Fast Enumeration of Smallest Engineering Strategies in Genome-Scale Metabolic Networks

Steffen Klamt

Research Group „Analysis and Redesign of Biological Networks“
Max Planck Institute Magdeburg

Outline

1) Introduction: Stoichiometric and Constraint-Based Analysis of Biochemical Reaction Networks

→ Focus: Elementary Modes and Minimal Cut Sets

---

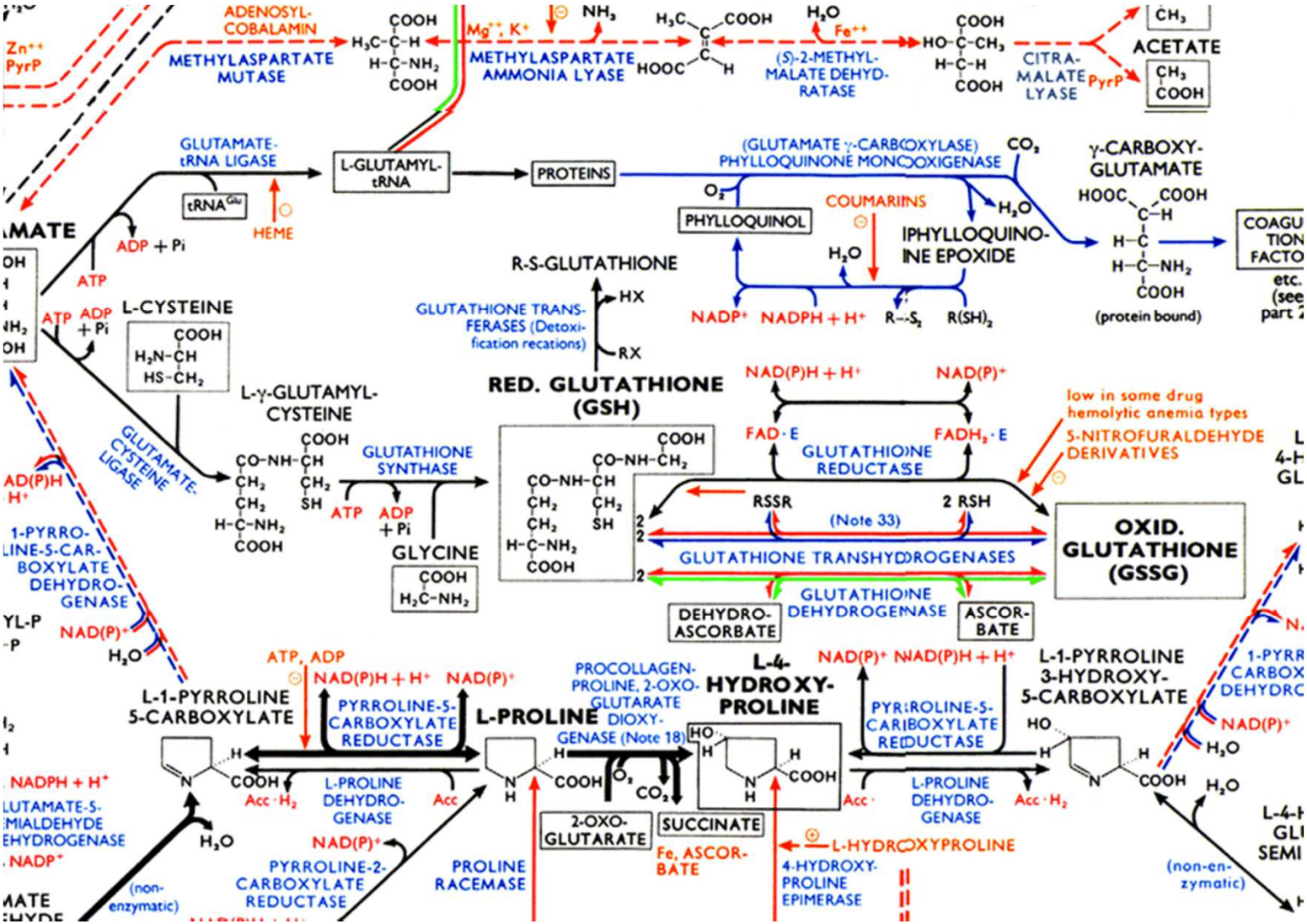
2) (Re)Designing Metabolic Networks by Minimal Cut Sets

3) Duality Between Modes and Cuts and Algorithms for Their Computation

Methods will be presented with focus on metabolic network analysis.

In principle, they are applicable to any network of chemical reactions or material flows.

Metabolic Reaction Networks



Metabolic Network Modeling

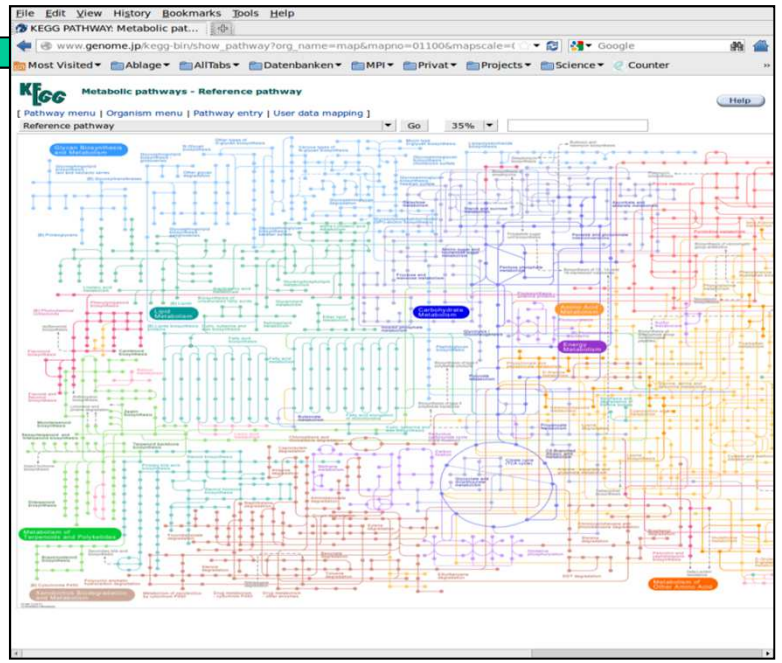
- metabolites
- reactions
  - stoichiometry
  - reaction directionality / reversibility
  - involved enzymes (+ encoding genes)

Stoichiometric Model

Metabolic Network Analysis

- detection of structural inconsistencies
- properties of global network architecture
- identification of meaningful pathways and cycles
- network functions and capabilities
- maximal/extremal/optimal performance
- prediction of knock-out effects
- computer-aided (re)design

Organism-specific metabolic networks
(databases: KEGG, BioCyc, …)



- Biochemical knowledge
- Genome-sequencing

Stoichiometric Models of Metabolic Networks

Meanwhile, >100 genome-scale reconstructions of organism-specific metabolic networks have been published.

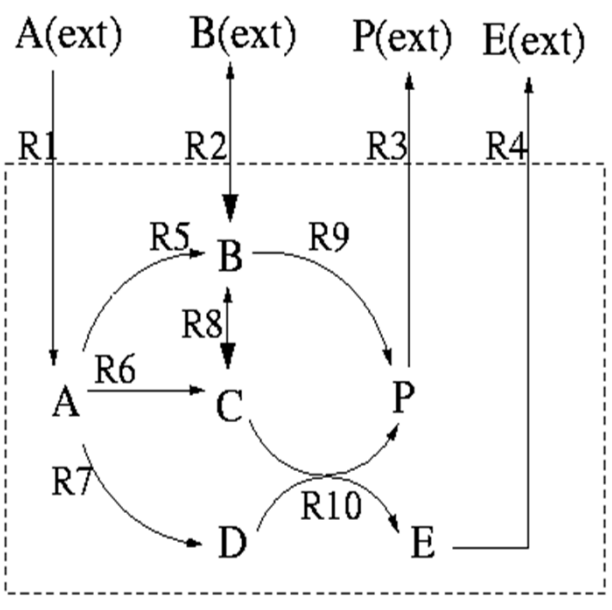
They typically contain 600 - 3000 or more reactions and metabolites.

For certain aspects, it is often sufficient to focus on the core of the central metabolism with 100-150 reactions.

Stoichiometric Matrix N

m x *q* stoichiometric matrix **N** **:** captures the stoichiometric coefficients

– *m* (internal) metabolites (external sinks, sources, ... often not considered in **N**)
– *q* reactions (may include pseudo reactions, e.g., transport or biomass synthesis)
– *Rev*: indices of reversible reactions; *Irrev* = indices of irreversible reactions



$$\mathbf{N} = \begin{array}{c} \begin{matrix} R1 & R2 & R3 & R4 & R5 & R6 & R7 & R8 & R9 & R10 \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array} \begin{matrix} \\ \leftarrow A \\ \leftarrow B \\ \leftarrow C \\ \leftarrow D \\ \leftarrow E \\ \leftarrow P \end{matrix}$$

Rev={R2,R8}, *Irrev*={R1,R3,R4,R5,R6,R7,R9,R10}

(store addtionally the enzymes/genes for each reaction)

Graphs vs. Hypergraphs

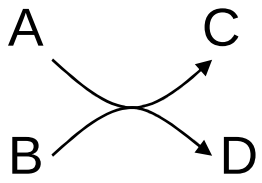
Network = Graph?!

(Metabolic) Reaction networks span a hypergraph and N is the incidence matrix of this hypergraph.

Graph: A → B **Edges** can only connect pairs of nodes.

A (bimolecular) reaction A + B → C + D cannot be represented in a graph.

Hypergraph: **Hyperedges** may connect two **sets** of nodes.



hyperedge ***h***={ {A,B}, {C,D} }

Hypergraphs can be converted to graphs where some important network properties can be studied. However, graph representations are limited for ***functional*** network analysis.

More on hypergraphs in biological network analysis:
Klamt S, Haus U, Theis F: Hypergraphs and Cellular Networks. 2009. PLoS Computational Biology 5:e1000385.

Stoichiometric Matrix N

$\mathbf{N}$ is fundamental for dynamic modeling of (metabolic) reaction networks (ODEs):

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N} \cdot \mathbf{r}(t)$$

$\mathbf{c}(t)$: metabolite concentrations at time t [mmol/gDW]

$\mathbf{r}(t)$: (net) reaction rates [mmol/(gDW · h)]

$\mathbf{r}(t) = f(\mathbf{c}(t), \mathbf{p})$ ⟶ detailed knowledge of reaction mechanism, kinetic parameters and regulation required (often not available)

Methods for stoichiometric / structural / topological network analysis rely on N and do / need not consider reaction kinetics.

Quasi-Steady-State Assumption

- high turnover of metabolites in central metabolism
  (typical turnover time: 0.1-60 sec)

- reasonable assumption: metabolism operates close to steady-state,
  in particular in micro-organisms (e.g., bacteria, yeast,…)

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{N} \cdot \mathbf{r}(t) \approx \mathbf{0}$$

$\mathbf{N} \cdot \mathbf{r} = \mathbf{0}$ — *Metabolite Balancing Equation*
(Production = Consumption
for each internal metabolite)

$\mathbf{r}_{Irrev} \geq \mathbf{0}$ — *Irreversibility Constraints*

→ equalities and inequalities (constraints)
→ „constraint-based modeling of metabolic networks“

Space of feasible steady state flux distributions

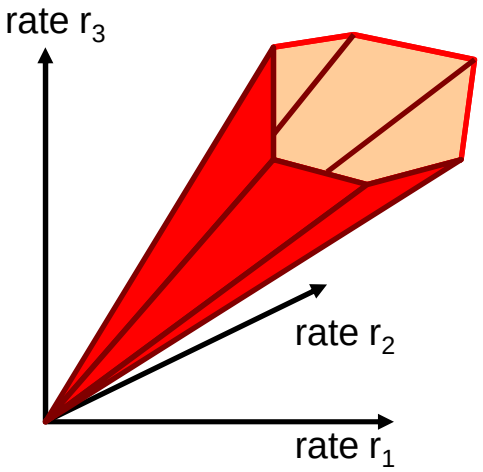
C1) *Quasi-steady-state*: **Nr = 0**

Solution space of flux vectors **r**:

null space (or kernel) of **N:** *ker*(**N**)

C1) *Quasi-steady-state*: **Nr = 0**
C2) *Irreversibility*: $\mathbf{r}_{Irrev} \geq \mathbf{0}$

Convex polyhedral cone (flux cone)



Constraint-based Methods:

- Metabolic Flux Analysis
  (particular flux distribution)
- Flux Balance Analysis
  (optimal flux distribution)
- Elementary Modes
- Minimal Cut Sets

Elementary (Flux) Modes

An **Elementary (flux) mode (EM):** is a flux vector **e** fulfilling

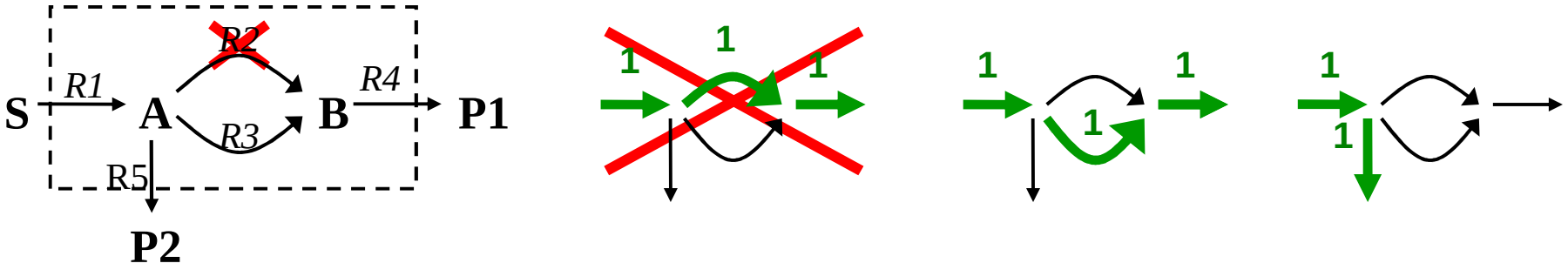
C1) *Steady State*: **Ne = 0** (*Schuster et al.*)

C2) *Irreversibility*: $\mathbf{e}_{Irrev} \geq \mathbf{0}$

C3) *Elementarity*: no vector $\mathbf{v} \neq \mathbf{0}$ satisfies C1 and C2 and supp(**v**) $\subset$ supp(**e**) where supp(**v**)={i | $v_i \neq 0$}

Elementary modes:

- support-minimal sets of reactions able to perform function in steady state
- correspond to pathways or cycles
- uniquely identified by its participating reactions (rate ≠0)



- deletion of a reaction: new set of EMs can easily be derived.

