seit 1558

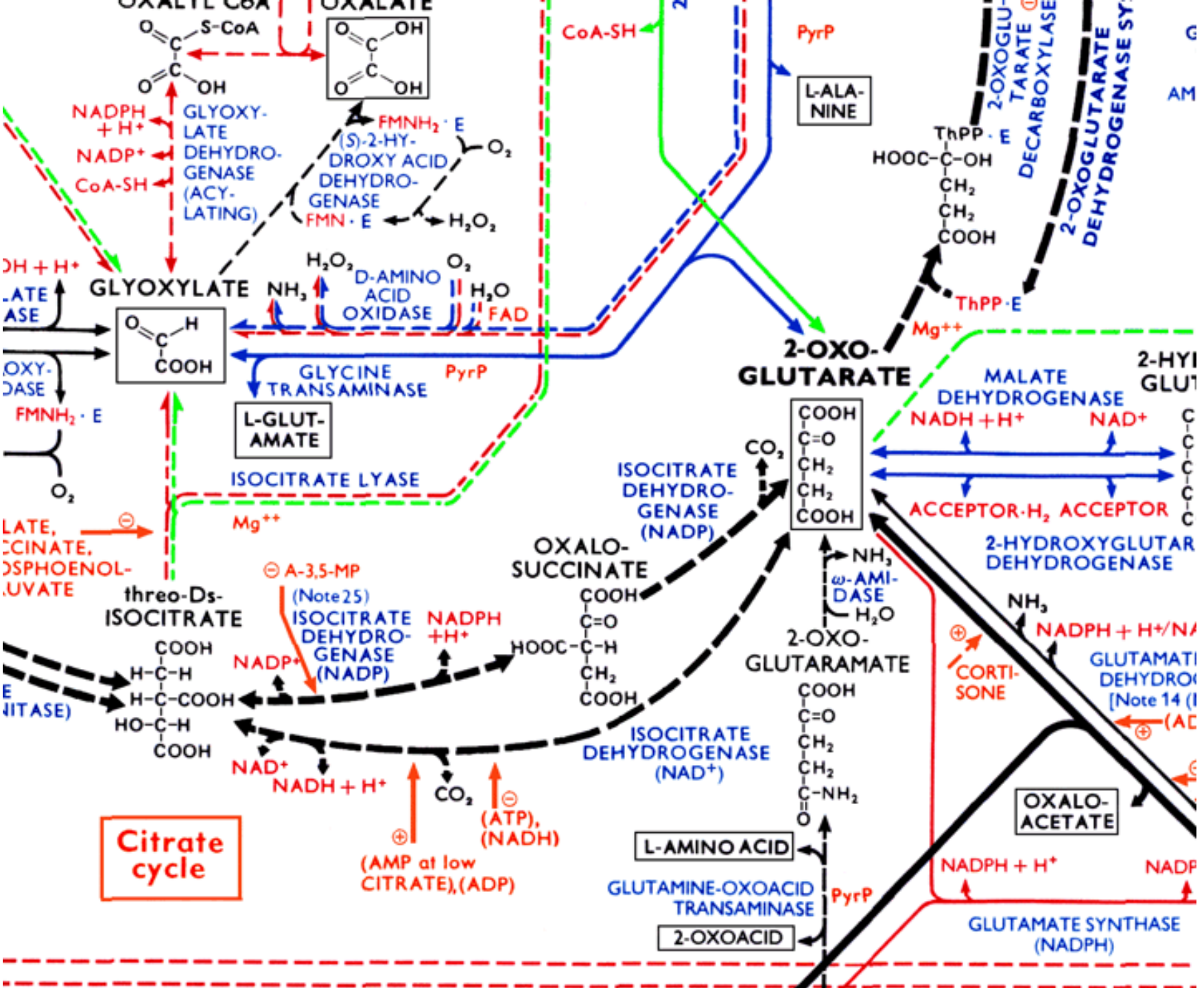
JCB
JENA CENTRE FOR BIOINFORMATICS

Réseaux métaboliques et modes élémentaires

Stefan Schuster

Friedrich Schiller University Jena

Dept. of Bioinformatics


OXALYL CoA
OXALATE
S-CoA
NADPH + H+
NADP+
CoA-SH
GLYOXYLATE DEHYDROGENASE (ACYLATING)
FMNH2 · E
(S)-2-HYDROXY ACID DEHYDROGENASE
FMN · E
H2O2
O2
CoA-SH
PyrP
L-ALANINE
ThPP · E
2-OXOGLUTARATE DECARBOXYLASE
2-OXOGLUTARATE DEHYDROGENASE SYSTEM
GLYOXYLATE
NH3
D-AMINO ACID OXIDASE
H2O
FAD
GLYCINE TRANSAMINASE
L-GLUTAMATE
2-OXOGLUTARATE
Mg++
MALATE DEHYDROGENASE
NADH + H+
NAD+
ACCEPTOR·H2
ACCEPTOR
2-HYDROXYGLUTARATE DEHYDROGENASE
FMNH2 · E
ISOCITRATE LYASE
ISOCITRATE DEHYDROGENASE (NADP)
CO2
OXALOSUCCINATE
NH3
ω-AMIDASE
H2O
2-OXOGLUTARAMATE
A-3,5-MP
(Note 25)
threo-Ds-ISOCITRATE
NADPH + H+
NADP+
ISOCITRATE DEHYDROGENASE (NADP)
CORTISONE
NADPH + H+/NA
GLUTAMATE DEHYDROG
[Note 14 (
NAD+
NADH + H+
CO2
(ATP), (NADH)
(AMP at low CITRATE), (ADP)
ISOCITRATE DEHYDROGENASE (NAD+)
OXALOACETATE
L-AMINO ACID
GLUTAMINE-OXOACID TRANSAMINASE
PyrP
2-OXOACID
NADPH + H+
NADP
GLUTAMATE SYNTHASE (NADPH)
Citrate cycle

Introduction

- **Analysis of metabolic systems requires theoretical methods due to high complexity**
- **Major challenge: clarifying 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**

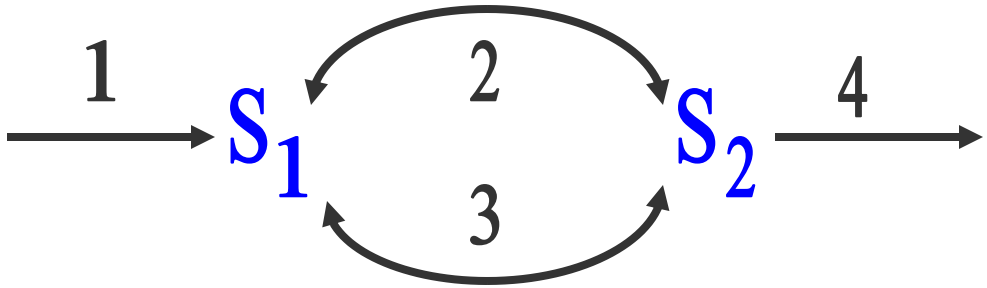
History of pathway analysis

- „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:

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

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

At any stationary state, this simplifies to:

$$\mathbf{N}\mathbf{V}(\mathbf{S}) = \mathbf{0}$$

Kernel of N

Steady-state condition $\mathbf{NV(S) = 0}$

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.

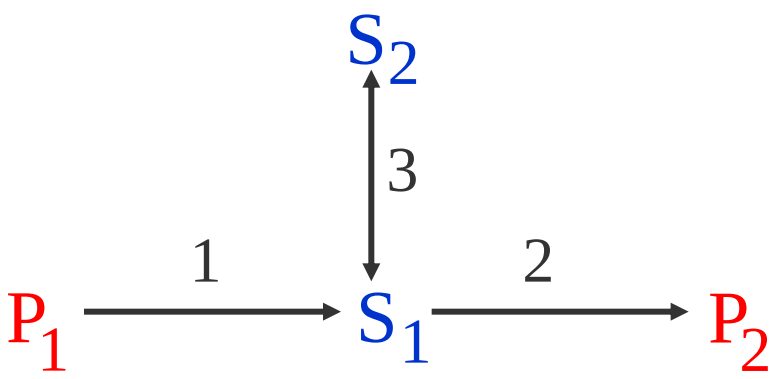
Mathematically: kernel (null-space) of **N**. Spanned by basis vectors. These are not unique.

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:



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

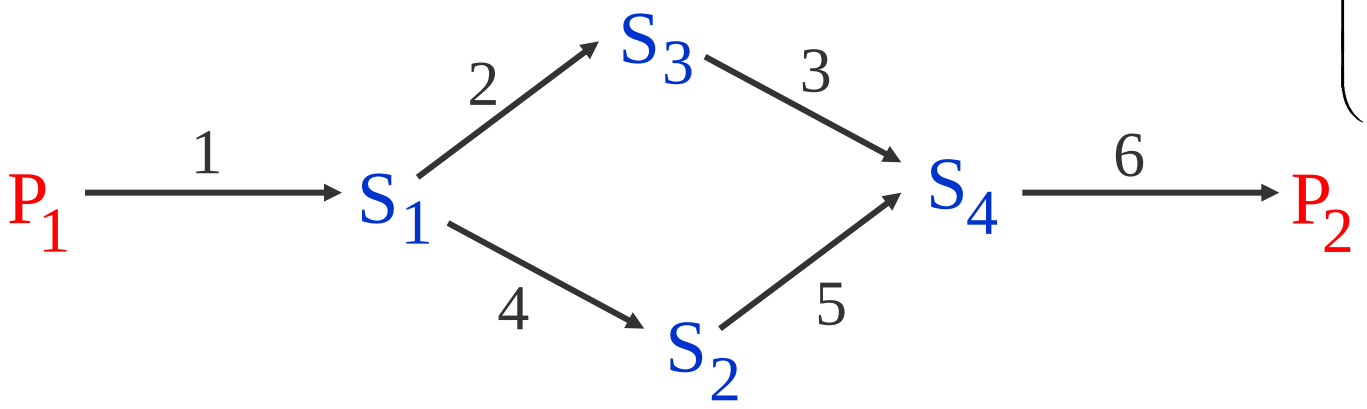
Use of null-space (2)

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}

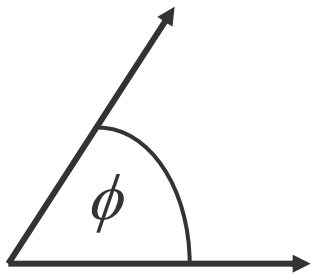
$$\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$



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

Extensions of the concept of „enzyme subsets“

Representation of rows of null-space matrix as vectors in space:



If $\cos(\phi) = \quad 1$, then the enzymes belong to the same subset

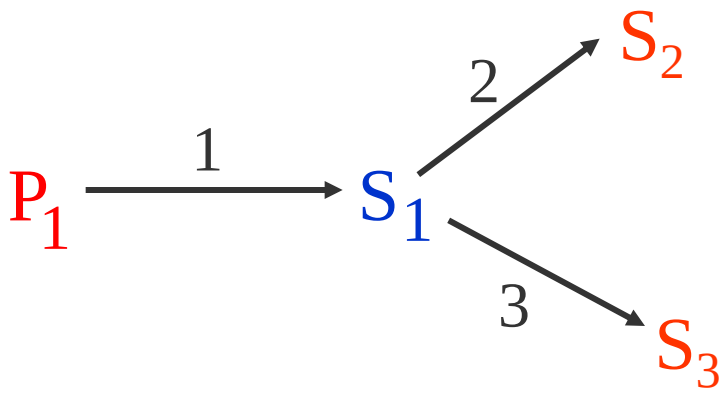
If $\cos(\phi) = 0$, then reactions uncoupled

Otherwise, enzymes partially coupled.

M. Poolman et al., *J. theor. Biol*. 249 (2007) 691–705

Extensions of the concept of „enzyme subsets“ (2)

Inclusion of information about irreversibility



If all reactions are irreversible, operation of enzyme 2 implies operation of enzyme 1.

