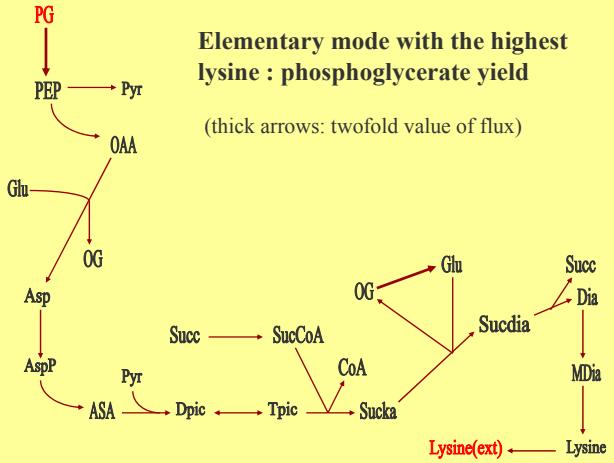
Optimization: Maximizing molar yields



ATP:G6P yield = 3 ATP:G6P yield = 2

Synthesis of lysine in E. coli





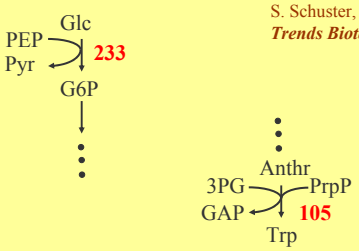
Elementary mode with the highest lysine : phosphoglycerate yield

(thick arrows: twofold value of flux)

Maximization of tryptophan:glucose yield

Model of 65 reactions in the central metabolism of *E. coli*. 26 elementary modes. 2 modes with highest tryptophan: glucose yield: 0.451.

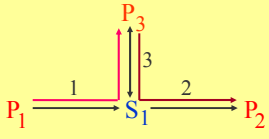
S. Schuster, T. Dandekar, D.A. Fell, ***Trends Biotechnol***. 17 (1999) 53



Convex basis

Minimal set of elementary modes sufficient to span the flux cone.

Example:

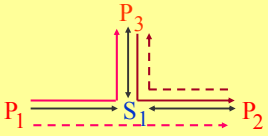


$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}$$

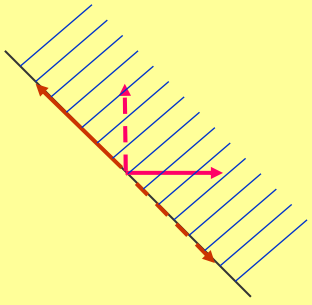
If the flux cone is pointed (all angles are less then 180°), then the convex basis is unique up to scaling.

Otherwise, it is not.

Example: Reactions 2 and 3 are reversible.



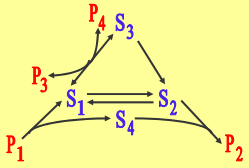
For the latter example, the flux cone is a half-plane:



The cone is not pointed.
Again, there are elementary modes in the interior.

Related concept: Extreme pathways

C.H. Schilling, D. Letscher and B.O. Palsson, *J. theor. Biol.* 203 (2000) 229
- distinction between internal and exchange reactions, all internal reversible reactions are split up into forward and reverse steps



Then, the convex basis is calculated.
Spurious cyclic modes are discarded.

Advantages of extreme pathways:
•Smaller number
•Correspond to edges of flux cone

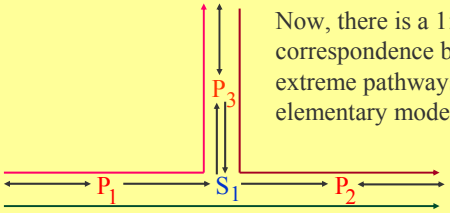
Drawbacks of extreme pathways:
•Flux cone is higher-dimensional
•Often not all relevant biochemical pathways represented
•Knock-outs not properly described
•Often route with maximal yield not covered

However, this depends on network configuration. Originally, Schilling et al. (2000) proposed adding exchange reaction for each external metabolite.