Elementary (Flux) Modes

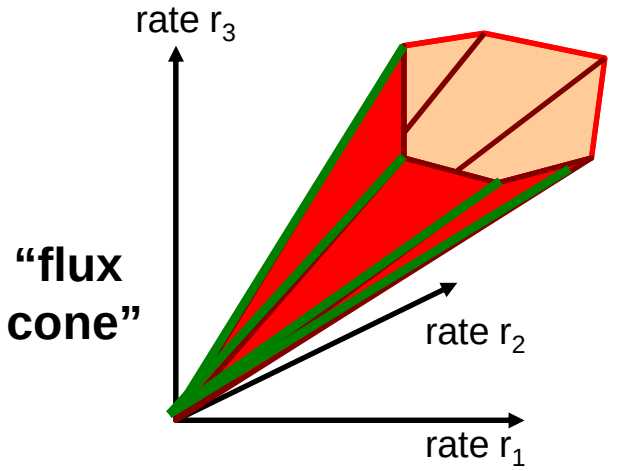
- elementary modes generate the flux cone (by non-negative linear combinations)

$$\mathbf{r} = \sum \gamma_i \mathbf{e}^i \quad ; \quad \gamma_i \geq 0$$

- each extreme ray of the flux cone is an EM (not all EMs are extreme rays)

- computation of EMs relies on methods for computing extreme rays of polyhedral cones

- combinatorial complexity: even in medium-scale networks millions of EMs may exist
  → applicable to medium-scale networks only (<150 reactions)



Elementary Modes: A Versatile Tool for Metabolic Pathway and Network Analysis

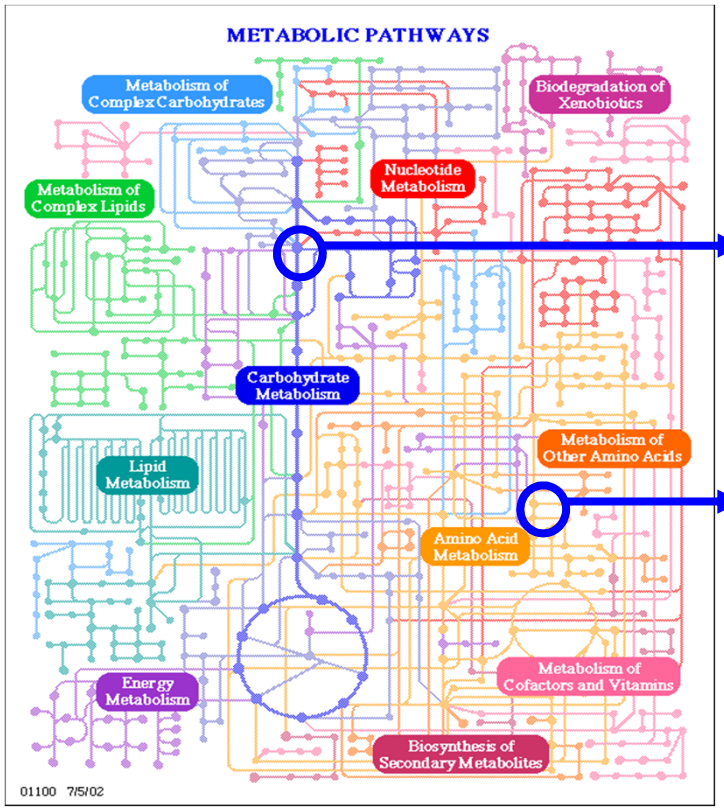
MAX-PLANCK-GESELLSCHAFT

Biochemical Network

Functionality of interest

Minimal Functional Units (Elementary Modes)

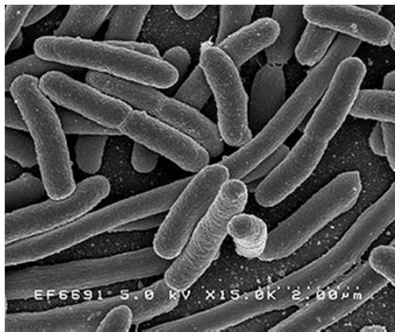
- pathways and routes
- flexibility to perform a task
- optimal performance
- functional importance of reactions
- phenotype predictions
- structural couplings



substrate (e.g. glucose)

biosynthetic product (e.g. tryptophan)

Central metabolism of *E.coli* : Elementary-modes analysis



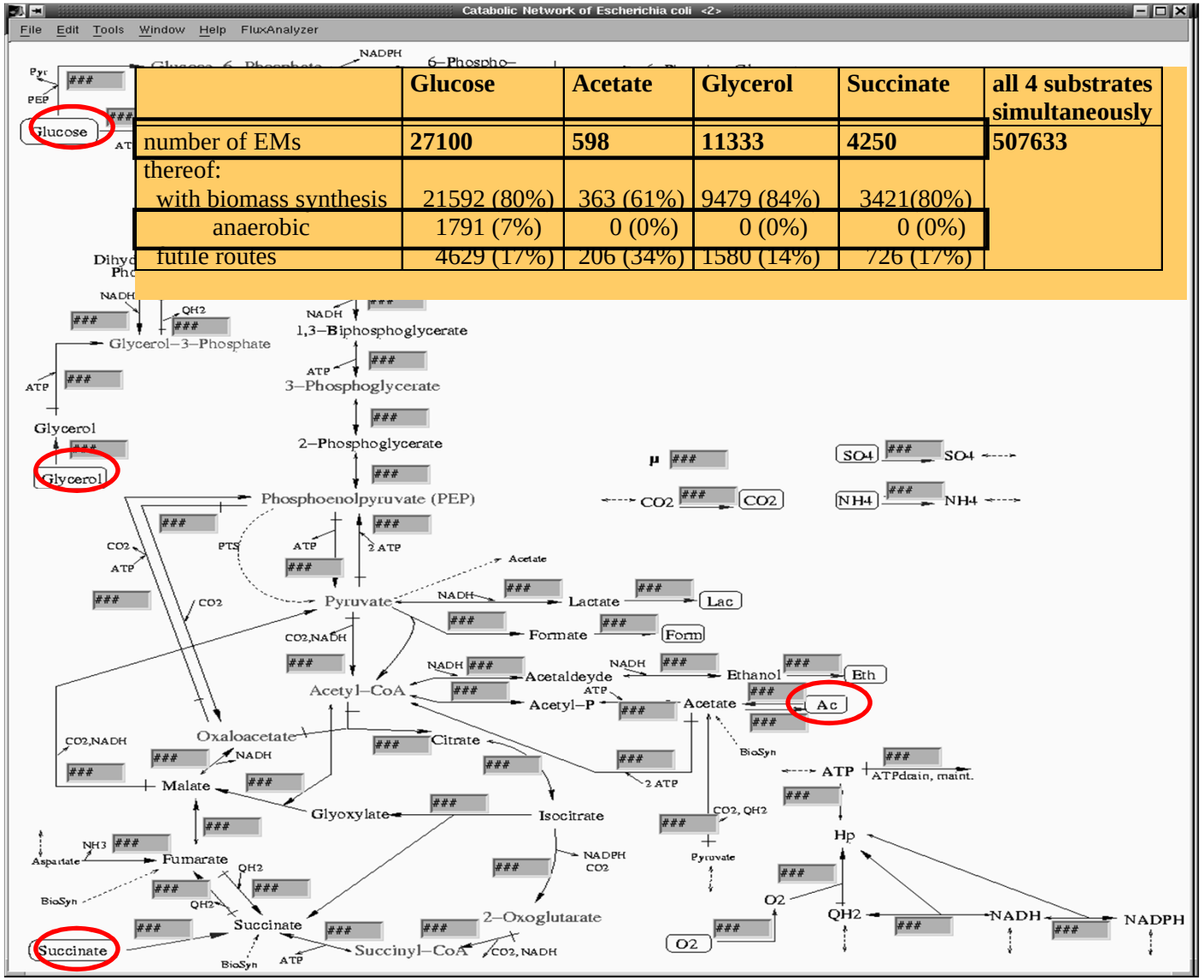
110 reactions
89 metabolites

(Central metabolism in detail; anabolic part compressed)

Substrates:

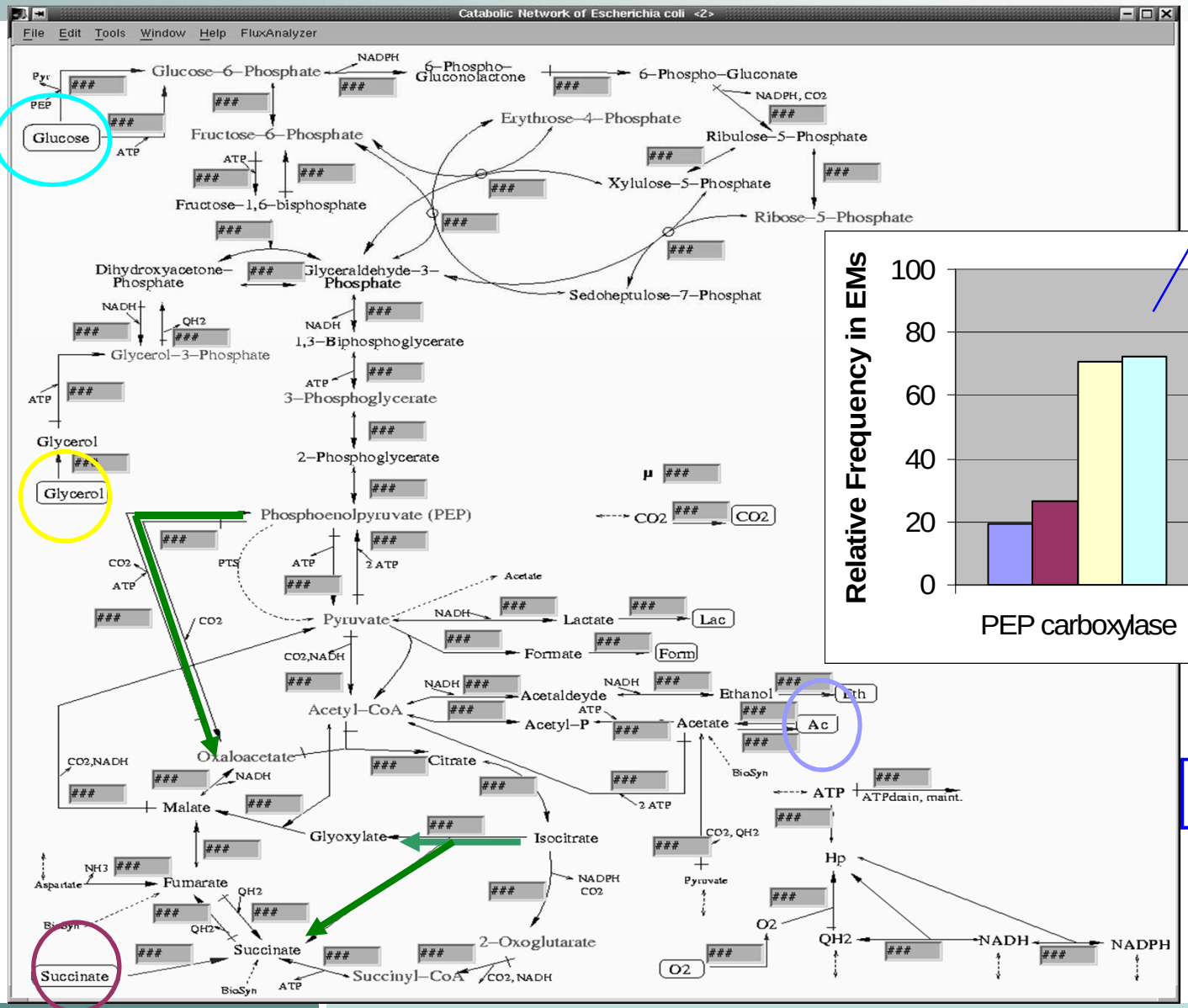
Glucose
Succinate
Glycerol
Acetate

Stelling, Klamt et al. ; Nature 420:190-193



	Glucose	Acetate	Glycerol	Succinate	all 4 substrates simultaneously
number of EMs	**27100**	**598**	**11333**	**4250**	**507633**
thereof:					
with biomass synthesis	21592 (80%)	363 (61%)	9479 (84%)	3421(80%)	
anaerobic	1791 (7%)	0 (0%)	0 (0%)	0 (0%)	
futile routes	4629 (17%)	206 (34%)	1580 (14%)	726 (17%)	