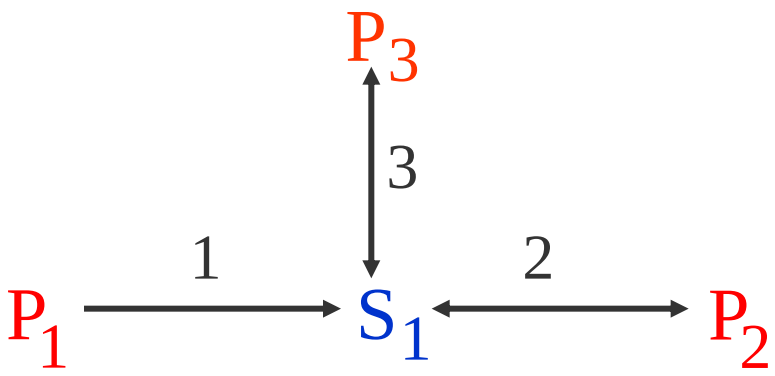
(1) **Directional coupling** ($v_1 \rightarrow v_2$), if a non-zero flux for v_1 implies a non-zero flux for v_2 but not necessarily the reverse.
(2) **Partial coupling** ($v_1 \leftrightarrow v_2$), if a non-zero flux for v_1 implies a non-zero, though variable, flux for v_2 and vice versa.
(3) **Full coupling** ($v_1 \Leftrightarrow v_2$), if a non-zero flux for v_1 implies not only a non-zero but also a fixed flux for v_2 and vice versa. – Enzyme subset.

Flux coupling analysis

A.P. Burgard et al. *Genome Research* 14 (2004) 301-312.

Drawbacks of null-space

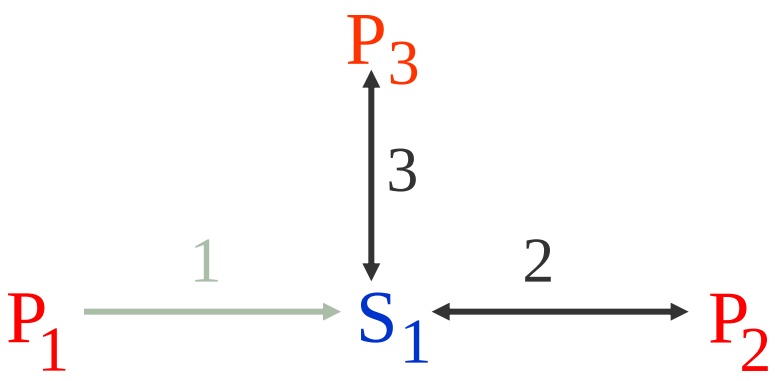
- The basis vectors are not given uniquely.
- They are not necessarily the simplest possible.
- They do not necessarily comply with the directionality of irreversible reactions.
- They do not always properly describe knock-outs.



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

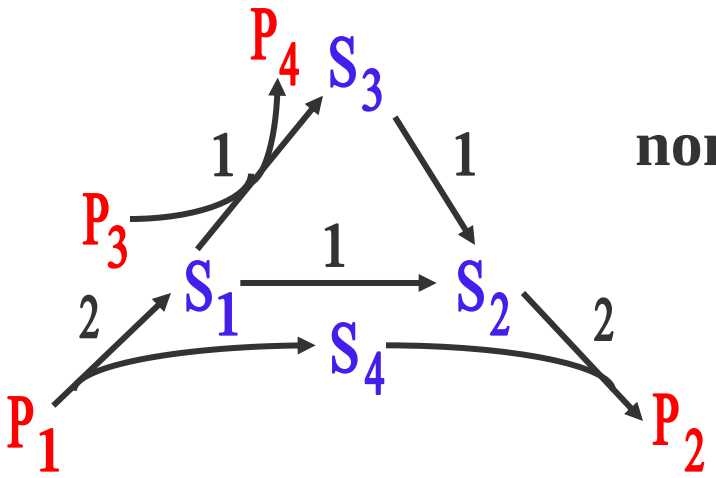
Drawbacks of null-space

They do not always properly describe knock-outs.

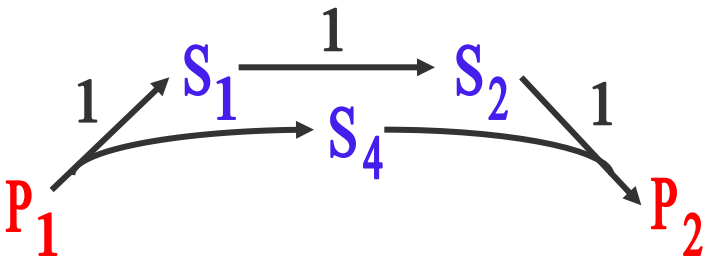


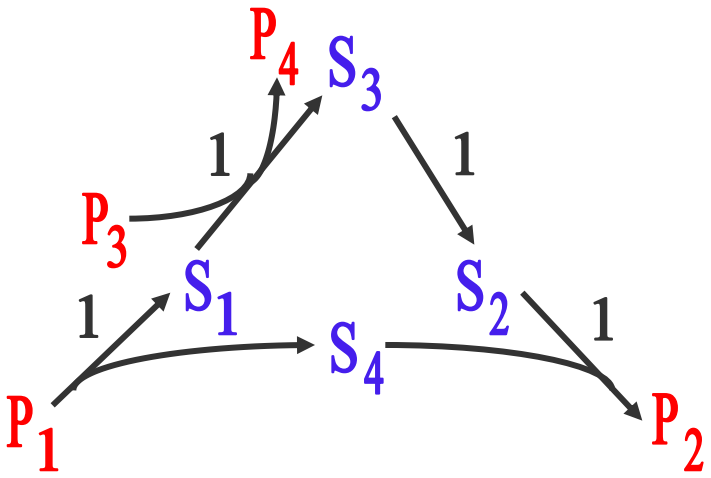
$$\mathbf{K} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

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
“ et al., *Nature Biotechnol*. 18 (2000) 326-332.

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.

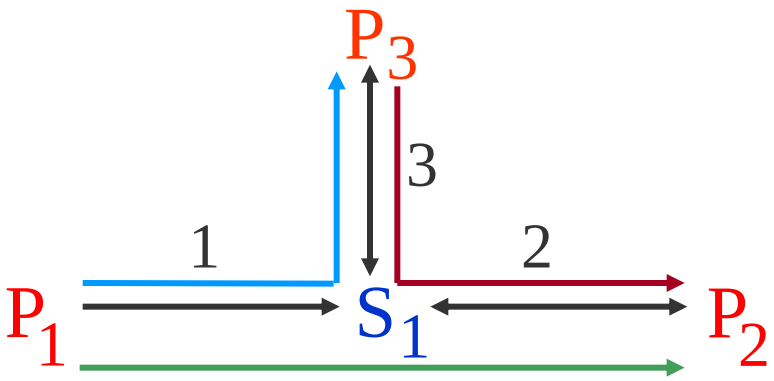
All flux distributions in the living cell are non-negative linear combinations of elementary modes

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: $\mathbf{V^{irr} \geqq 0}$

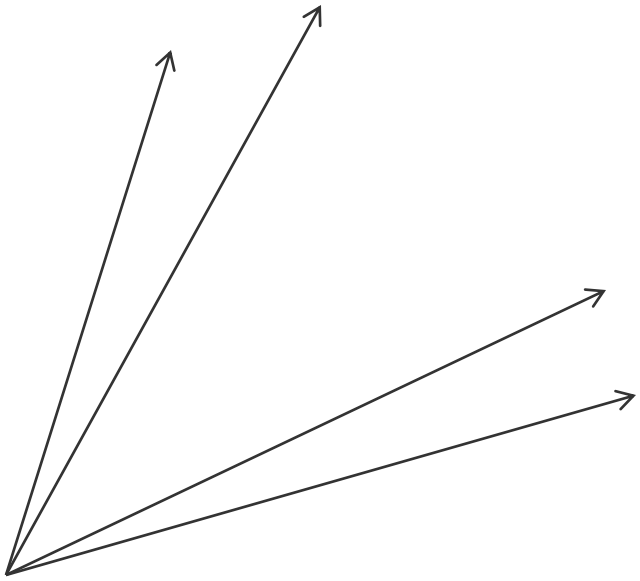
This represents a linear equation/inequality system.

Solution is a convex region.

All edges correspond to elementary modes.

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

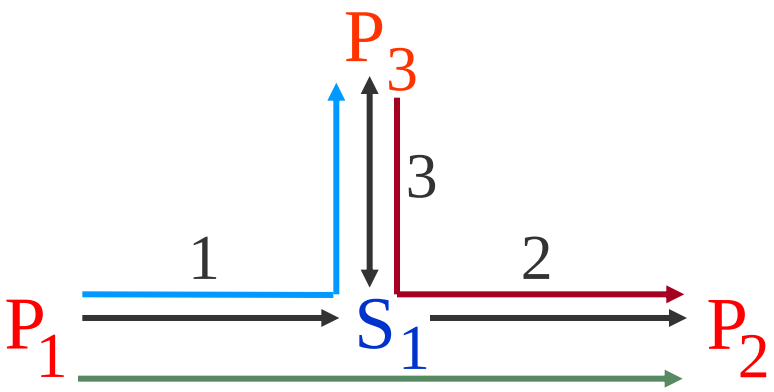
Geometrical interpretation



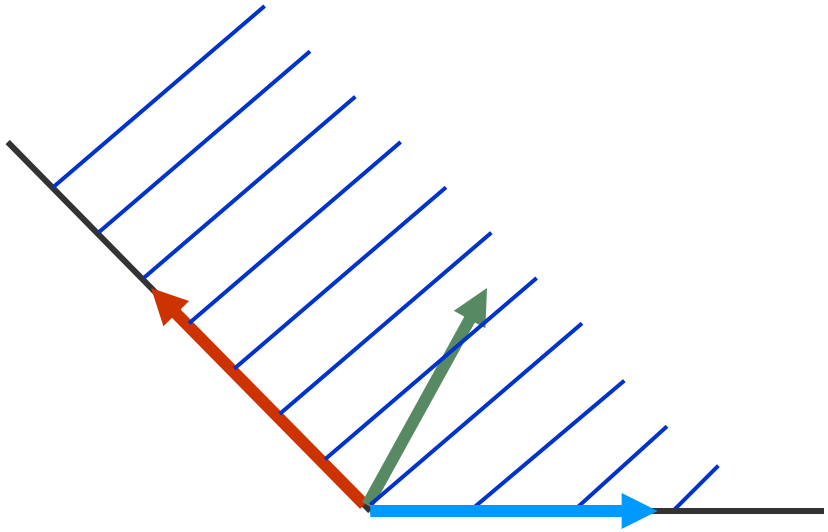
Elementary modes correspond to generating vectors (edges) of a convex polyhedral cone (= pyramid) in flux space (if all reactions are irreversible)

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

Example:



Flux cone:

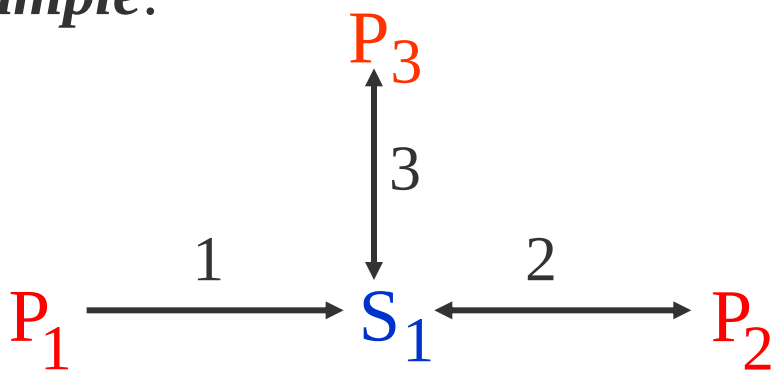


There are elementary modes in the interior of the cone.

Mathematical properties of elementary modes

Any vector representing an elementary mode involves at least dim(null-space of **N**) − 1 zero components.

Example:



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

dim(null-space of **N**) = 2

Elementary modes:

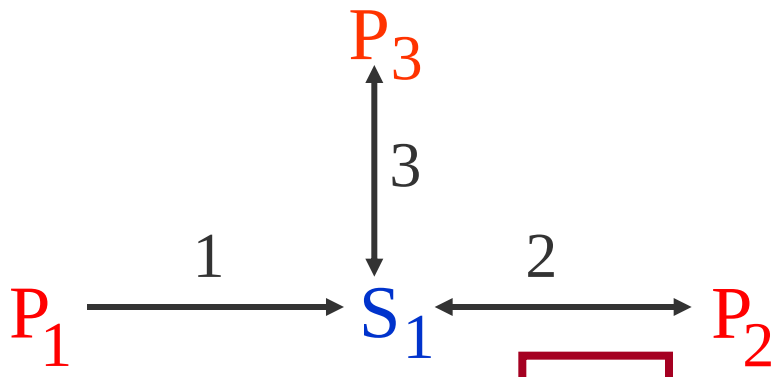
$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

(Schuster et al., *J. Math. Biol.* 2002, after results in theoretical chemistry by Milner et al.)