Network reconfiguration

1. Decomposition of internal reversible reactions into forward and reverse steps

2. Optionally: inclusion of (non-decomposed) exchange reactions for each external metabolite.



Now, there is a 1:1 correspondence between extreme pathways and elementary modes!

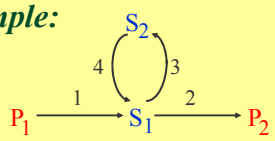
Algorithm for computing elementary modes

Related to Gauss-Jordan method

Starts with tableau ($\mathbf{N}^T$ $\mathbf{I}$)

Pairwise combination of rows so that one column of $\mathbf{N}^T$ after the other becomes null vector

Example:



$$\mathbf{T}^{(0)} = \left(\begin{array}{cc:cccc} 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

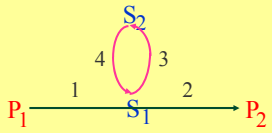
$$\mathbf{T}^{(0)} = \left(\begin{array}{cc:cccc} 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\mathbf{T}^{(1)} = \left(\begin{array}{cc:cccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

These two rows should not be combined

Final tableau:

$$\mathbf{T}^{(2)} = \left(\begin{array}{cc:cccc} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right)$$



Algorithm is faster, if this column is processed first.

$$\mathbf{T}^{(0)} = \left(\begin{array}{cc:cccc} 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Software for computing elementary modes

EMPATH (in SmallTalk) - J. Woods

METATOOL (in C) - Th. Pfeiffer, F. Moldenhauer, A. von Kamp, M. Pachkov

Included in GEPASI - P. Mendes and JARNAC - H. Sauro

part of METAFLUX (in MAPLE) - K. Mauch

part of FluxAnalyzer (in MATLAB) - S. Klamt

part of ScrumPy (in Python) - M. Poolman

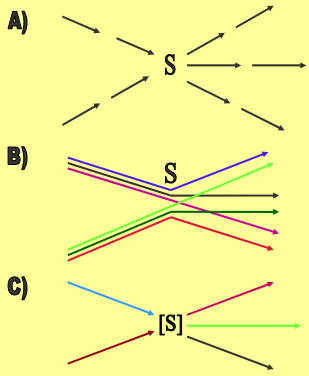
Alternative algorithm in MATLAB – C. Wagner (Bern)

On-line computation:

pHpMetatool - H. Höpfner, M. Lange

http://pgrc-03.ipk-gatersleben.de/tools/phpMetatool/index.php

Combinatorial explosion of elementary modes

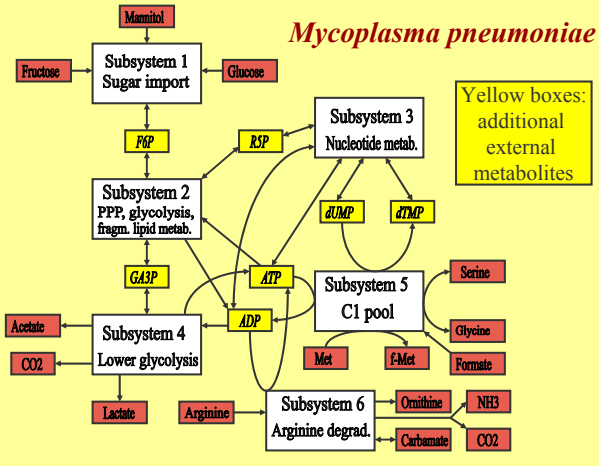


2*3 modes

S external: 2+3 modes

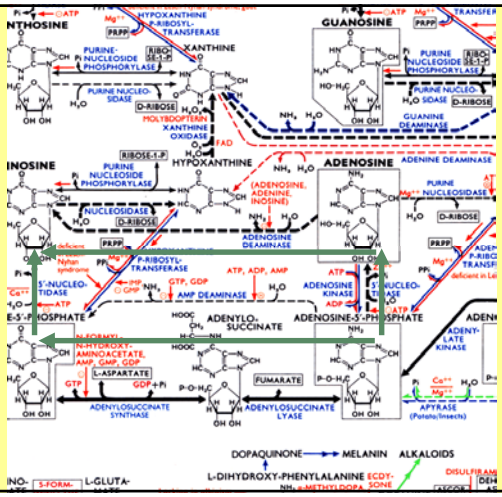
Proposed decomposition procedure

- **In addition to the pre-defined external metabolites, set all metabolites participating in more than 4 reactions to external status**
- **Thus, the network disintegrates into subnetworks**
- **Determine the elementary flux modes of the subnetworks separately**