Relative Importance of Single Reactions



Correlate with gene expression

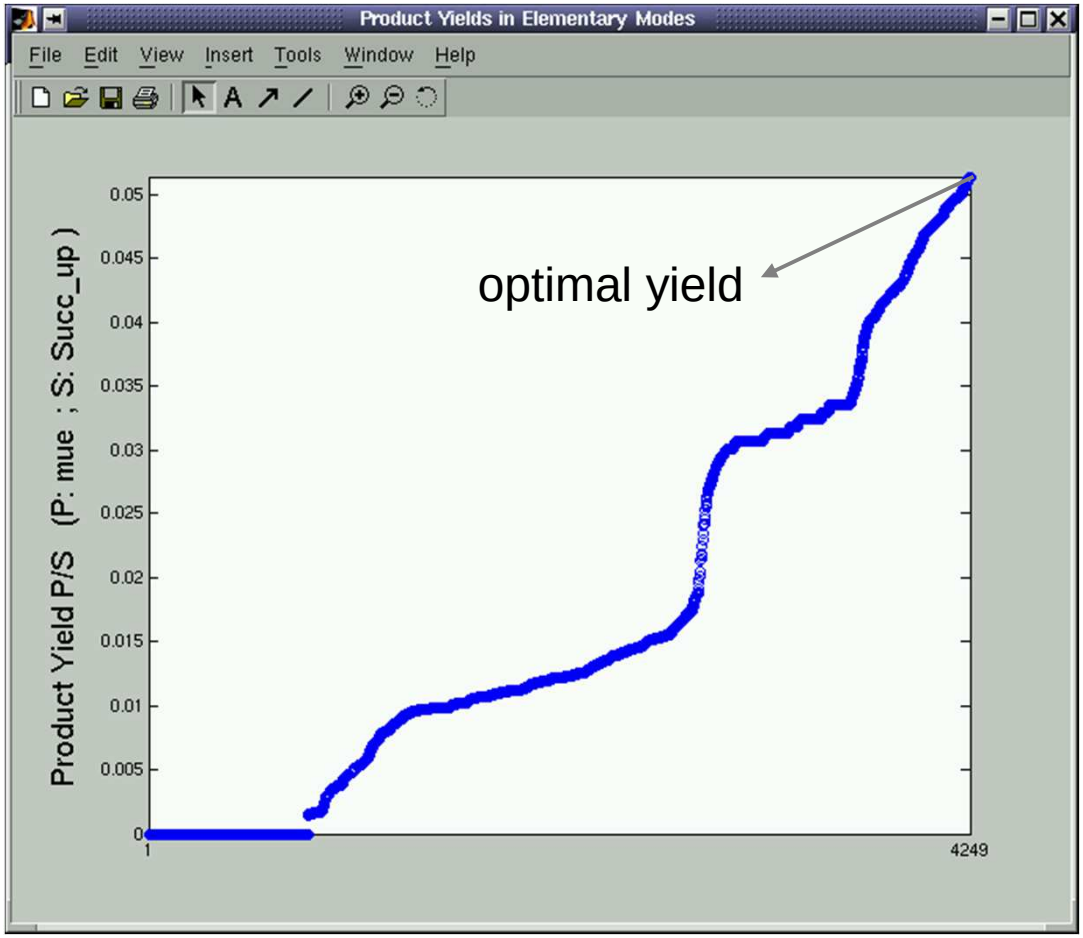
essential

→ phenotype predictions

Product yields and optimal EMs

Optimal yield can be identified AND all the modes by which the optimum is achieved

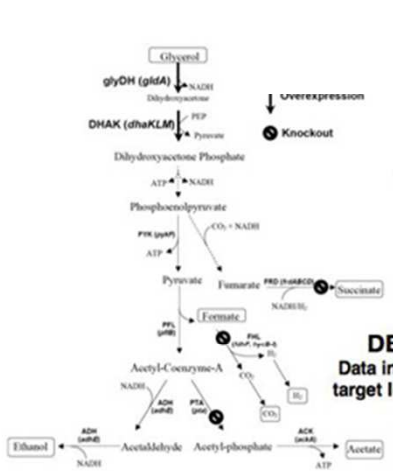
Biomass Yield
gDW biomass per mmol substrate

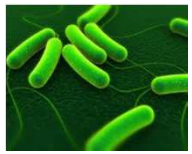


Example: biomass yield of *E.coli* on (substrate) succinate

Metabolic Engineering / Biotechnology

Metabolic Engineering: experimental + theoretical techniques for the targeted improvement of metabolic pathways in microorganisms for the production of fuels and chemicals from renewable resources.



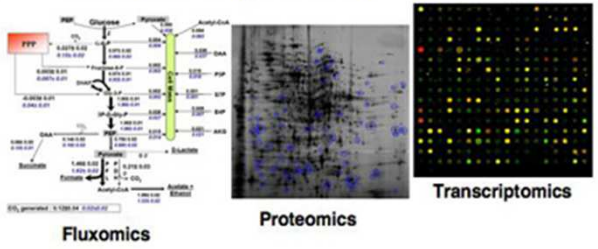


STRAIN CONSTRUCTION
Gene(s) "knockins" & "knockouts"
Mutagenesis & Selection

STRAIN CHARACTERIZATION
Tubes, Flask, Bioreactors

DESIGN
Data integration &
target Identification

SYSTEMS ANALYSIS
Functional genomics tools



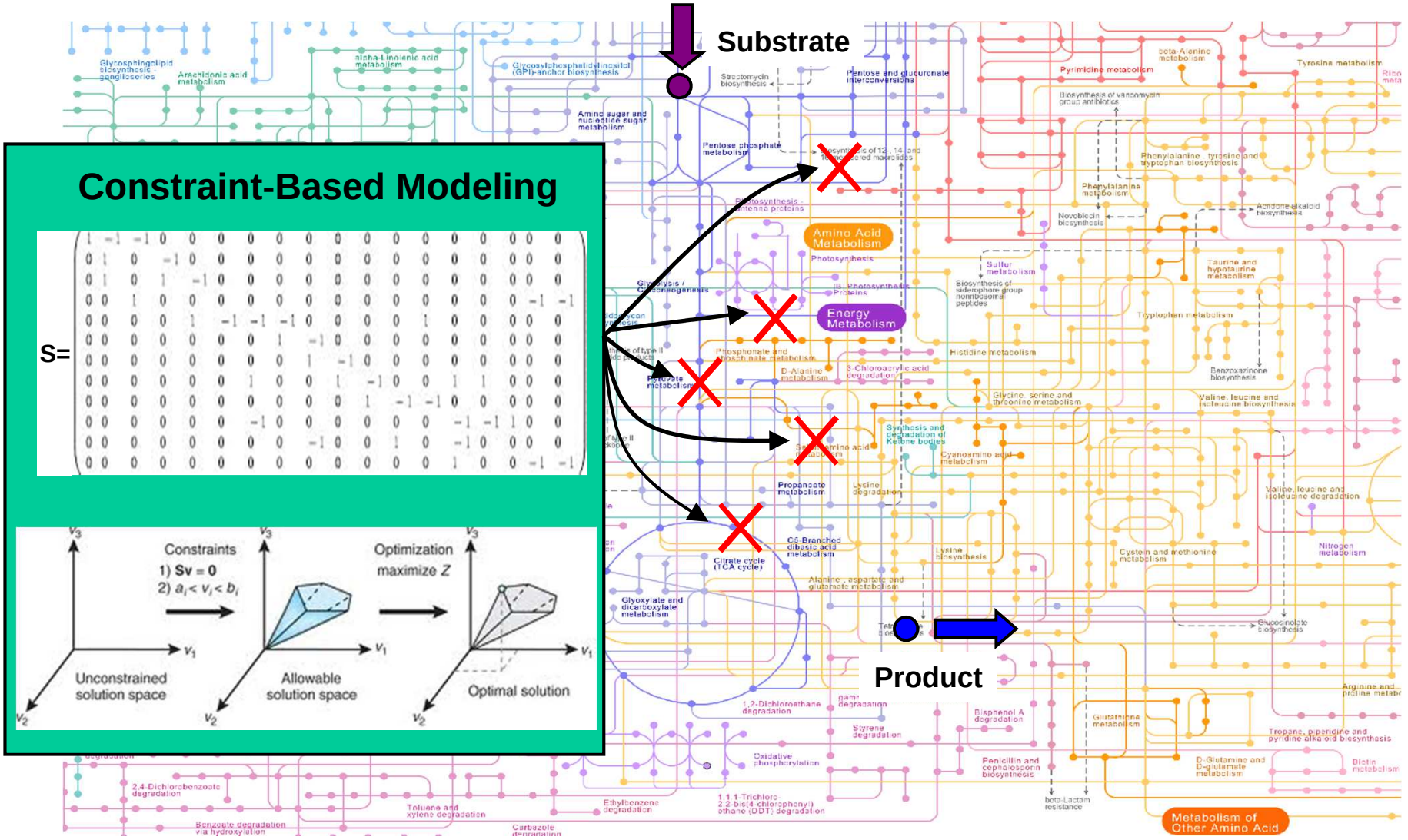
Rational Strain Design Using Computational Methods

Metrics: Concentration (g/L), Yield (g/g), Rate (g/L/h and g/gDW/h)

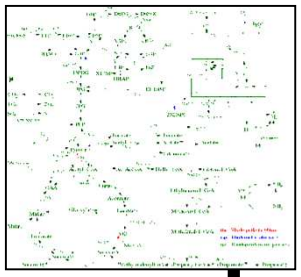
Large-scale production process

Computational Strain Design
(Metabolic Engineering / Biotechnology)

MAX-PLANCK-GESELLSCHAFT



Metabolic Engineering: (Re-)Design of Cellular Factories



Stochiometric network model

Intervention goal:
bio-synthesis of a certain chemical from a given substrate

Via projection of EMs

Computational Strain Design

Model-based search for metabolic intervention strategies optimizing product yield

Maximal product yield

Desired region of operation

Product yield

Metabolic Engineering *(targeted knock-out or/and overexpression of metabolic genes)*

Biomass yield

Maximial biomass yield (objective of the cell!)

Constraint-Based Methods for Computational Strain Design

Bi-level optimization problems

- knockout strategies enforcing coupled biomass and product synthesis
- OptKnock and variants OptReg, OptForce (Maranas et al.); OptORF (Reed at al.); EMILiO (Mahadevan et al.); and others

→ MILP problems solved at genome-scale
→ difficult: #knockouts >10 ; #potential targets >300;
→ #alternative strategies found <100