Mathematical properties of elementary modes (2)

A flux mode **V** is elementary if and only if the null-space of the submatrix of **N** that only involves the reactions of **V** is of dimension one.

Klamt, Gagneur und von Kamp, *IEE Proc. Syst. Biol*. 2005, after results in convex analysis by Fukuda et al.

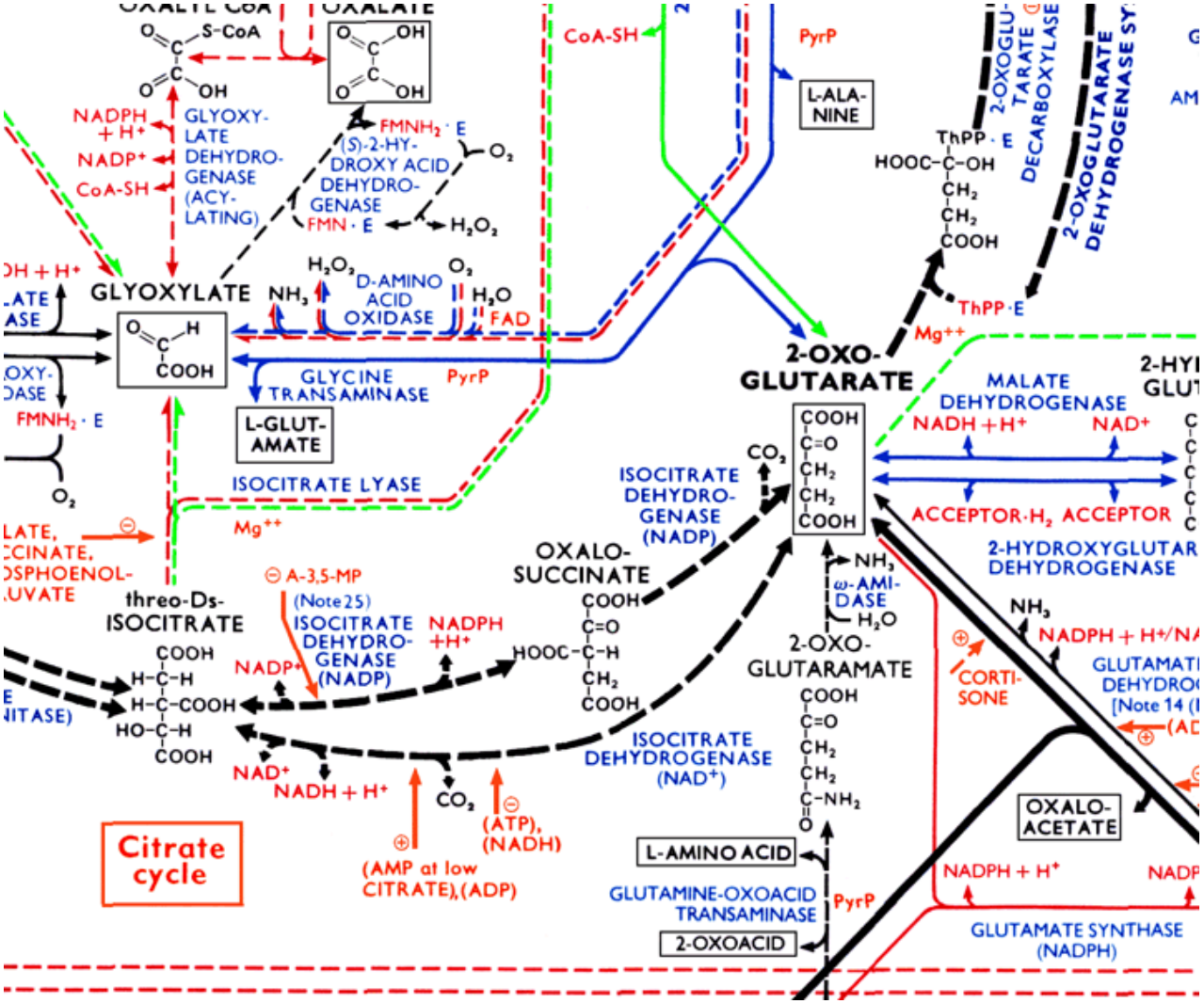


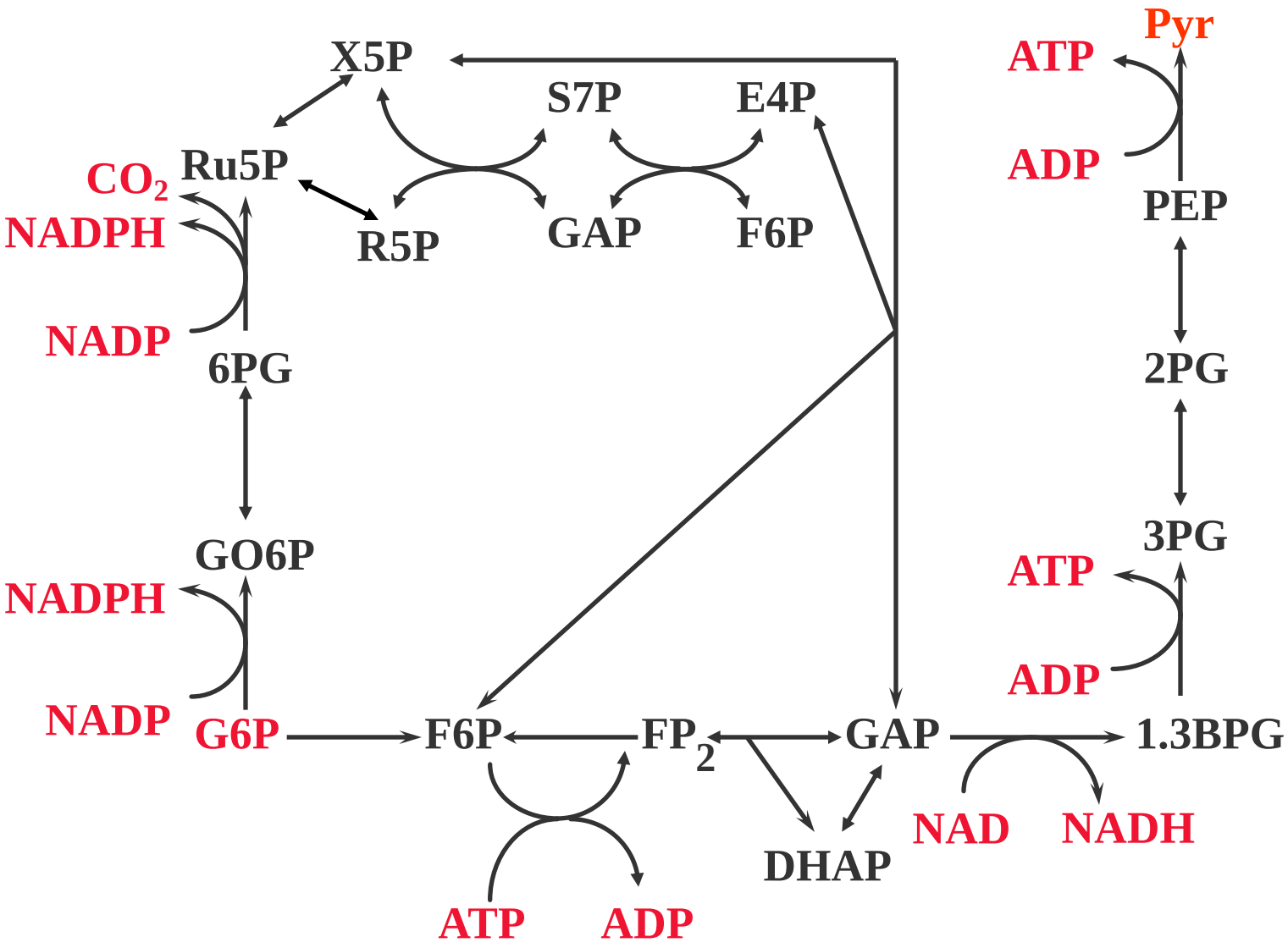
e.g. elementary mode:

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

N = (1 –1) → dim = 1

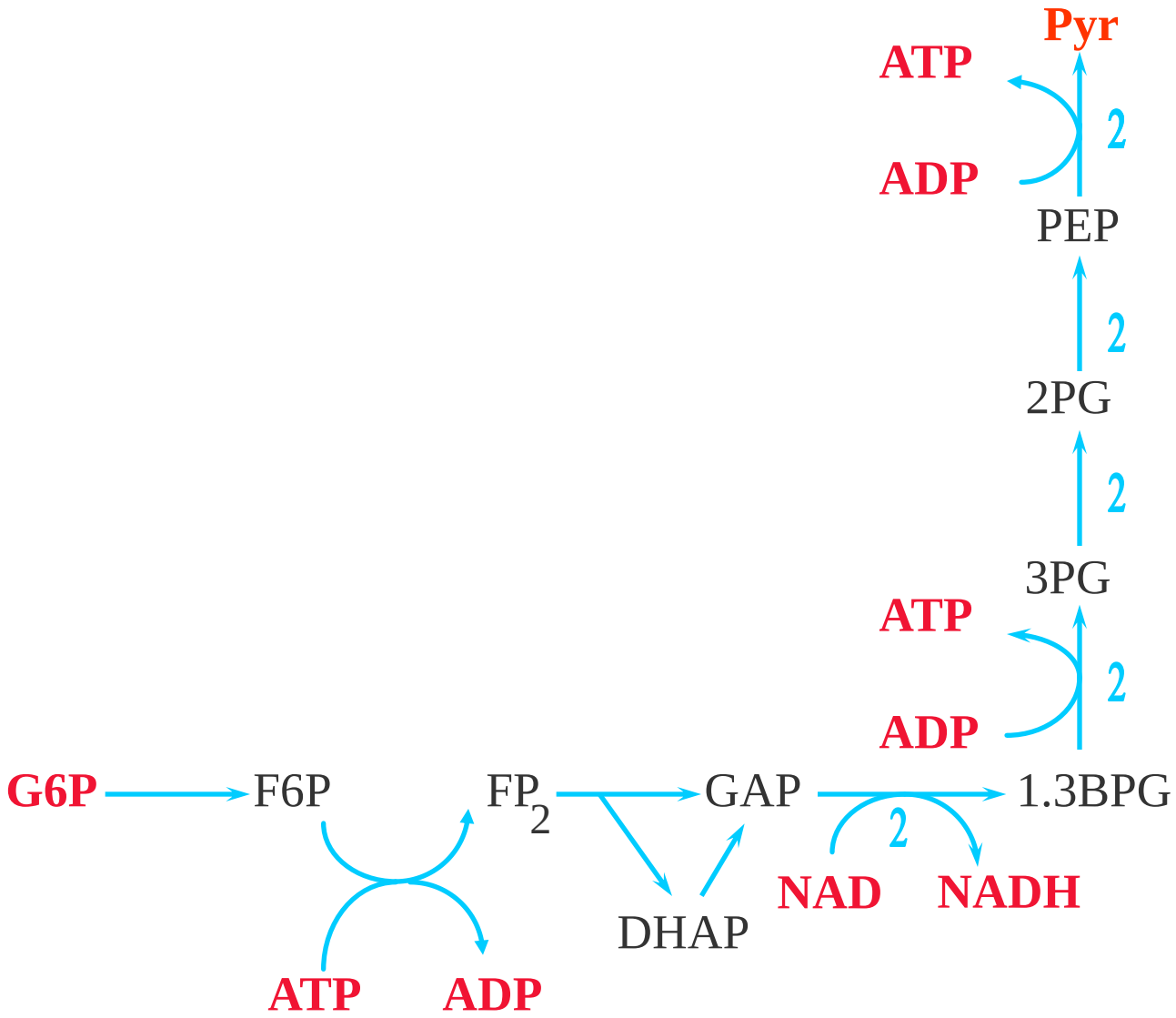
Biochemical examples



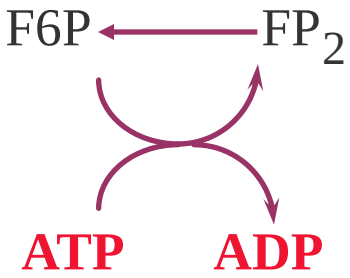


Part of monosaccharide metabolism

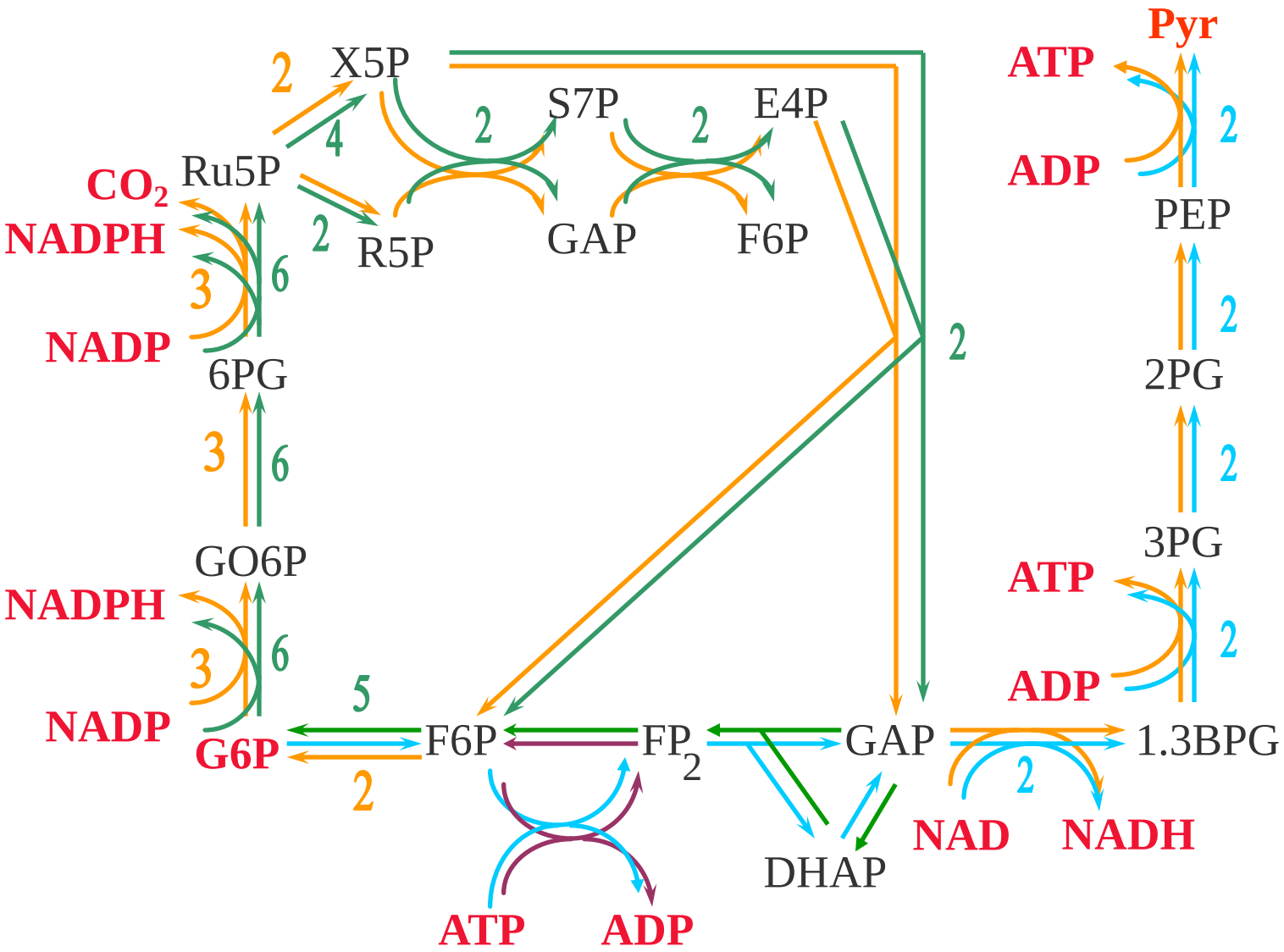
Red: external metabolites



1st elementary mode: glycolysis

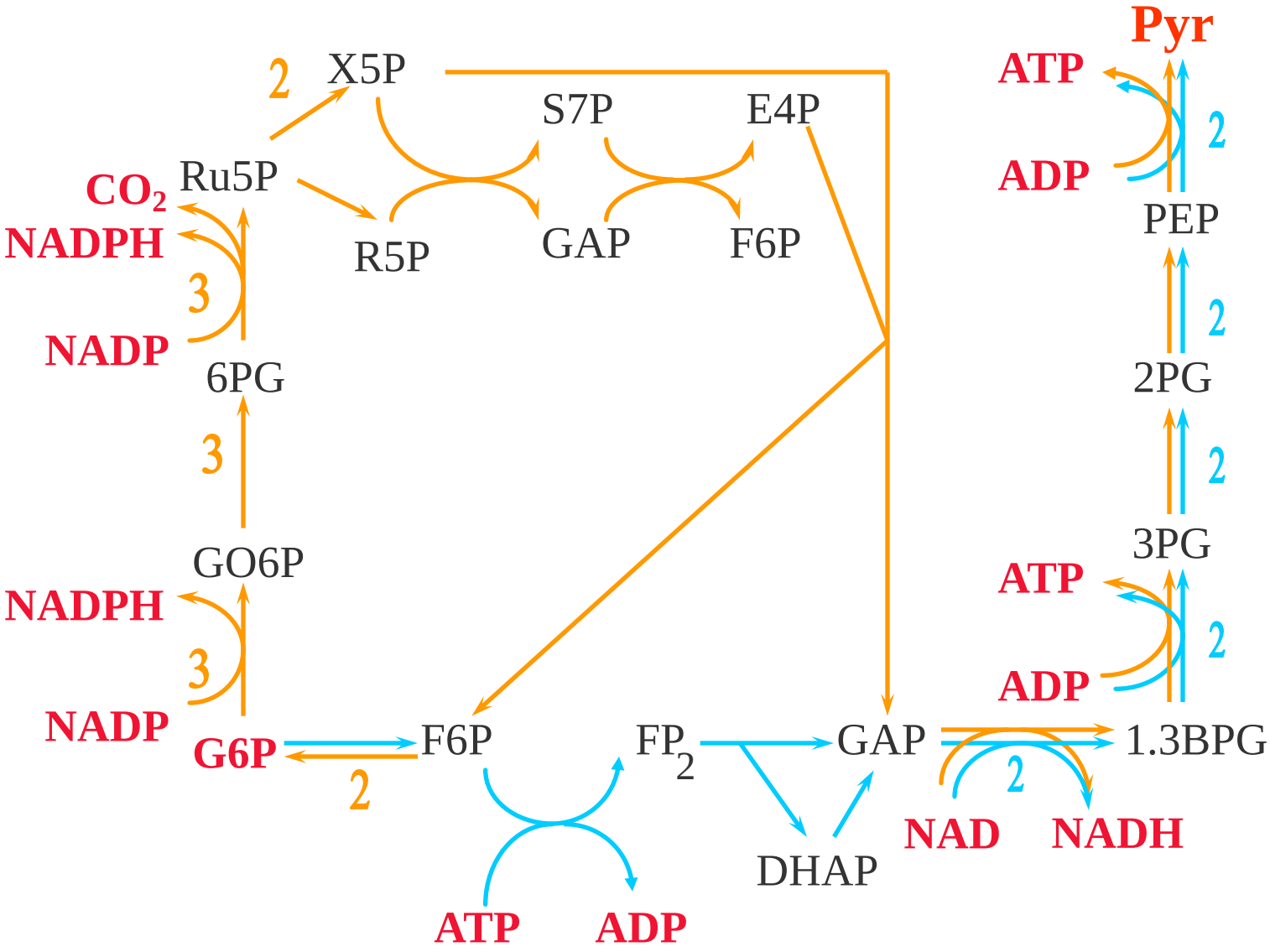


2nd elementary mode: fructose-bisphosphate cycle



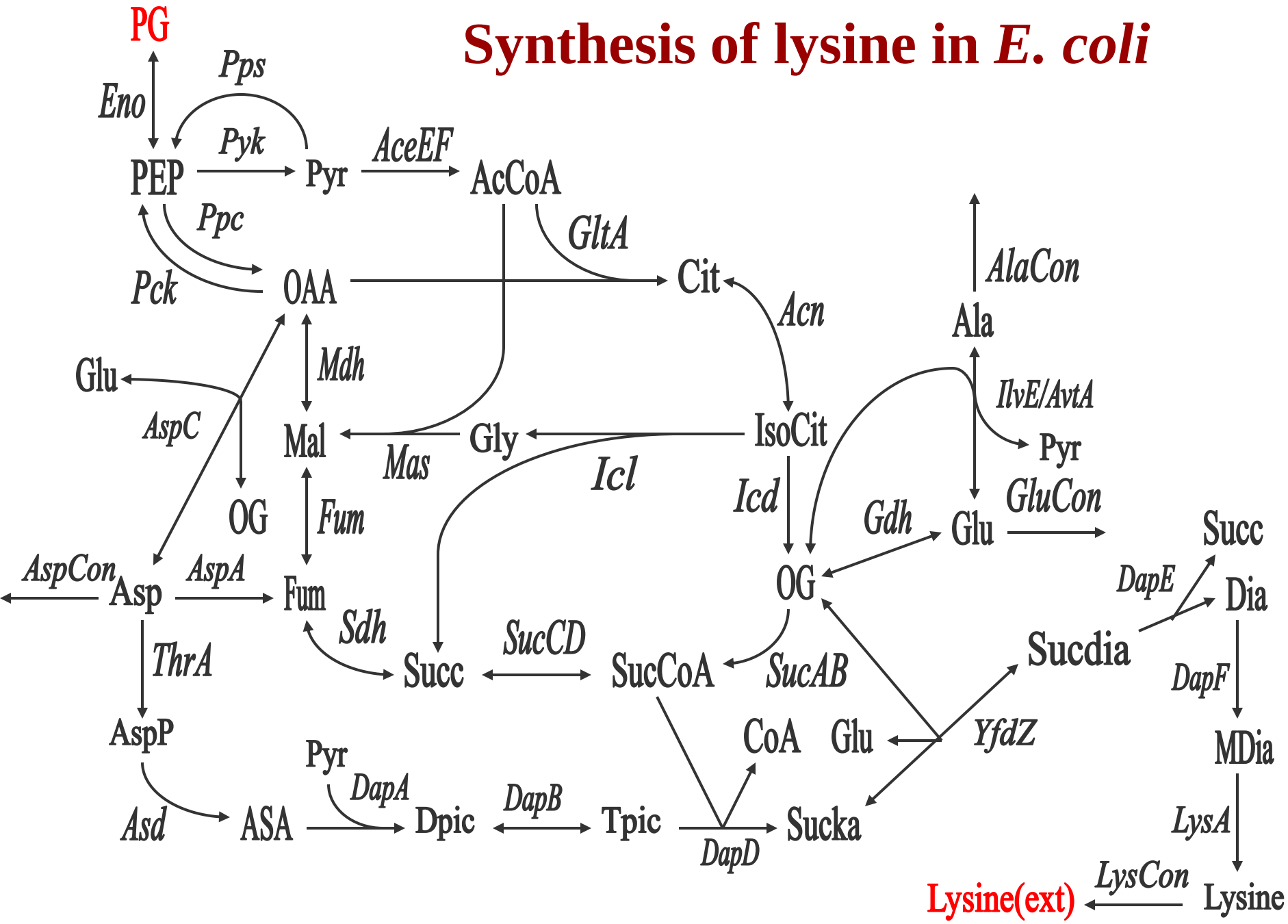
4 out of 7 elementary modes in glycolysis-pentose-phosphate system