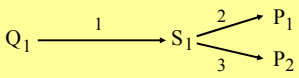
Robustness of metabolism

- Number of elementary modes leading from a given substrate to a given product can be considered as a measure of redundancy
- This is then also a rough estimate of robustness and of flexibility, because it characterizes the number of alternatives between which the network can switch if necessary



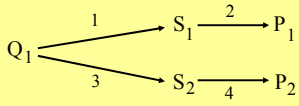
Number of elementary modes is not the best measure of robustness

A)



Knockout of enzyme 1 implies deletion of 2 elem. modes

B)



Proposed measure of network robustness

$$R_1 = \frac{\sum_{i=1}^{r} z^{(i)}}{r \cdot z}$$

r: number of reactions
z: number of elem. modes
z_i: number of elem. modes remaining after knockout of enzyme i.

T. Wilhelm, J. Behre and S. Schuster: Analysis of structural robustness of metabolic networks. *Syst. Biol.*, 1 (2004) 114 - 120.

Metabolic network/ essential products	Number of elementary modes	R_1 (robustness)
Human erythrocyte		
ATP, hypoxanthine, NADPH, 2,3DPG	667	0.3834
E. coli		
Ala, Arg, Asn, His§	667	0.5084
Arg, Asn, His, Ile	656	0.5211
Arg, Asn, Ile, Leu	567	0.5479
Arg, Asn, Leu, Pro	540	0.5360
His, Ile, Leu, Lys	802	0.5112
Ile, Leu, Pro, Val	597	0.5488

Conclusions

- **Elementary modes are an appropriate concept to describe biochemical pathways in wild-type and mutants. Complies with irreversibility constraints.**
- **Information about network structure can be used to derive far-reaching conclusions about performance of metabolism, e.g. about viability of mutants.**
- **Elementary modes reflect specific characteristics of metabolic networks such as steady-state mass flow, thermodynamic constraints and molar yields.**

Conclusions (2)

- **Pathway analysis is well-suited for computing maximal and submaximal molar yields**
- **Relevant medical application: enzyme deficiencies**
- **Worthwile investigating double and triple mutants.**
- **Work still to be done on decomposition methods (combinatorial explosion)**

Cooperations

•Steffen Klamt, Jörg Stelling, Ernst Dieter Gilles (MPI Magdeburg)
•Thomas Dandekar (U Würzburg)
•David Fell (Brookes U Oxford)
•Thomas Pfeiffer, Sebastian Bonhoeffer (ETH Zürich)

•Peer Bork (EMBL Heidelberg)
•Reinhart Heinrich, Thomas Höfer (HU Berlin)
•Hans Westerhoff (VU Amsterdam)
• and others

•Acknowledgement to DFG and BMBF for financial support

Applications of elementary-modes analysis by other authors

Rohwer & Botha, Analysis of sucrose accumulation in the sugar cane culm on the basis of in vitro kinetic data.
Biochem. J. 358 (2001) 437.

Förster, Gombert, & Nielsen, A functional genomics approach using metabolomics and in silico pathway analysis.
Biotechnol. Bioeng. 79 (2002) 703.

Van Dien & Lidstrom, Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1, with application to reconstruction of C(3) and C(4) metabolism.
Biotechnol. Bioeng. 78 (2002) 296.

Applications of elementary-modes analysis by other authors (2)

Carlson, Fell & Srienc, Metabolic pathway analysis of a recombinant yeast for rational strain development.
Biotechnol. Bioeng. 79 (2002) 121.

Poolman, Fell & Raines, Elementary modes analysis of photosynthate metabolism in the chloroplast stroma.
Eur. J. Biochem. 270 (2003) 430.

and meanwhile several more…