Elementary-modes-based optimization

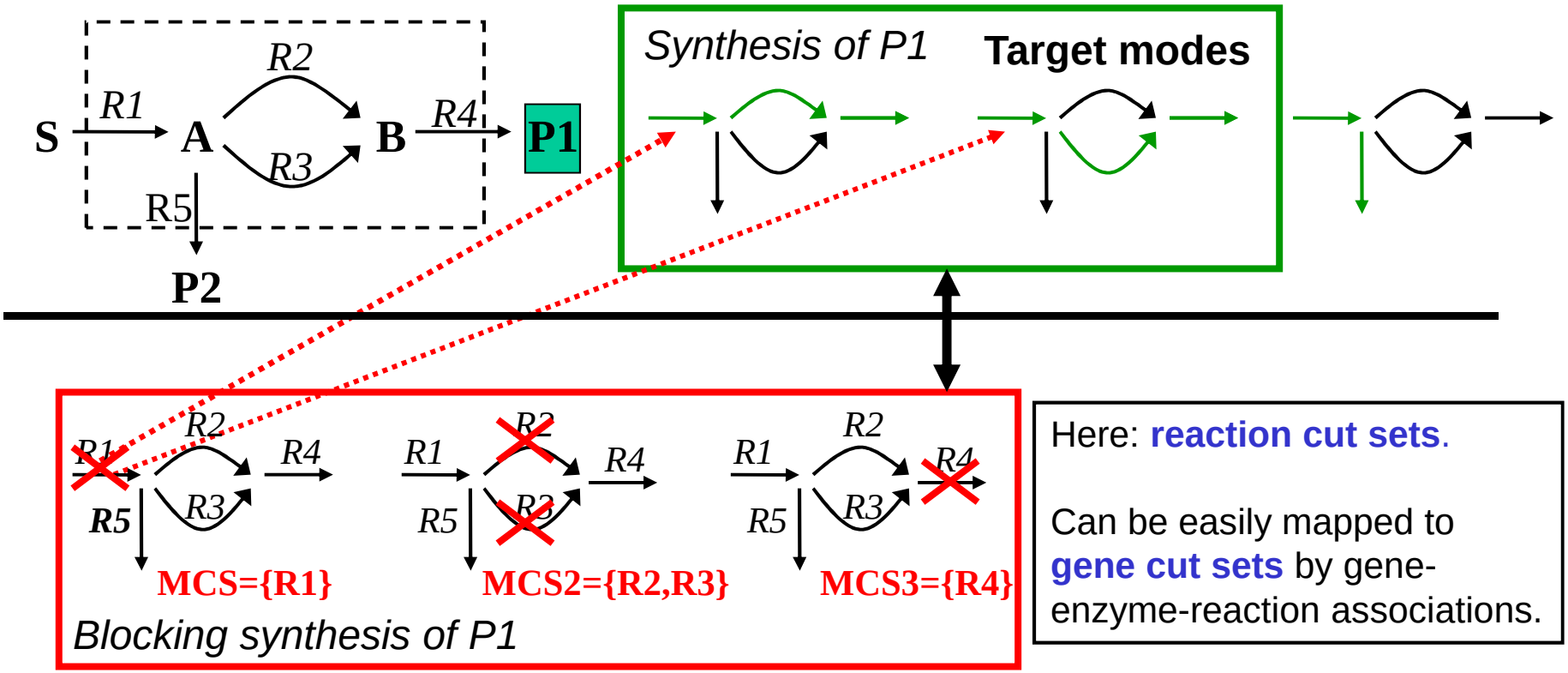
• Minimal metabolic functionality (Trinh, Srienc et al.); FluxDesign (Wittmann et al.); CASOP (Hädicke et al.); and others

→ not applicable to genome-scale networks

(Constrained) Minimal Cut Sets: systematic enumeration of all minimal knockout sets that block *undesired* while keeping *desired* metabolic functions.
• enumerative approach
• calculation based on elementary modes
→ so far: not applicable to genome-scale networks

From Elementary Modes to Minimal Cut Sets

Elementary Modes: minimal functional units (pathways) that can operate in steady state



Here: **reaction cut sets**.

Can be easily mapped to **gene cut sets** by gene-enzyme-reaction associations.

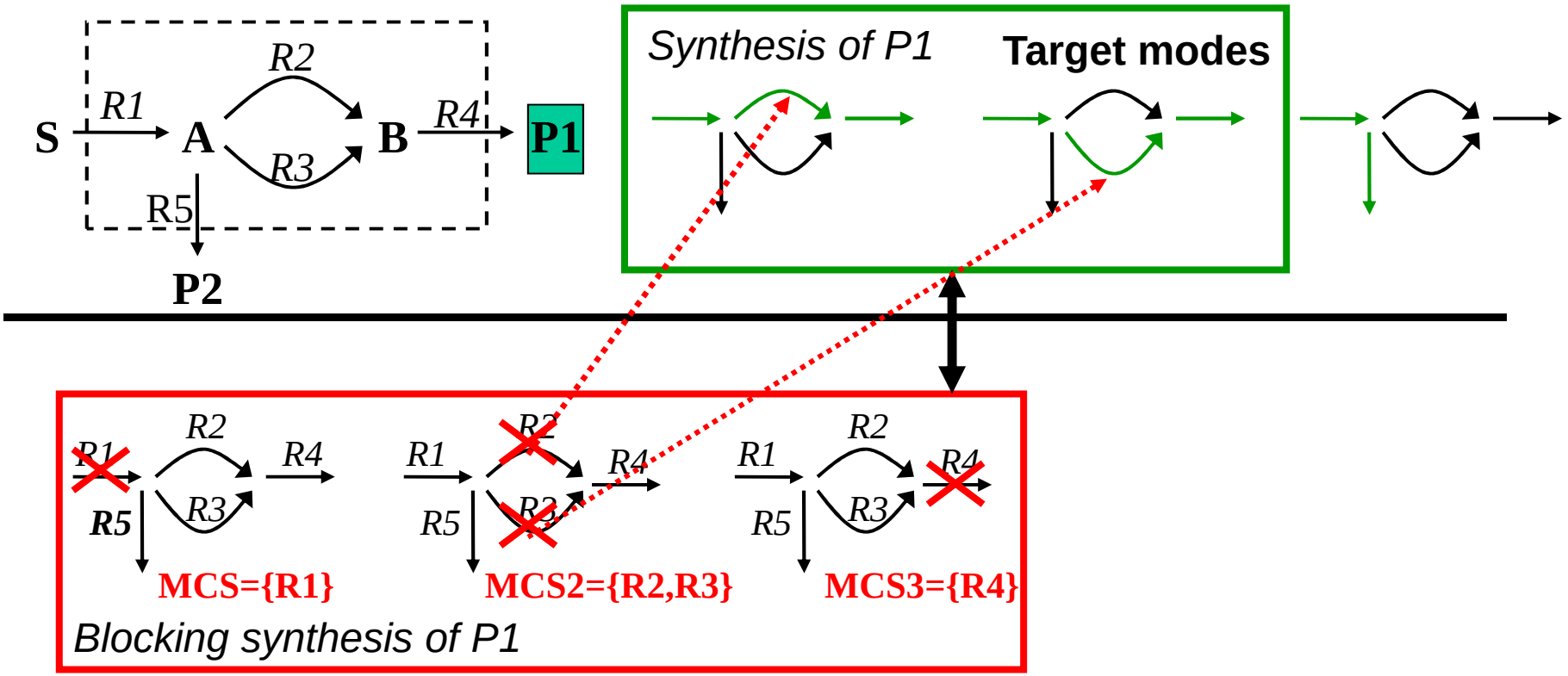
Minimal Cut Set (MCS)*:* minimal set of cuts (knockouts) blocking certain functions in steady state

***MCSs** are the minimal hitting sets of the corresponding EMs*

Klamt S: BioSystems 2006, 83:233-237.

From Elementary Modes to Minimal Cut Sets

Elementary Modes: minimal functional units (pathways) that can operate in steady state



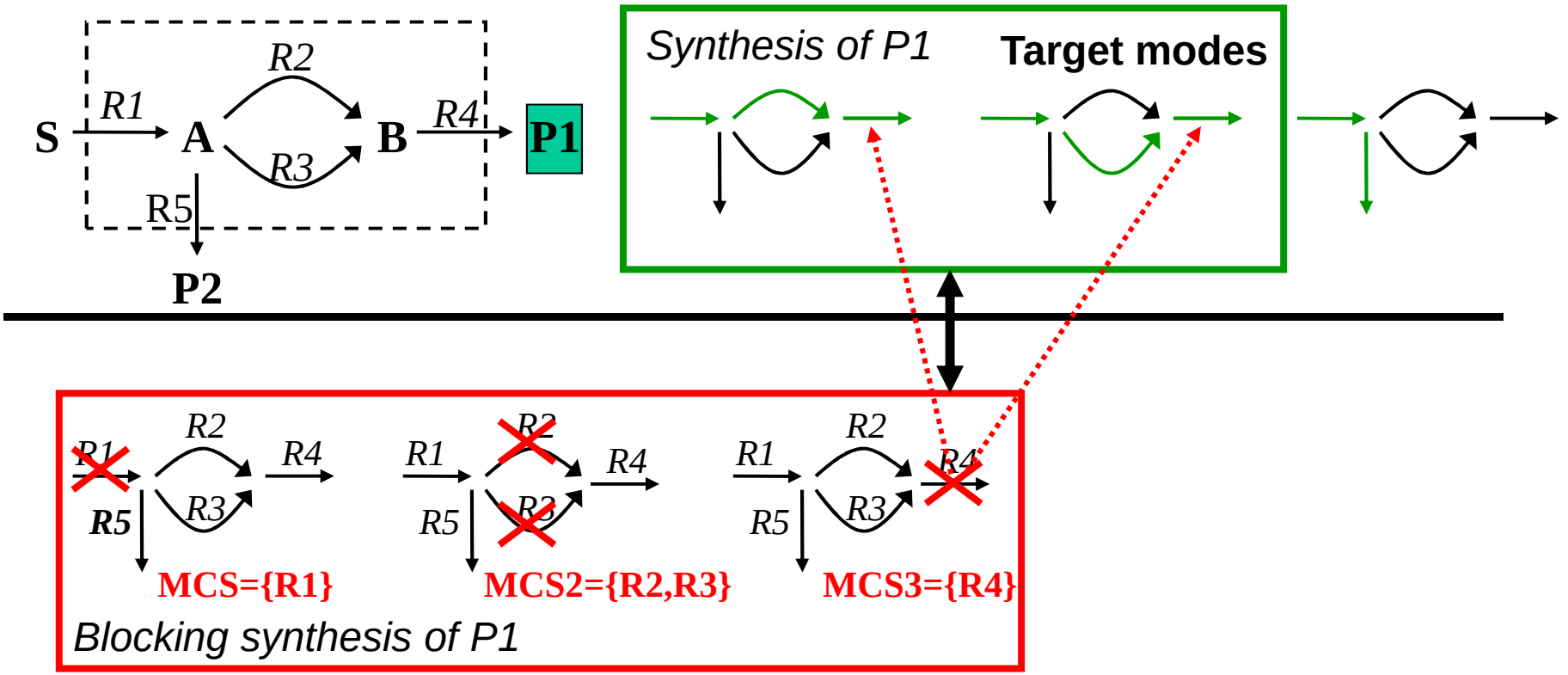
Minimal Cut Set (MCS): minimal set of cuts (knockouts) blocking certain functions in steady state

***MCSs** are the minimal hitting sets of the corresponding EMs*

Klamt S: BioSystems 2006, 83:233-237.

From Elementary Modes to Minimal Cut Sets

Elementary Modes (EM)*:* minimal functional units that can operate in steady state



Minimal Cut Set (MCS)*:* minimal set of cuts (knockouts) blocking certain functions in steady state

***MCSs** are the minimal hitting sets of the corresponding EMs*

Klamt S: BioSystems 2006, 83:233-237.

Elementary Modes, Minimal Cut Sets and Minimal Hitting Sets

Computation of Minimal Hitting Sets (here: for a given set of Target Modes):

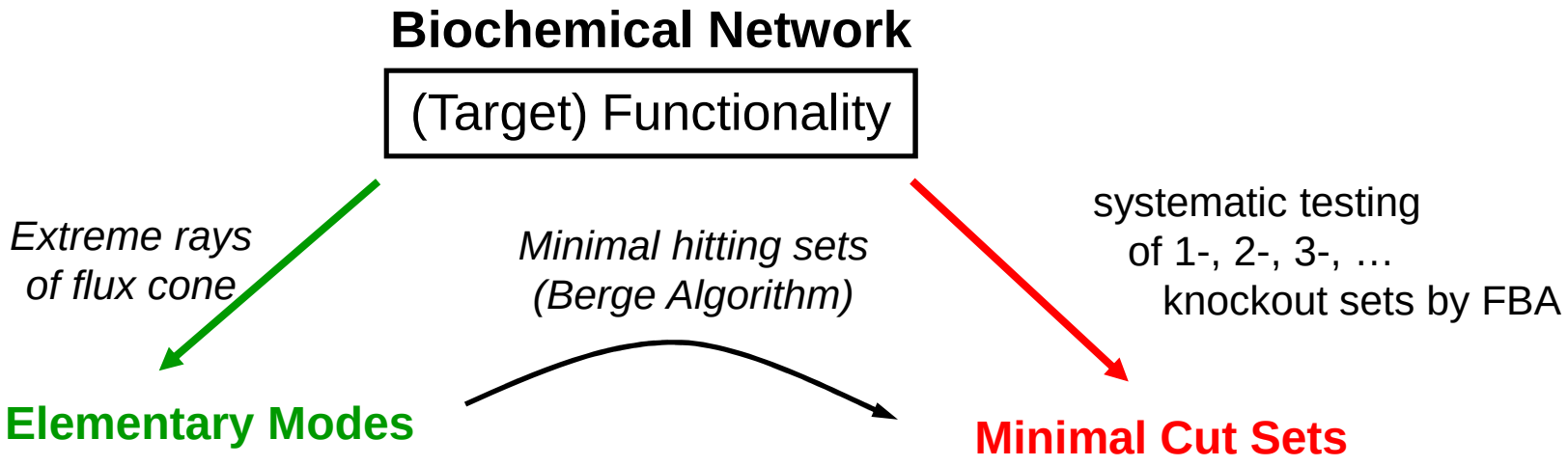
Many algorithms are known from hypergraph theory.

Algorithm of Claude Berge [1] performs well for metabolic networks [2].