Optimization: Maximizing molar yields



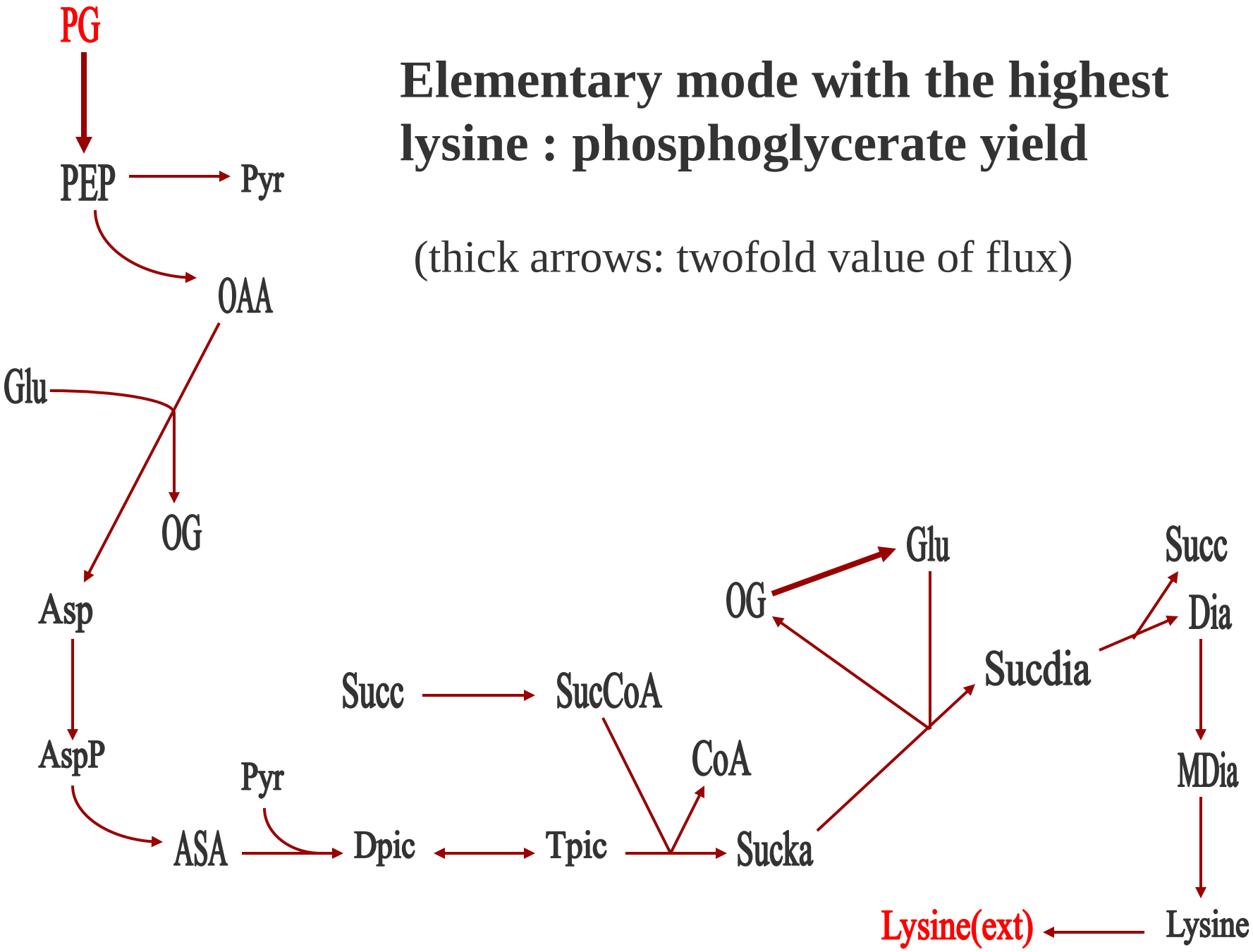
ATP:G6P yield = 3 **ATP:G6P yield = 2**

Synthesis of lysine in *E. coli*



Elementary mode with the highest lysine : phosphoglycerate yield

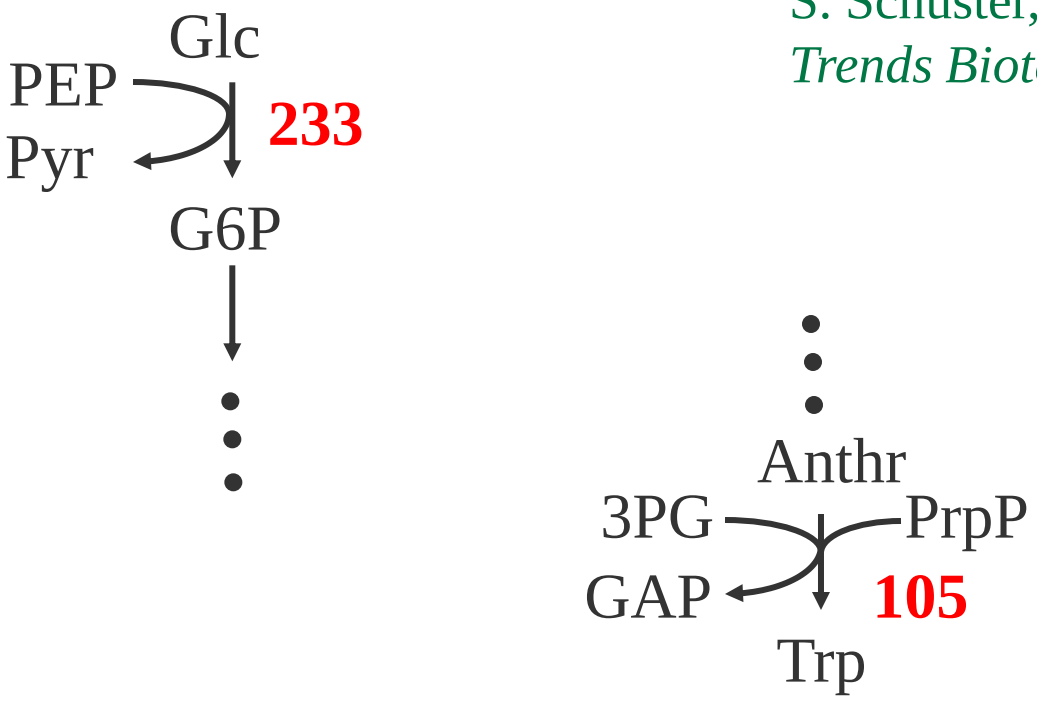
(thick arrows: twofold value of flux)



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



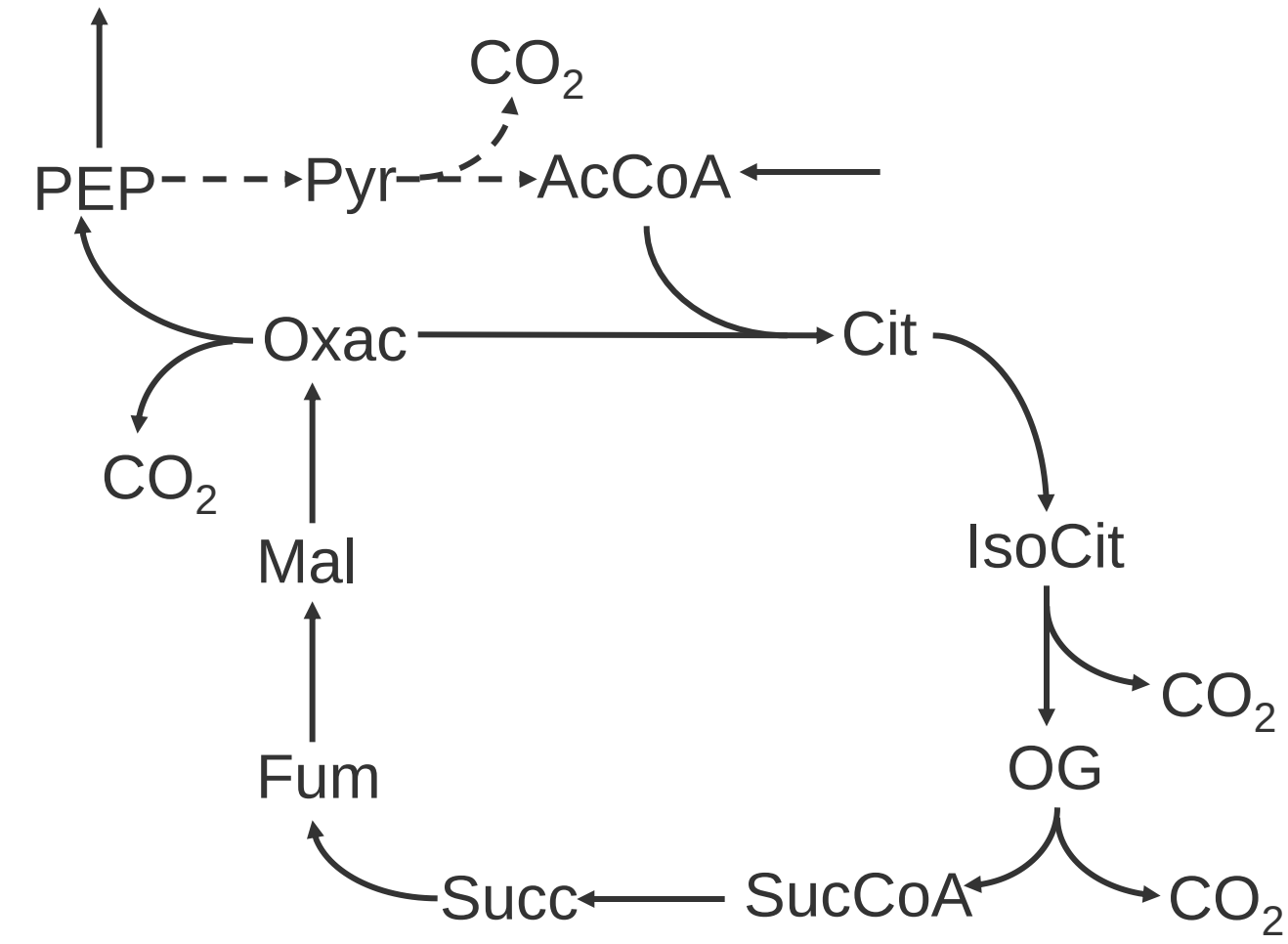
Can fatty acids be transformed into sugar?

- Excess sugar in human diet is converted into storage lipids, mainly triglycerides
- Is reverse transformation feasible? Triglyceride → sugar?

Triglycerides

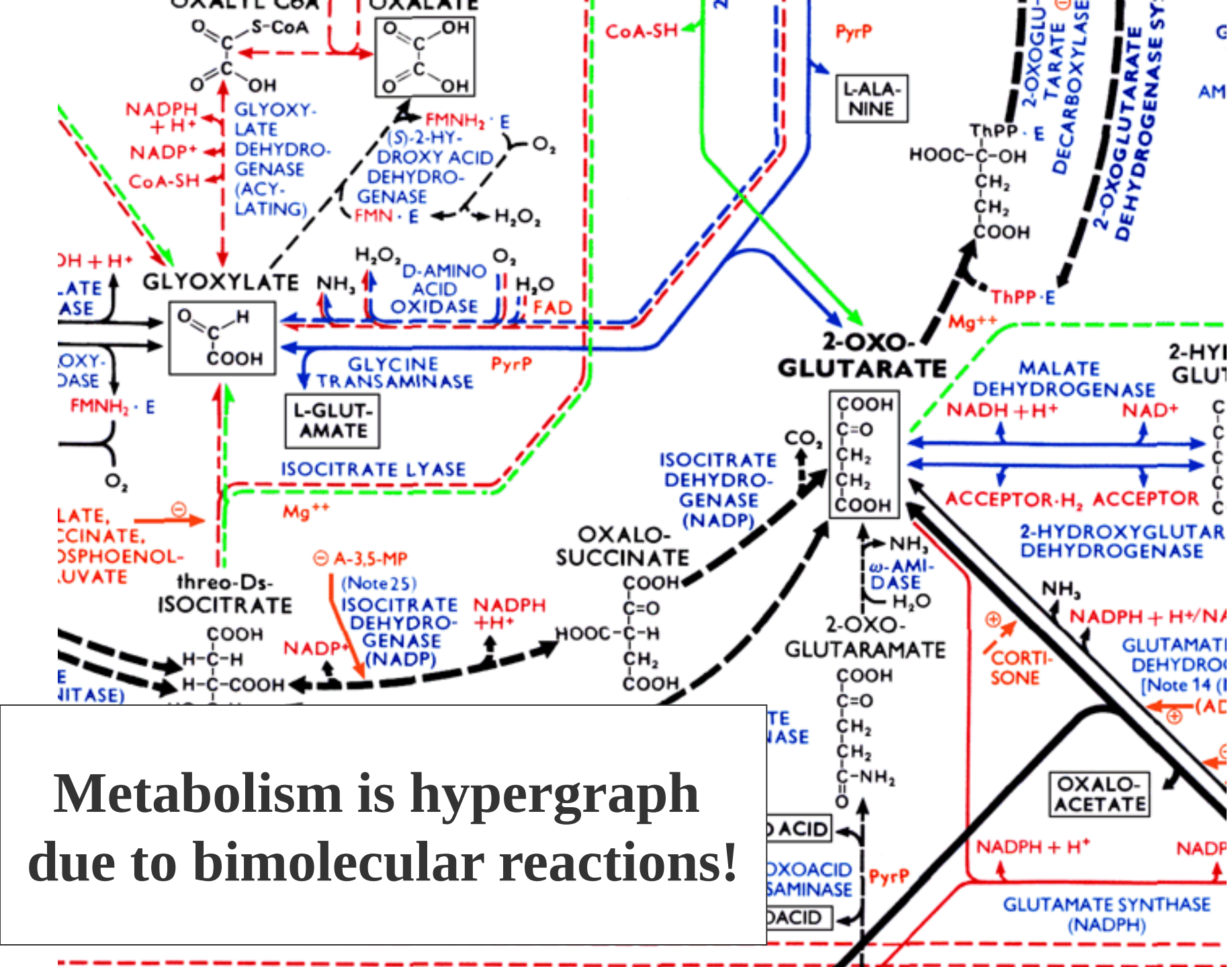
- 1 glycerol + 3 even-chain fatty acids (odd-chain fatty acids only in some plants and marine organisms)
- Glycerol → glucose OK (gluconeogenesis)
- (Even-chain) fatty acids → acetyl CoA (β-oxidation)
- Acetyl CoA → glucose?

Exact reversal of glycolysis and AcCoA formation is impossible because pyruvate dehydrogenase and some other enzymes are irreversible. Nevertheless, AcCoA is linked with glucose by a chain of reactions via the TCA cycle.



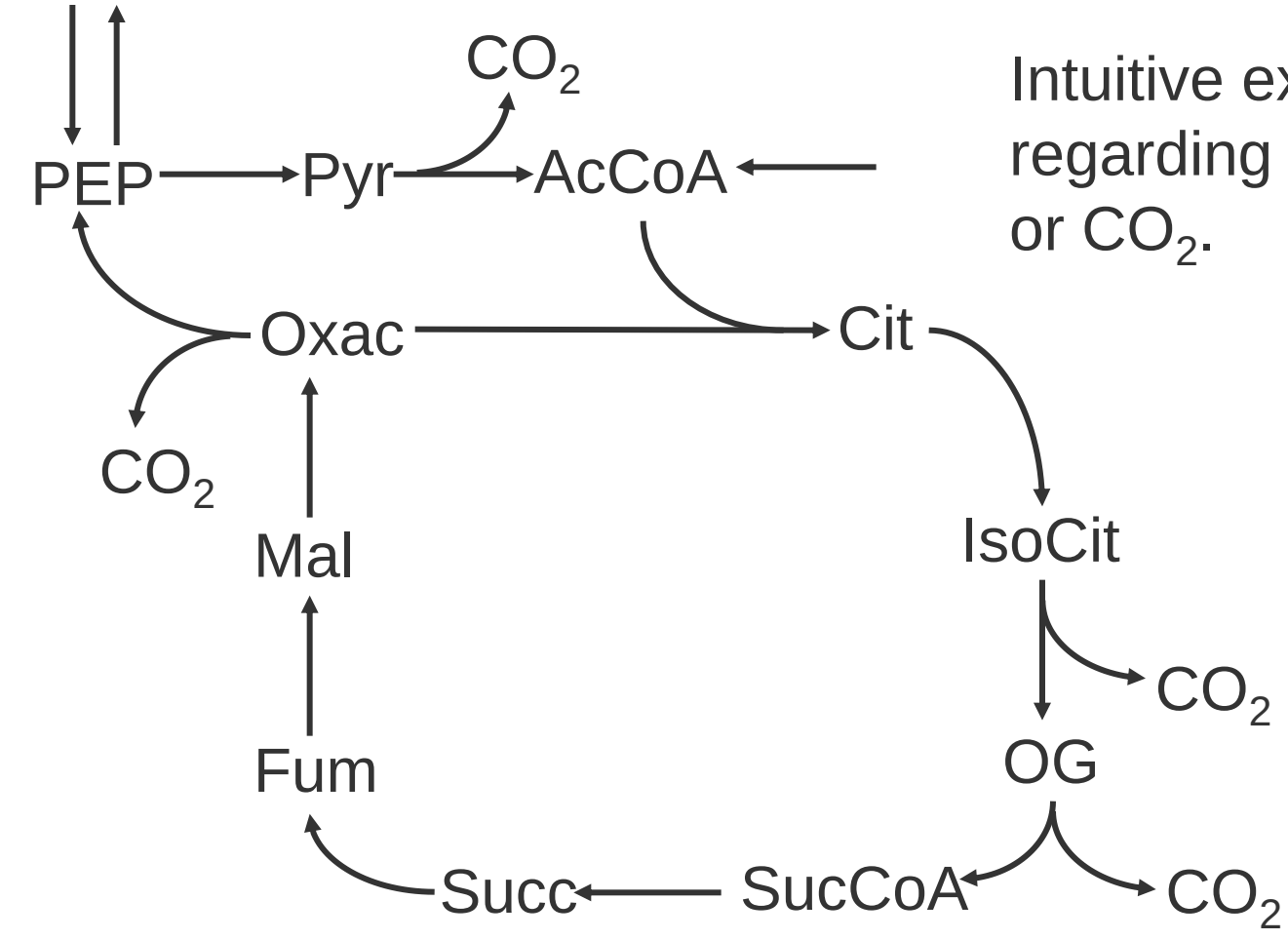
Graph theory vs. experiment

- By graph theory, it may be assumed that the conversion in question would be feasible.
- Experimental observation: If fatty acids are radioactively labelled, part of tracer indeed arrives at glucose.
- However, sustained formation of glucose at steady state is observed in humans only at very low rates.



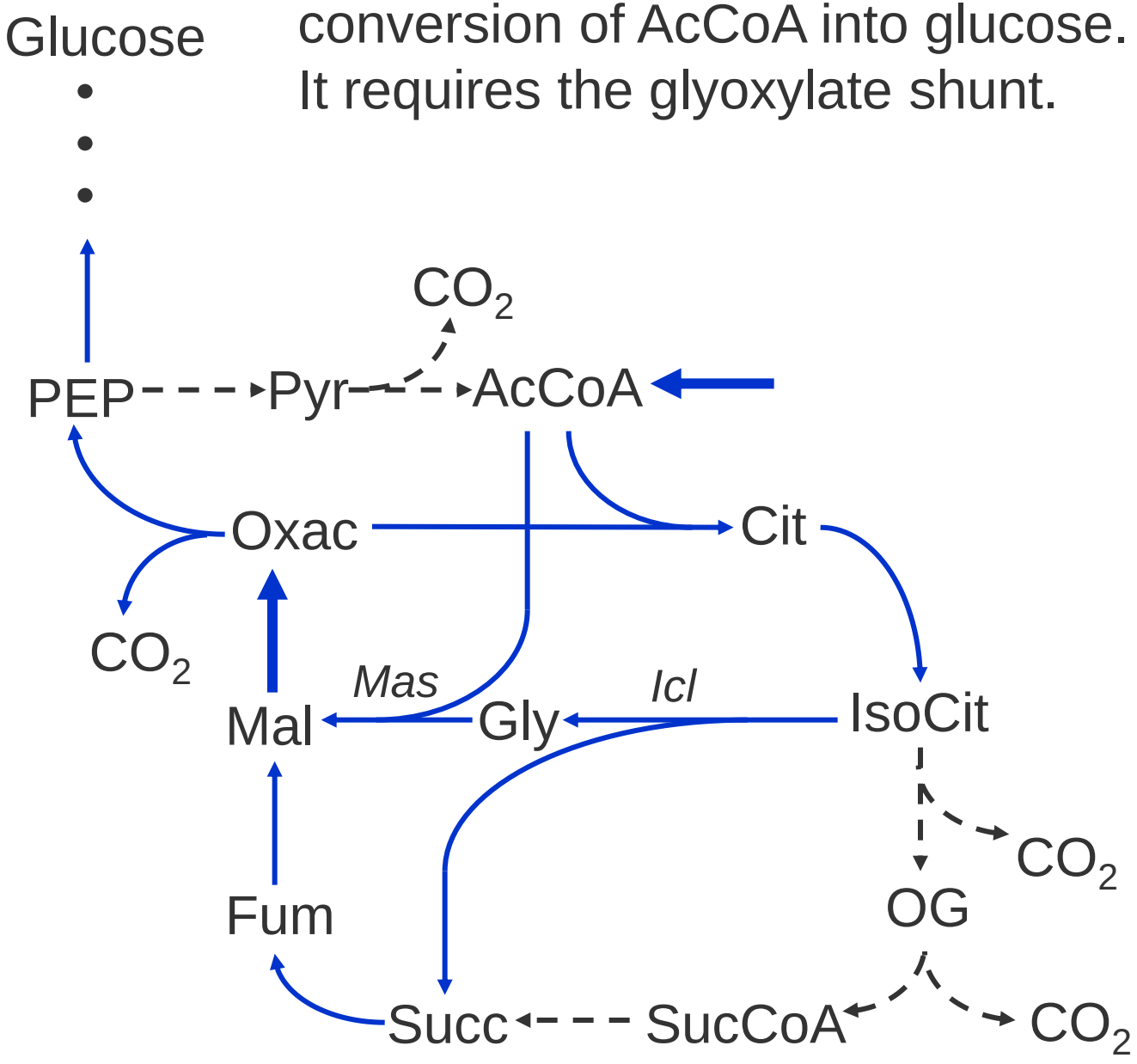
Metabolism is hypergraph due to bimolecular reactions!

If AcCoA, glucose, CO_2 and all cofactors are considered external, there is NO elementary mode consuming AcCoA, nor any one producing glucose.



Intuitive explanation by regarding oxaloacetate or CO_2.

Elementary mode representing conversion of AcCoA into glucose. It requires the glyoxylate shunt.



Animals versus plants

- Green plants can what we can‘t.
- Sugar is storage substance.
- In animals: brain cells, red blood cells and many other cells feed on glucose. Thus, starvation is a problem…
- Animals who died from starvation may still have fat reservoirs.

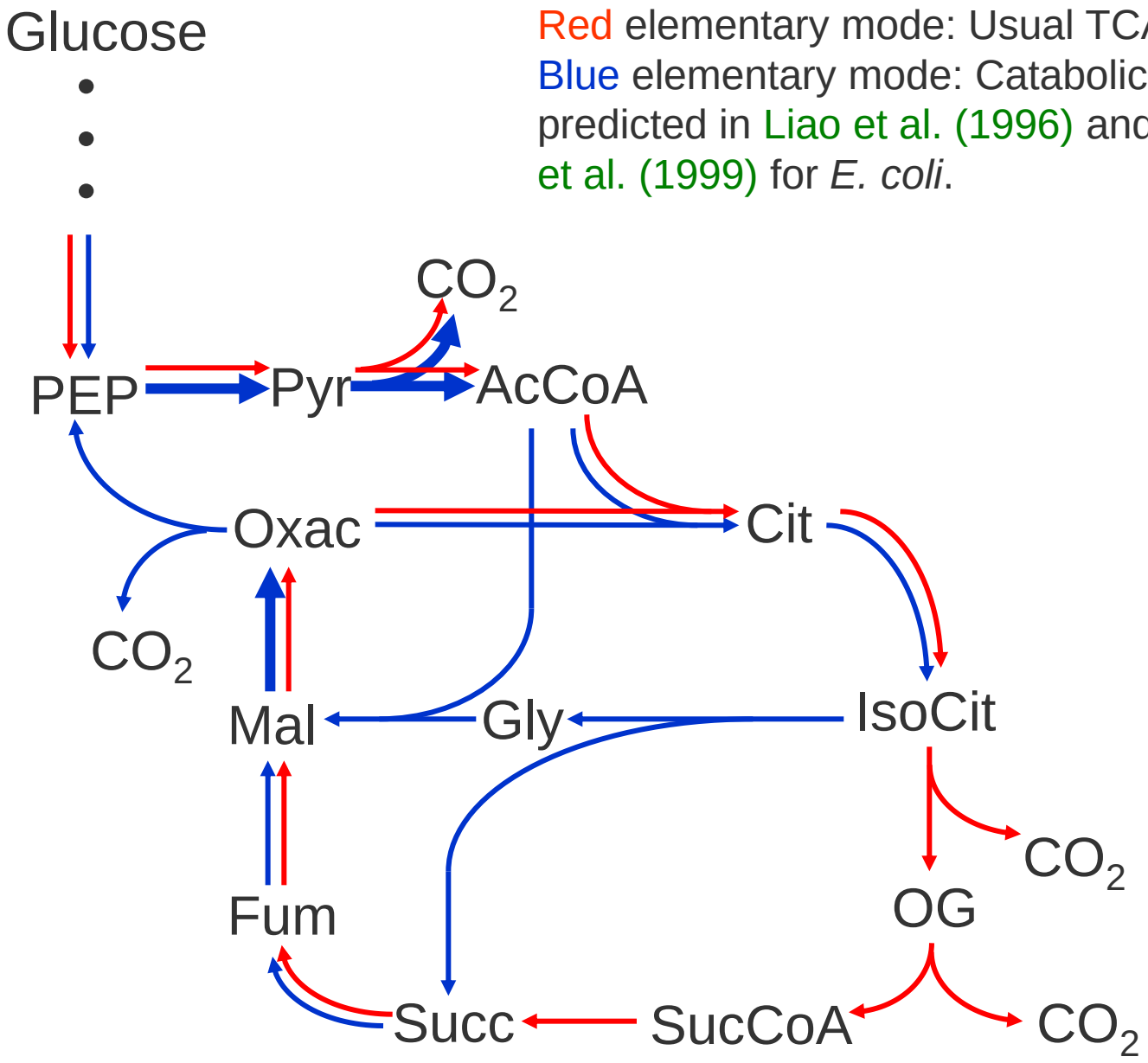
The glyoxylate shunt is present in green plants, yeast, many bacteria (e.g. *E. coli*) and others and – as the only clade of animals – in nematodes.