[1] Claude Berge. (1989): Hypergraphs. Combinatorics of finite sets.(North-Holland, Amsterdam)

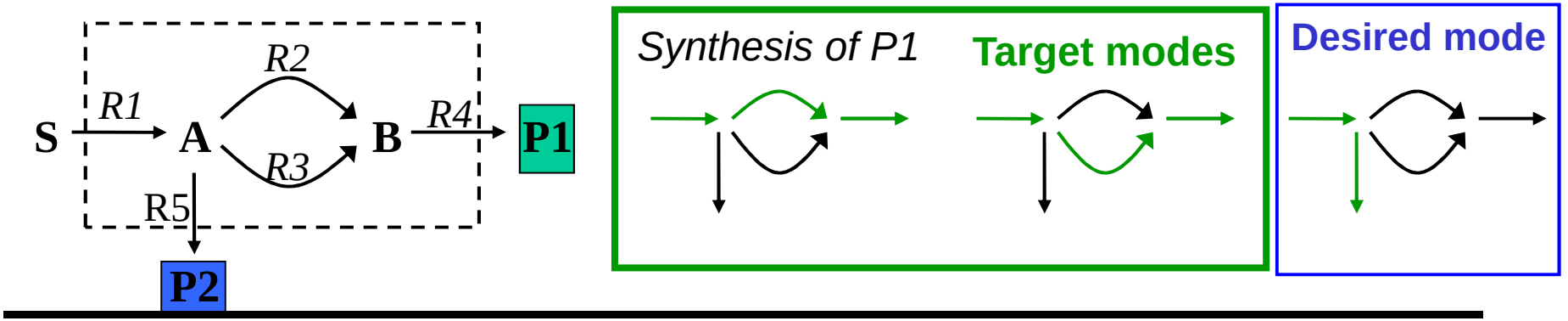
[2] Haus U, Klamt S, Stephen T. (2008) Computing knockout strategies in metabolic networks. J Comp Biol 15: 259-268.

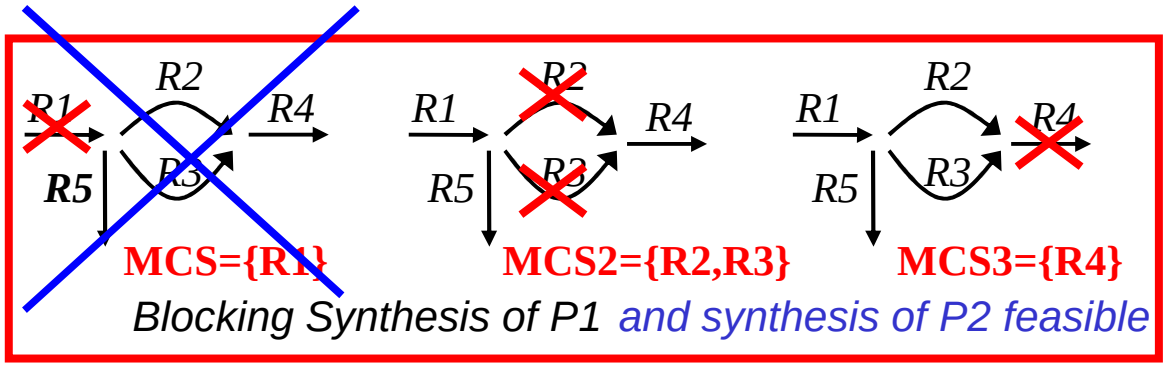
Elementary Modes, Minimal Cut Sets and Minimal Hitting Sets



Generalization:
Constrained Minimal Cut Sets

Elementary Modes*:* minimal functional units (pathways) that can operate in steady state





Constrained MCSs:
- specify „desired modes“ (must not be hit by MCSs)

Constrained Minimal Cut Set (MCS)*:* minimal cut sets blocking undesired while keeping desired metabolic behaviors.

Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

Constrained MCSs: A Tool for Computer-aided Redesign of Metabolic Networks

Constrained Minimal Cut Set (cMCSs) problem for reaction networks:

- set $\mathcal{T}$ of target EMs (to be blocked)
- set $\mathcal{D}$ of desired EMs
- n: minimal number of desired EMs in $\mathcal{D}$ that must not be hit by the MCSs

- algorithm for minimal hitting set calculation adapted for cMCSs:
  a) check on-the-fly whether MCS candidates keep desired modes (often faster)
  b) identify constrained MCSs from unconstrained MCSs during post-processing

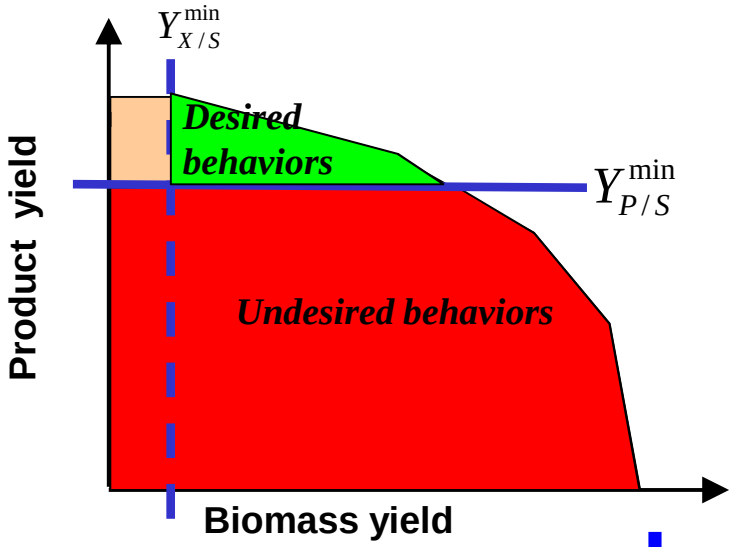
- large variety of complex intervention problems can be conveniently formulated and solved by cMCSs

- exhaustive enumeration of all knockout strategies
  (... if the network is not too large)

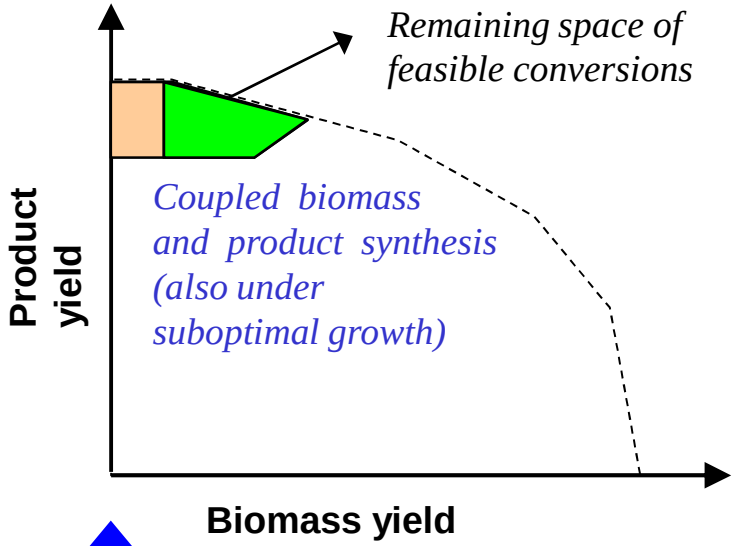
Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

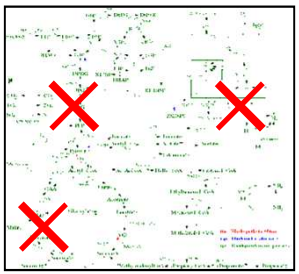
Using Constrained Minimal Cut Sets for Strain Design

(1) Compute and analyze space of feasible conversions and specify desired and undesired behaviors



(3) Example of the remaining solution space for one specific constrained MCS





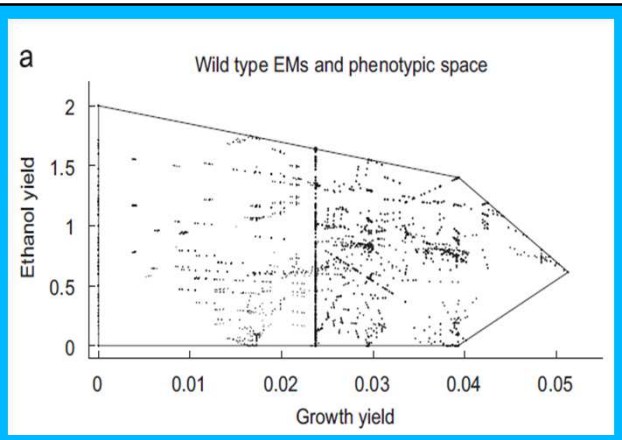
(2) Compute all constrained MCSs, that eliminate undesired flux distributions and keep at least some of the desired behaviors

Example: Coupled Biomass and Ethanol Synthesis in *E.coli*

Goal: search for interventions that lead to high (anaerobic) ethanol synthesis on glucose while still enabling some formation of biomass

1) Compute elementary modes and inspect phenotypic space

(each dot = one mode)

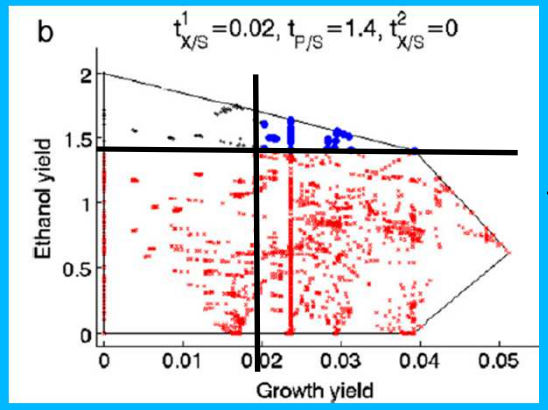


2) Specify target and desired modes

Red dots = target modes

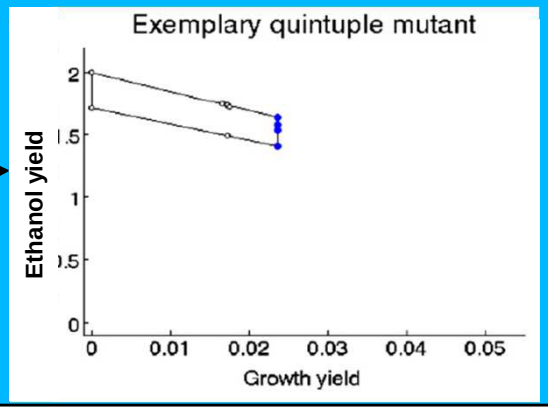
Blue dots = desired modes

Black dots = „don‘t care“



with n=1

3) Compute constrained MCSs and implement the most practical



Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

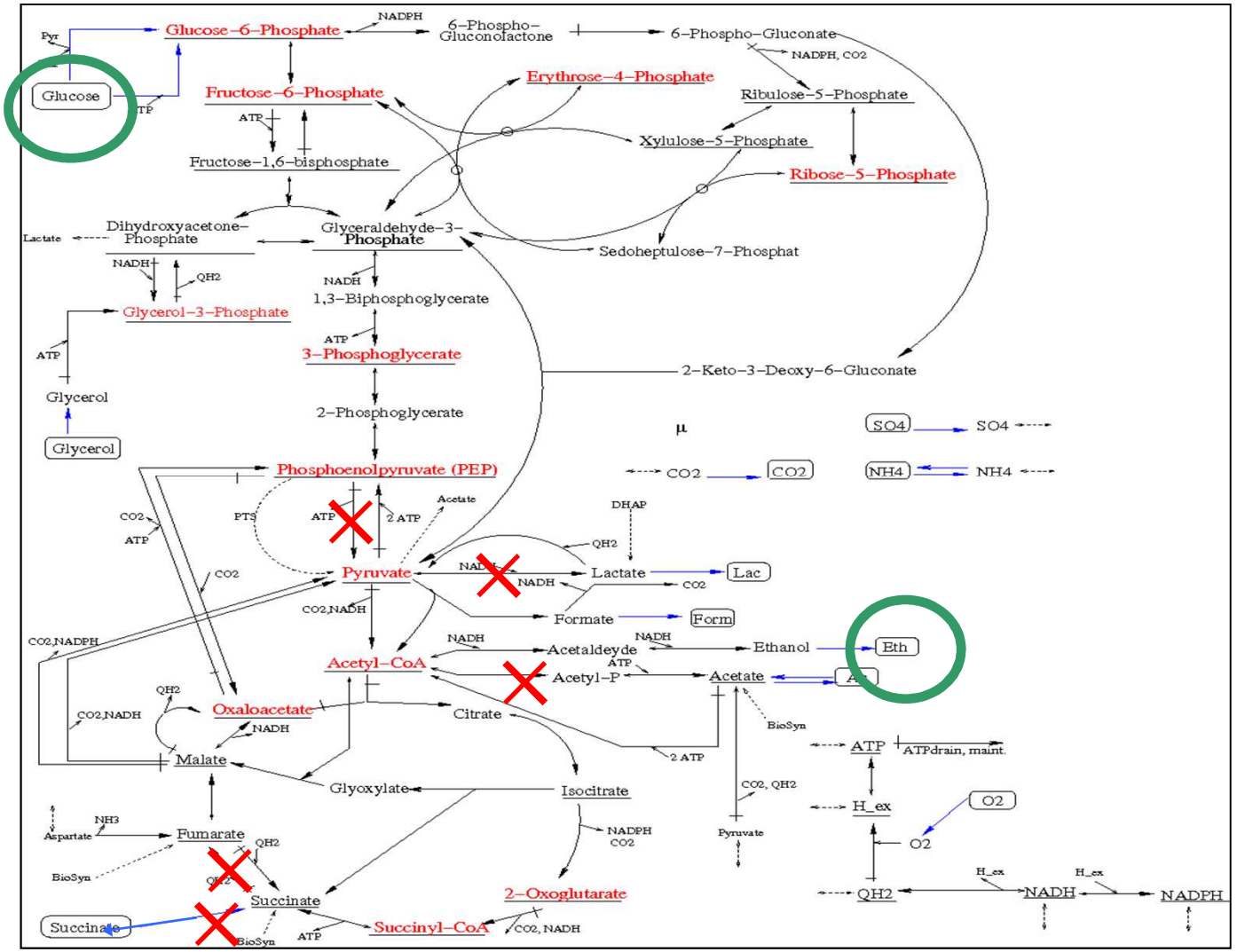
Example: Coupled Biomass and Ethanol Synthesis in *E.coli*

Intervention goal:

<u>Ethanol yield</u>
> 1.4 mol/(mol glucose)

<u>Biomass yield:</u>
> 20gDW/(mol glucose)

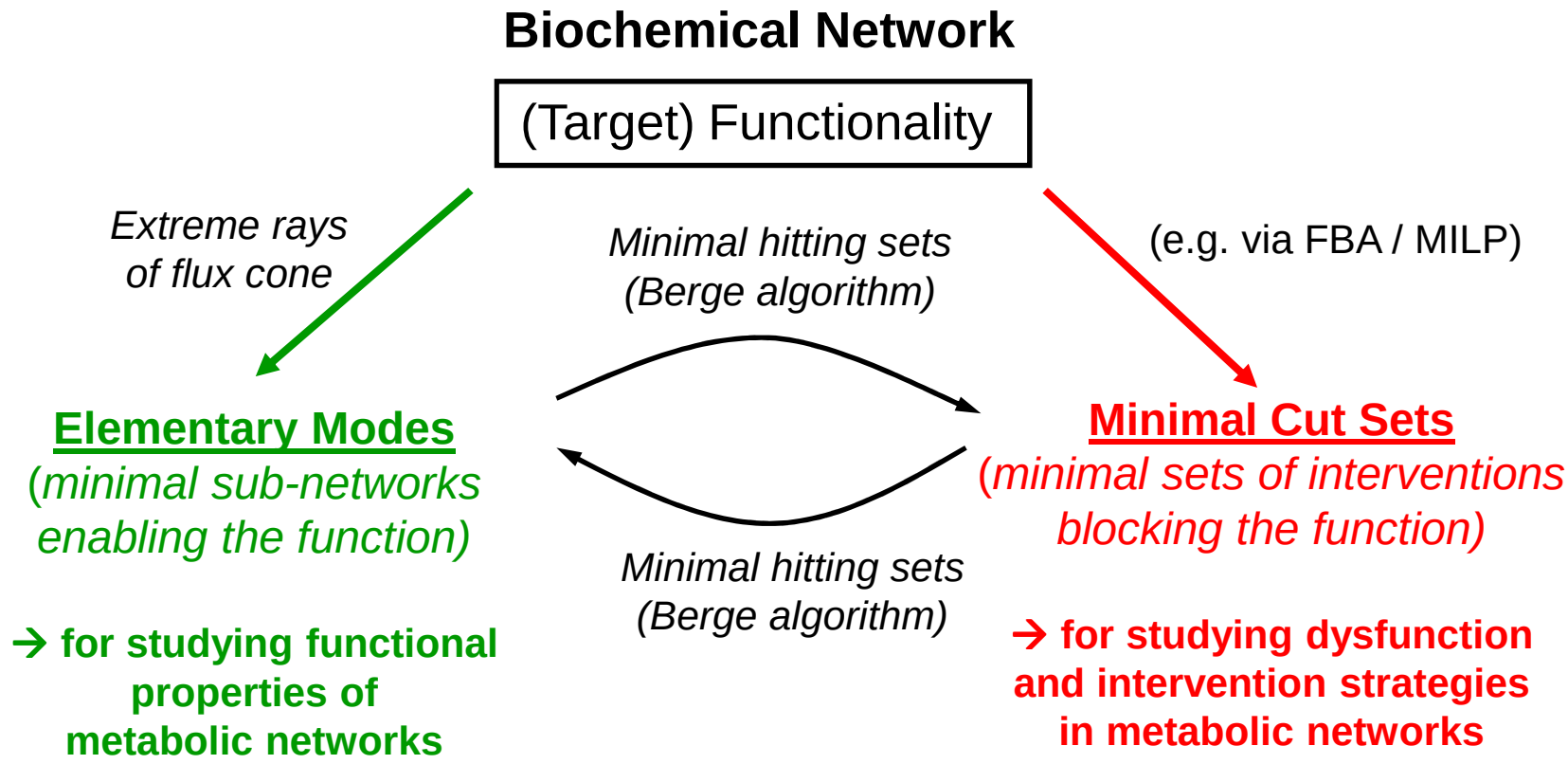
→ at least 5 knockouts
→ ca. 2000 cMCS



Hädicke and Klamt (2011): Computing Complex Metabolic Intervention Stratgeies using Constrained Minimal Cut Sets. Metabolic Engineering, 13:204-213.

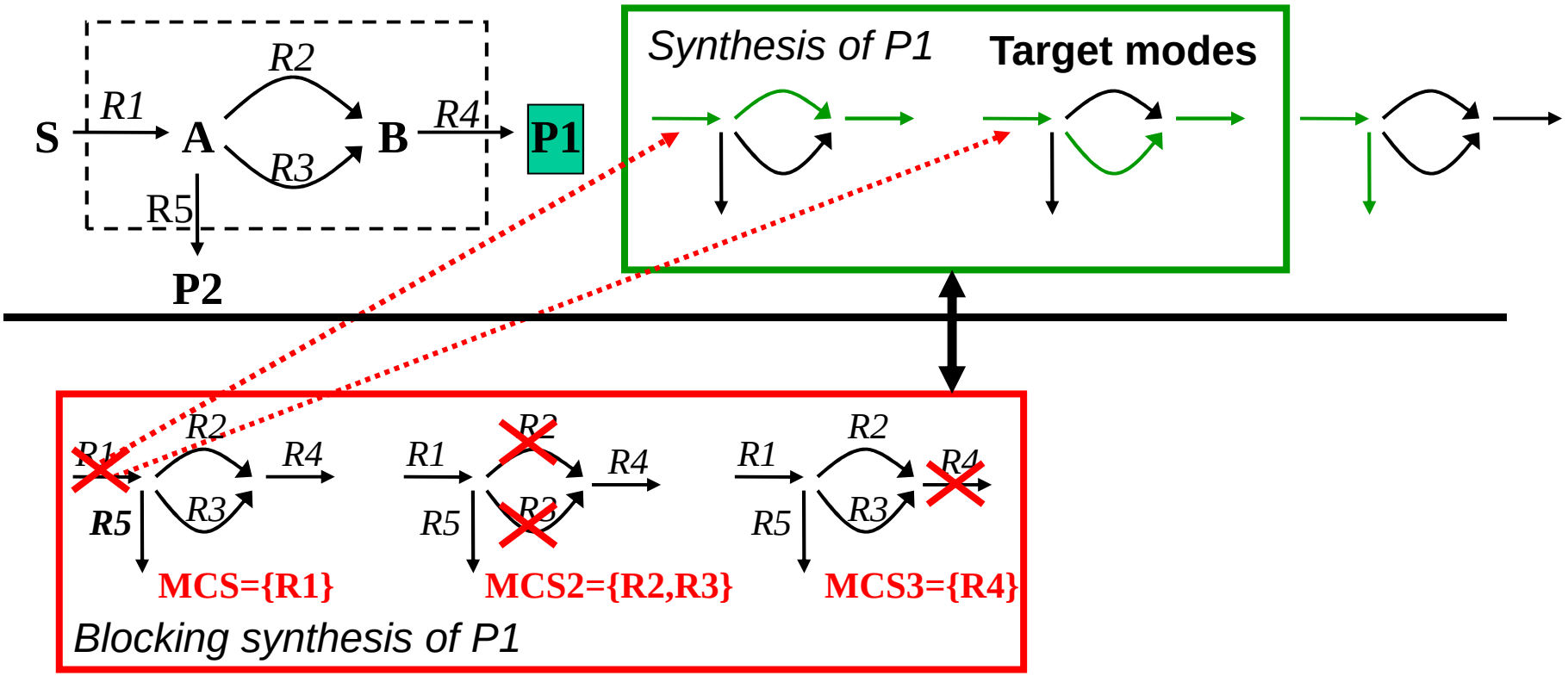
Duality: Elementary Modes and Minimal Cut Sets

EMs and MCSs are **dual representations** of network functions.



MCSs are Minimal Hitting Sets of EMs

Elementary Modes: minimal functional units (pathways) that can operate in steady state

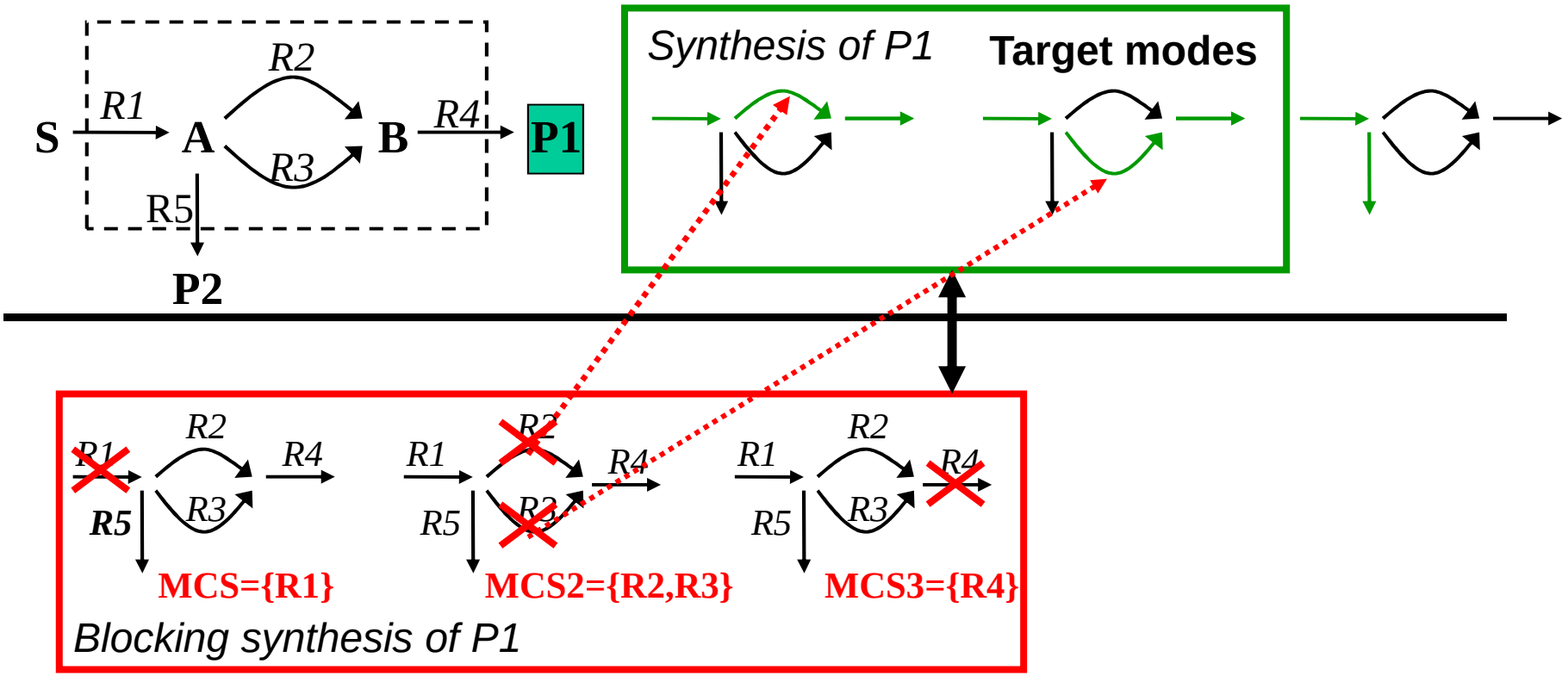


Minimal Cut Set (MCS)*:* minimal set of cuts (knockouts) blocking certain functions in steady state

Klamt S: BioSystems 2006, 83:233-237.