This example shows that a description by usual graphs in the sense of graph theory is insufficient…

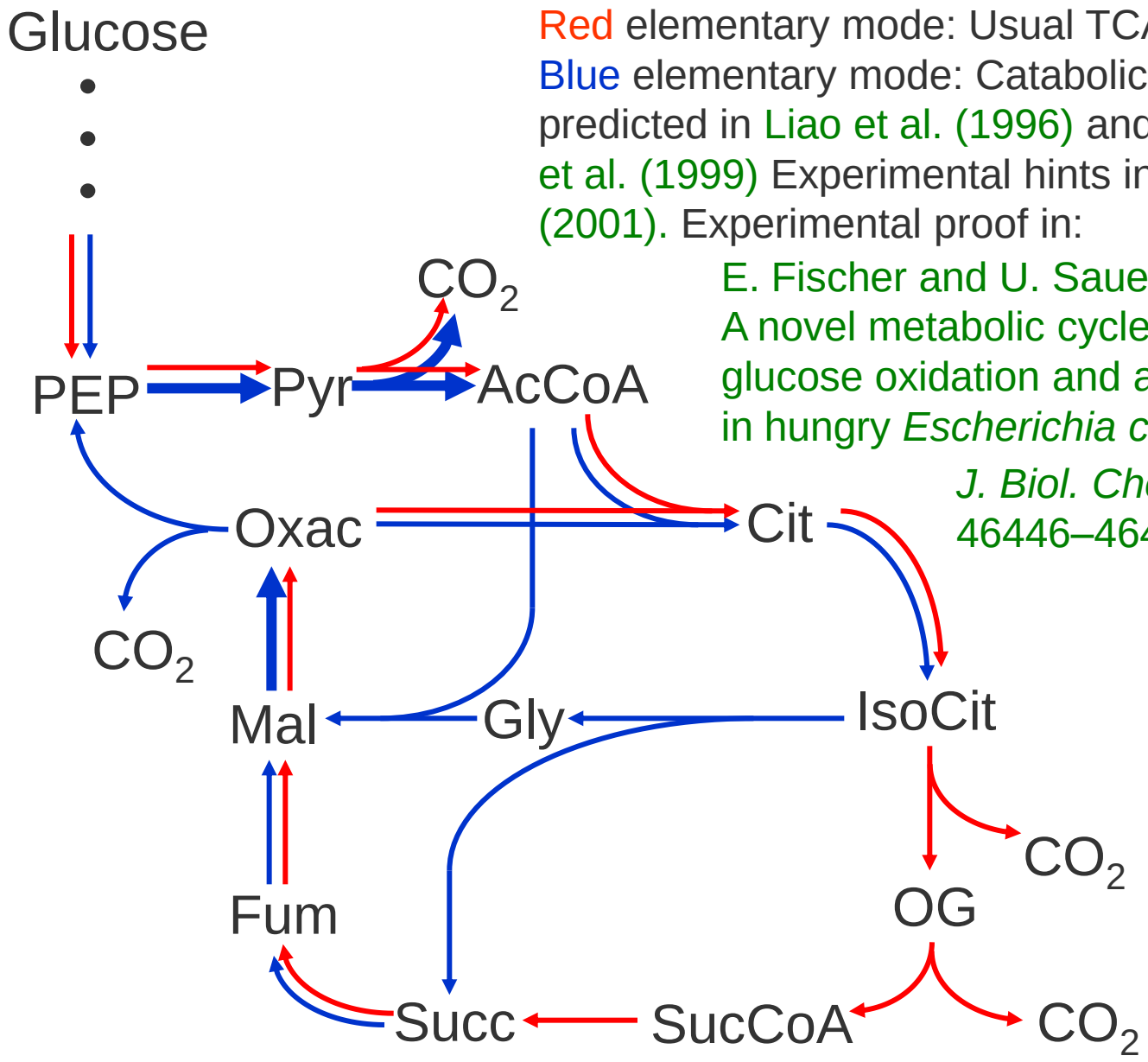
S. Schuster, D.A. Fell: Modelling and simulating metabolic networks. In: *Bioinformatics: From Genomes to Therapies* (T. Lengauer, ed.) Wiley-VCH, Weinheim 2007, pp. 755-805.

L. Figuereido, S. Schuster, C. Kaleta, D.A. Fell: Can sugars be produced from fatty acids? *Bioinformatics*, under revision

A successful theoretical prediction



A successful theoretical prediction



Red elementary mode: Usual TCA cycle
Blue elementary mode: Catabolic pathway predicted in Liao et al. (1996) and Schuster et al. (1999) Experimental hints in Wick et al. (2001). Experimental proof in:

E. Fischer and U. Sauer:
A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry *Escherichia coli,*

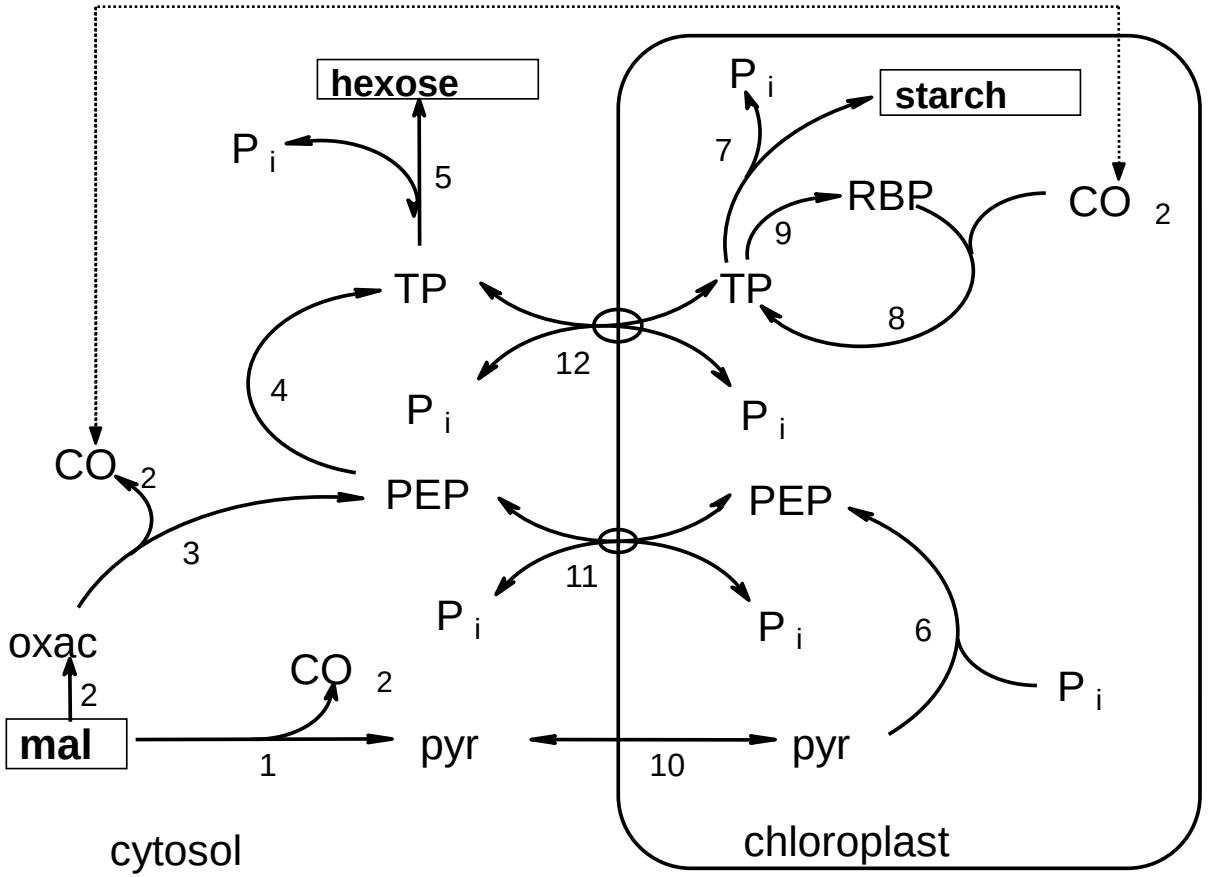
J. Biol. Chem. 278 (2003) 46446–46451

Crassulacean Acid Metabolism (CAM)

(Work with David Fell, Oxford)

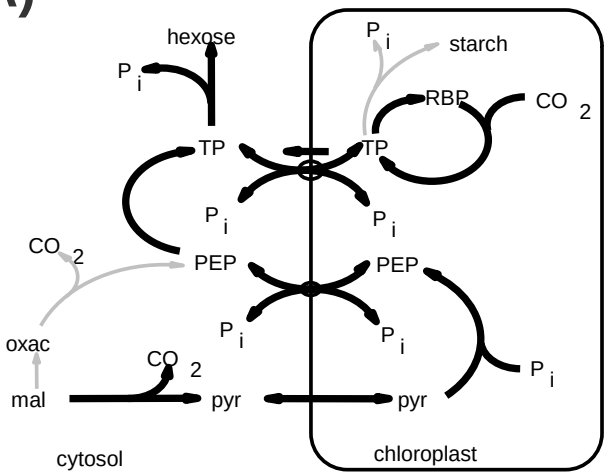
- Variant of photosynthesis employed by a range of plants (e.g. cacti) as an adaptation to arid conditions
- To reduce water loss, stomata are closed during daytime
- At nighttime, PEP + CO_2 → oxaloacetate → malate
- At daytime, malate → pyruvate (or PEP) + CO_2 → carbohydrates

CAM metabolism during daytime



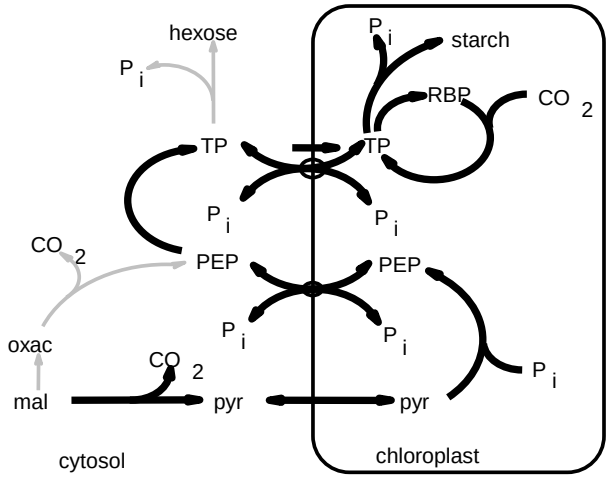
Elementary modes

A)