MCSs are Minimal Hitting Sets of EMs

Elementary Modes: minimal functional units (pathways) that can operate in steady state

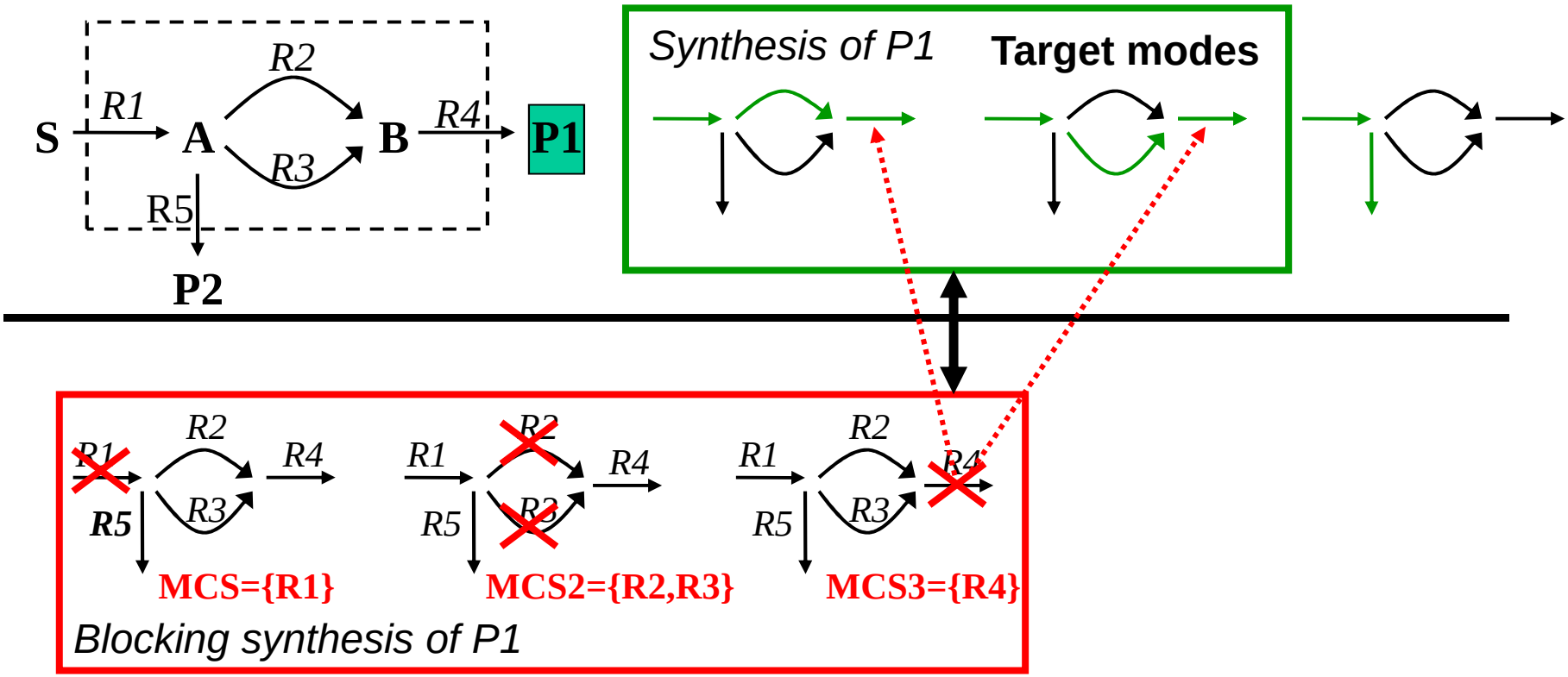


Minimal Cut Set (MCS): minimal set of cuts (knockouts) blocking certain functions in steady state

Klamt S: BioSystems 2006, 83:233-237.

MCSs are Minimal Hitting Sets of EMs

Elementary Modes: minimal functional units (pathways) that can operate in steady state

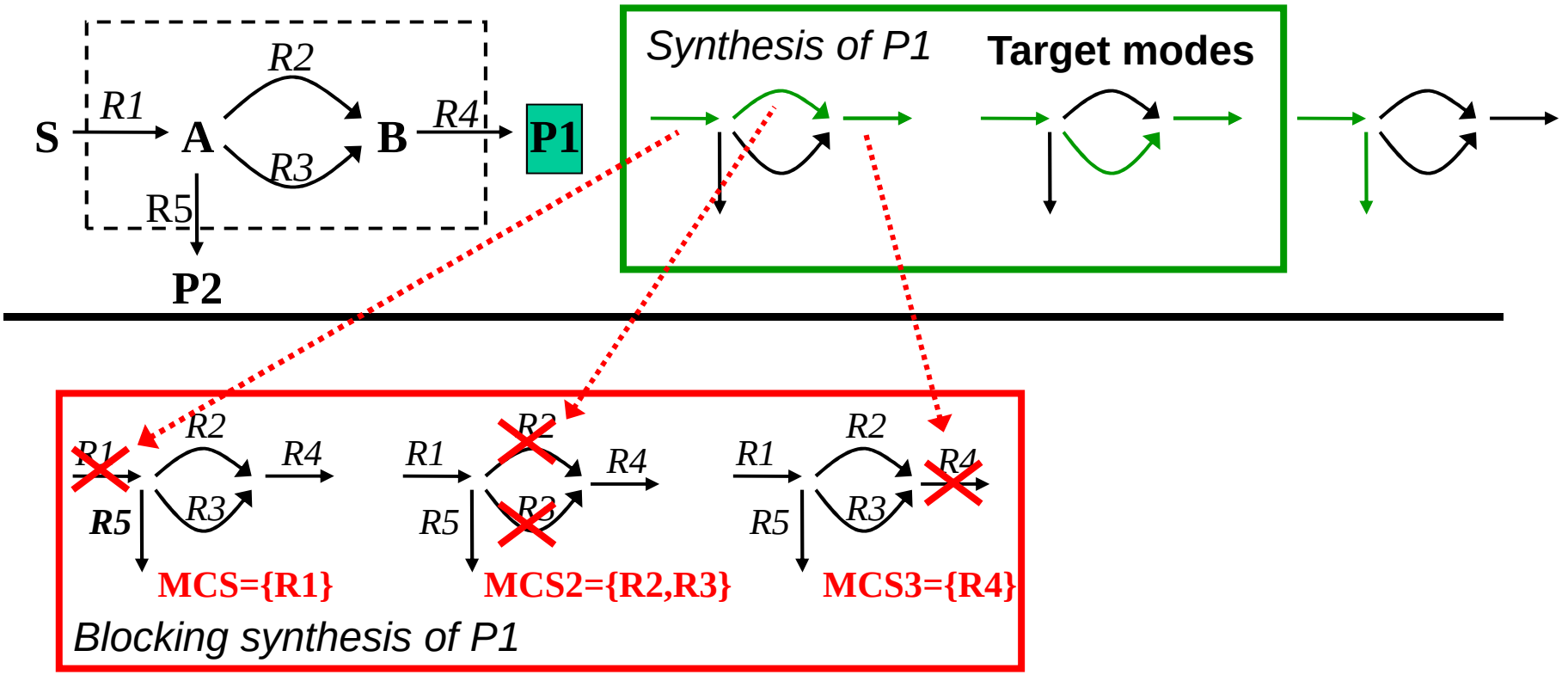


Minimal Cut Set (MCS): minimal set of cuts (knockouts) blocking certain functions in steady state

Klamt S: BioSystems 2006, 83:233-237.

EMs are Minimal Hitting Sets of MCSs!

Elementary Modes: minimal functional units (pathways) that can operate in steady state

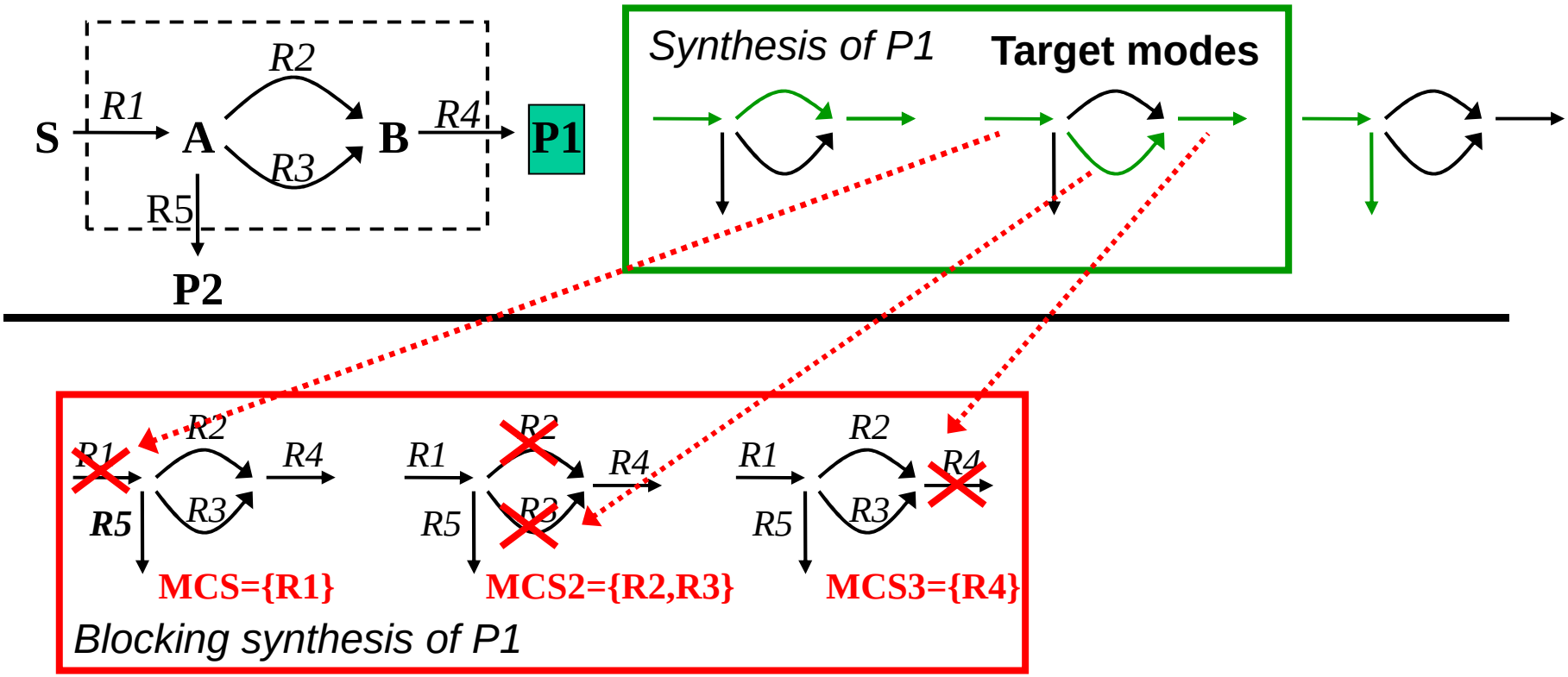


Minimal Cut Set (MCS)*:* minimal set of cuts (knockouts) blocking certain functions in steady state

Klamt S: BioSystems 2006, 83:233-237.

EMs are Minimal Hitting Sets of MCSs!

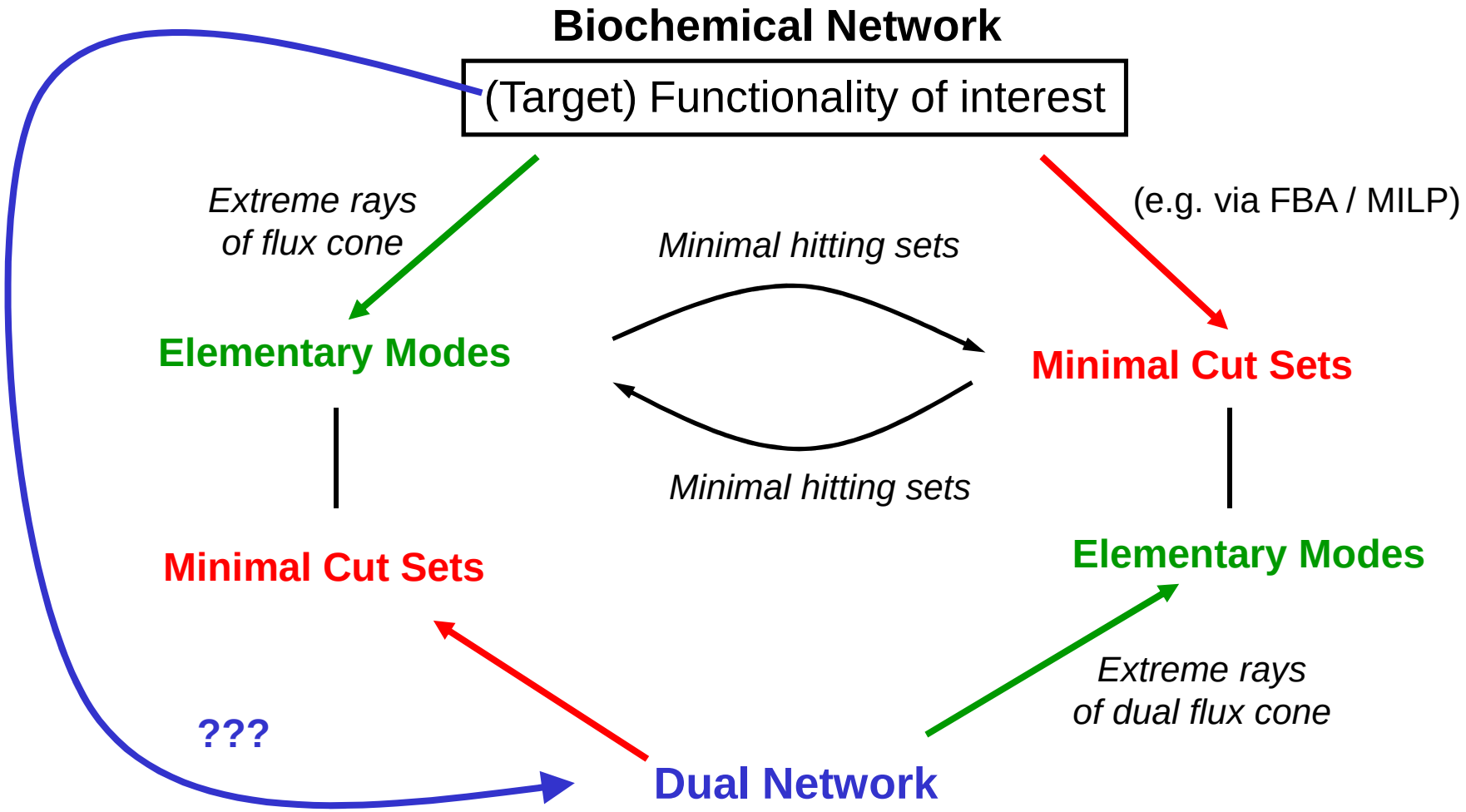
Elementary Modes: minimal functional units (pathways) that can operate in steady state