Hexose synthesis via malic enzyme as occurring in Agavaceae and Dracaenaceae

B)

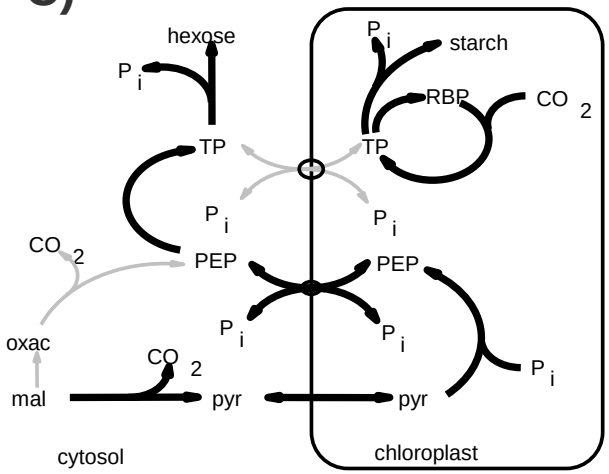


Starch synthesis via malic enzyme as occurring in Cactaceae and Crassulacea

Dracaena

Ferocactus

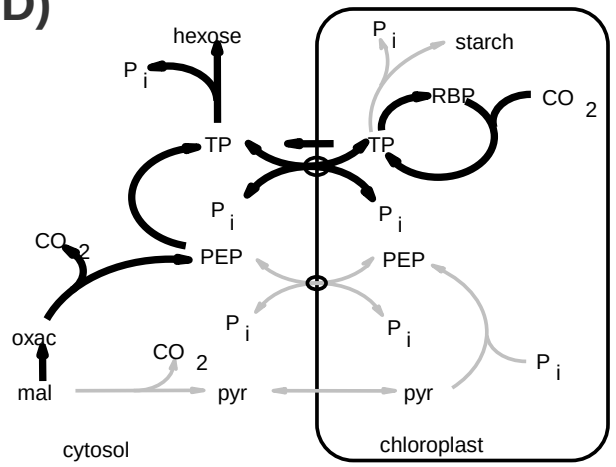
C)



Simultaneous starch and hexose synthesis via malic enzyme as occurring in:

Clusia minor

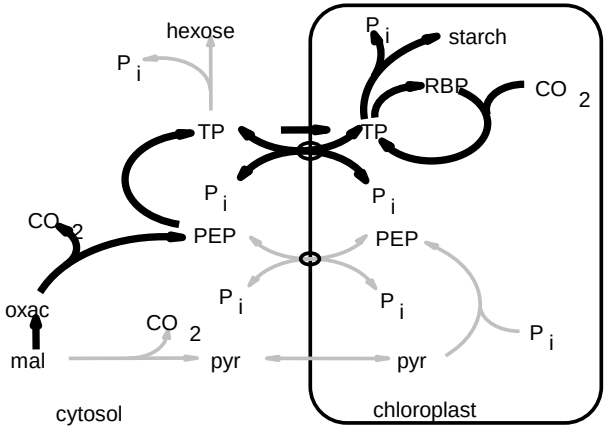
D)



Hexose synthesis via PEPCK as occurring in *Clusia rosea* and in:

Ananus comosus = pineapple

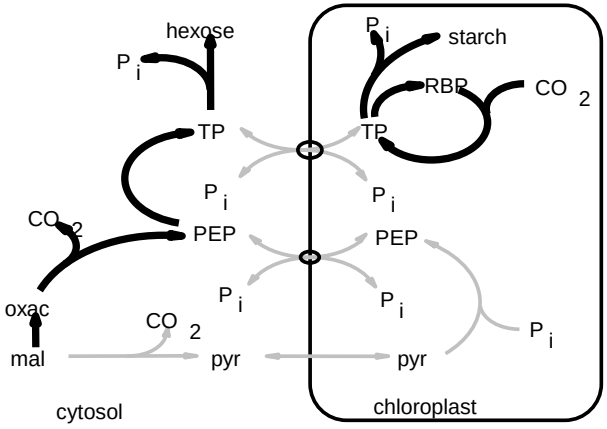
E)



Starch synthesis via PEPCK as occurring in Asclepidiaceae

Caralluma hexagona

F)



Simultaneous starch and hexose synthesis via PEPCK as occurring in:

Aloe vera

„Pure“ pathways

- In a review by Christopher and Holtum (1996), only cases A), B), D), and E) were given as “pure” functionalities. F) was considered as a superposition, and C) was not mentioned.
- However, F) is an elementary mode as well, although it produces two products. It does not use the triose phosphate transporter
- The systematic overview provided by elementary modes enables one to look for missing examples. Case C) is indeed realized in *Clusia minor* (Borland et al, 1994).
- Interestingly, (almost) pure elementary modes are realized here. No redundancy?

S. Schuster, D.A. Fell: Modelling and simulating metabolic networks. In: *Bioinformatics: From Genomes to Therapies* (T. Lengauer, ed.) Wiley-VCH, Weinheim, Vol. 2, 755-805.

Algorithms for computing elementary modes

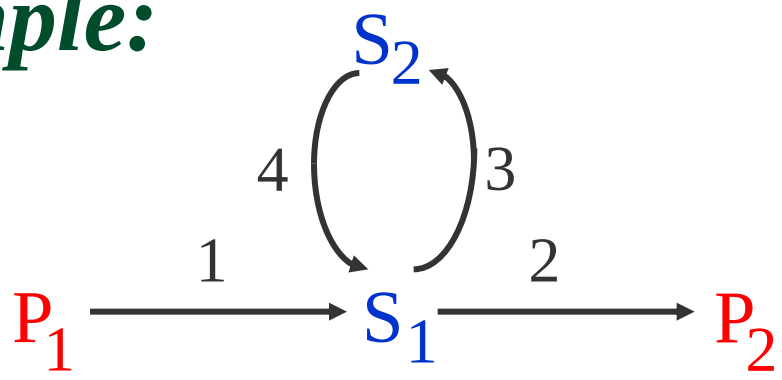
1. Modified Gauss-Jordan method starting with tableau ($\mathbf{N}^T$ $\mathbf{I}$). Pairwise combination of rows so that one column of $\mathbf{N}^T$ after the other becomes null vector.

S. Schuster et al., *Nature Biotechnol*. 18 (2000) 326-332.
J. Math. Biol. 45 (2002) 153-181.

2. Column operations on the null-space matrix. Empirically faster than 1. on biochemical networks.

C. Wagner, *J. Phys.Chem*. B 108 (2004) 2425–2431.
R. Urbanczik, C. Wagner, Bioinformatics. 21 (2005) 1203-1210.

Example:



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

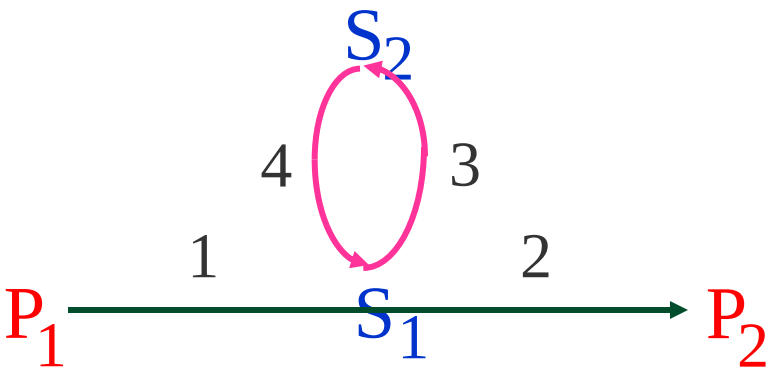
$$\mathbf{T}^{(0)} = \left(\begin{array}{cc:cccc} 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

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

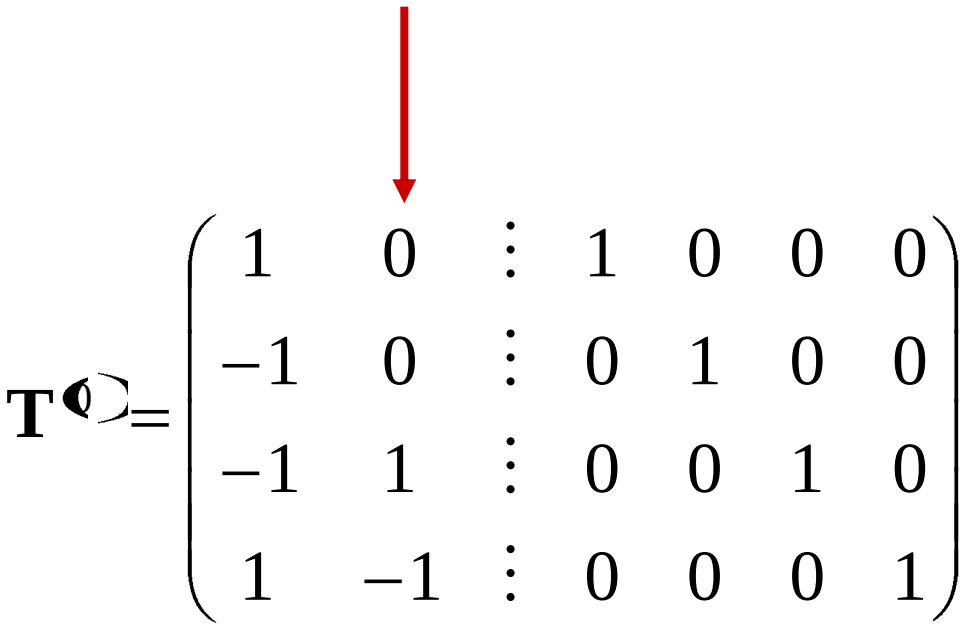
These two rows should not be combined

Final tableau:

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



Algorithm is faster, if this column is processed first.



Runtime complexity

- Not yet completely clear
- **V. Acuña, ..., M.-F. Sagot, L. Stougie: Modes and Cuts in Metabolic Networks: Complexity and Algorithms, *BioSystems*, 2009**
- **Theorem 9**. Given a matrix **N**, *counting* the number of elementary modes is ♯P-complete.
- **Theorem 10**. In case all reactions in a metabolic network are reversible, the elementary modes can be *enumerated* in polynomial time.
- **Open question**: Can elementary modes be *enumerated* in polynomial time if some reactions are irreversible?

Software involving routines for computing elementary modes

METATOOL - Th. Pfeiffer, F. Moldenhauer,
A. von Kamp (In versions 5.x, Wagner algorithm)
GEPASI - P. Mendes
JARNAC - H. Sauro
In-Silico-Discovery™ - K. Mauch
CellNetAnalyzer (in MATLAB) - S. Klamt
ScrumPy - M. Poolman
Alternative algorithm in MATLAB – C. Wagner, R. Urbanczik
PySCeS – B. Olivier et al.
YANAsquare (in JAVA) - T. Dandekar
EFMTool – M. Terzer, J. Stelling
On-line computation:
pHpMetatool - H. Höpfner, M. Lange

#P (sharp P) Complexity class

- An **NP** problem is often of the form, "Are there any solutions that satisfy certain constraints?" For example:
- Are there any subsets of a list of integers that add up to zero? (subset sum problem)
- Are there any Hamiltonian cycles in a given graph with cost less than 100?
- The corresponding **#P** problems ask "how many" rather than "are there any". For example:
- How many subsets of a list of integers add up to zero?
- How many Hamiltonian cycles in a given graph have cost less than 100?

Summary

- **Elementary modes are an appropriate concept to describe biochemical pathways in wild-type and mutants.**
- **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.**

Summary (2)

- **Pathway analysis is well-suited for computing maximal and submaximal molar yields**
- **Many metabolic systems in various organisms have been analysed in this way. In some cases new pathways discovered**
- **Relevant applications: knockout studies (biotechnology) and enzyme deficiencies (medicine)**
- **Work still to be done on decomposition methods (combinatorial explosion)**

Acknowledgements

- David Fell (Brookes U Oxford)
- Thomas Dandekar (U Würzburg)
- Steffen Klamt (MPI Magdeburg)
- Jörg Stelling (ETH Zürich)
- Thomas Pfeiffer (Harvard)
- late Reinhart Heinrich (HU Berlin)
- Ina Koch (TFH and MPI Berlin)
- and many others

•Acknowledgement to DFG, BMBF and FCT (Portugal) for financial support