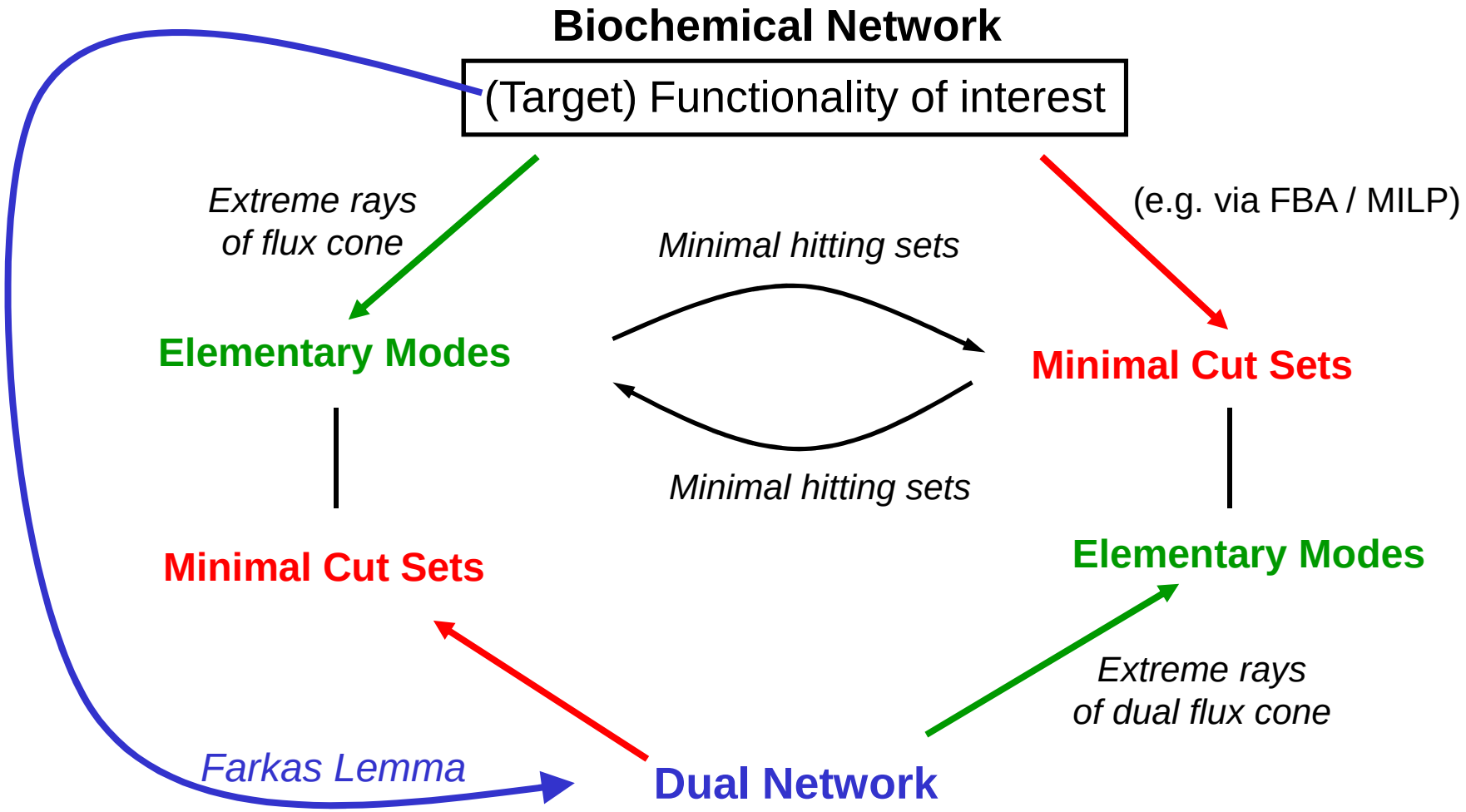
Minimal Cut Set (MCS)*:* minimal set of cuts (knockouts) blocking certain functions in steady state

Klamt S: BioSystems 2006, 83:233-237.

Duality: Elementary Modes and Minimal Cut Sets

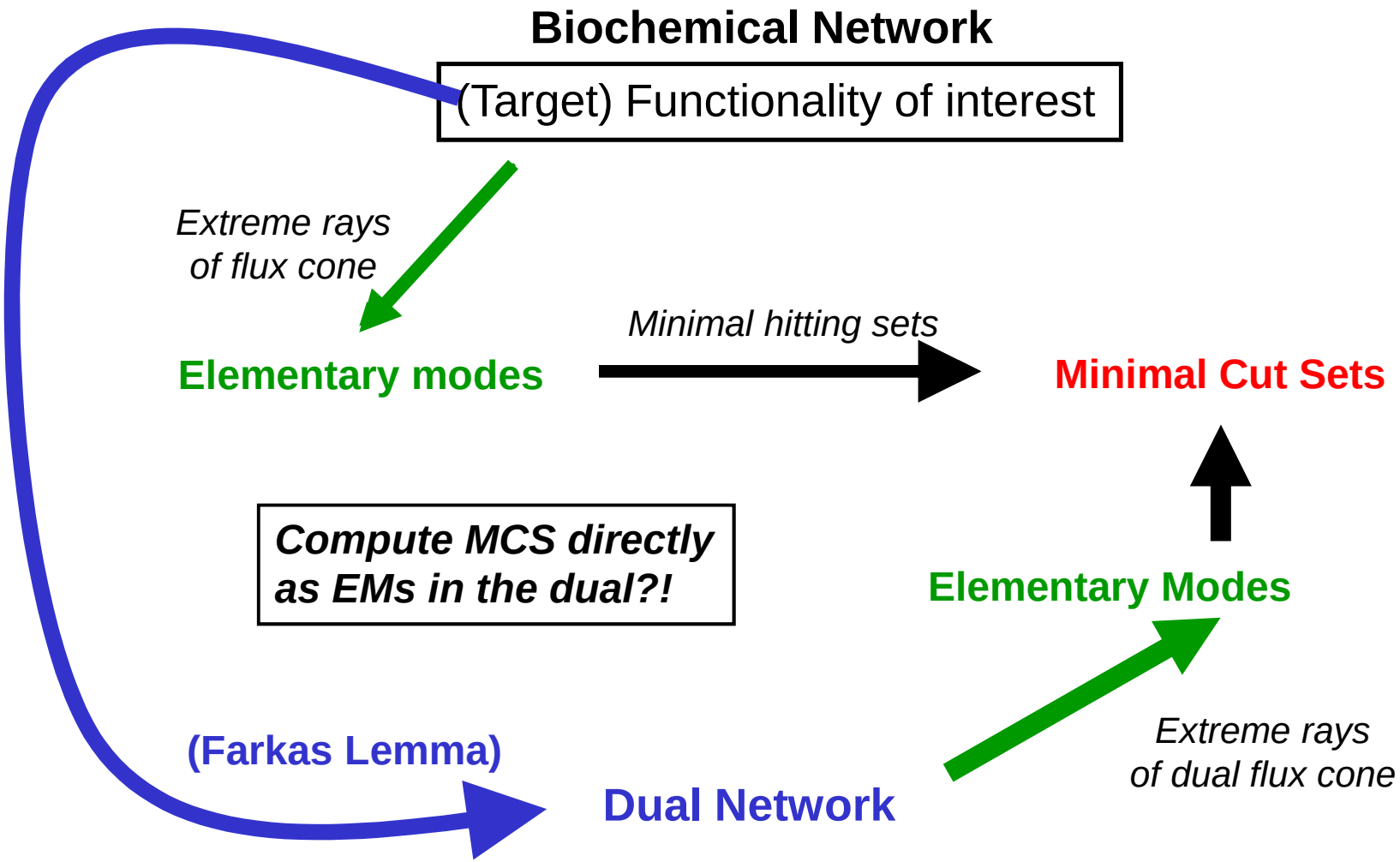


Duality: Elementary Modes and Minimal Cut Sets



Ballerstein K, von KampA, Klamt S and Haus UU. (2012) Minimal cut sets in a metabolic network are elementary modes in a dual network. Bioinformatics 18: 381-387.

Elementary Modes and Minimal Cut Sets: Algorithms and Duality



Ballerstein K, von KampA, Klamt S and Haus UU. (2012) Minimal cut sets in a metabolic network are elementary modes in a dual network. Bioinformatics 18: 381-387.

Computation of MCSs

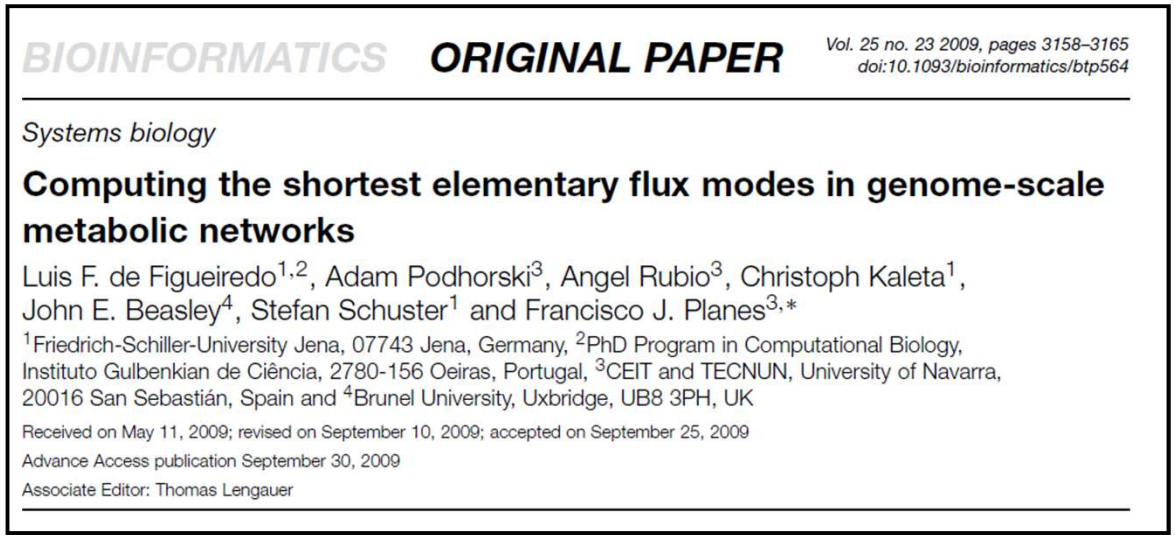
Unfortunately:

- computation of MCS as EMs in the dual does not pay off in most cases!

Full enumeration of MCSs/EMs anyway not realistic in genome-scale networks!

Focus on smallest MCSs reasonable for applications!
→ But: how to compute?

MILP algorithm for calculating shortest EMs in genome-scale networks



BIOINFORMATICS ORIGINAL PAPER
Vol. 25 no. 23 2009, pages 3158–3165
doi:10.1093/bioinformatics/btp564

Systems biology

Computing the shortest elementary flux modes in genome-scale metabolic networks

Luis F. de Figueiredo[1,2], Adam Podhorski[3], Angel Rubio[3], Christoph Kaleta[1], John E. Beasley[4], Stefan Schuster[1] and Francisco J. Planes[3,*]

[1]Friedrich-Schiller-University Jena, 07743 Jena, Germany, [2]PhD Program in Computational Biology, Instituto Gulbenkian de Ciência, 2780-156 Oeiras, Portugal, [3]CEIT and TECNUN, University of Navarra, 20016 San Sebastián, Spain and [4]Brunel University, Uxbridge, UB8 3PH, UK

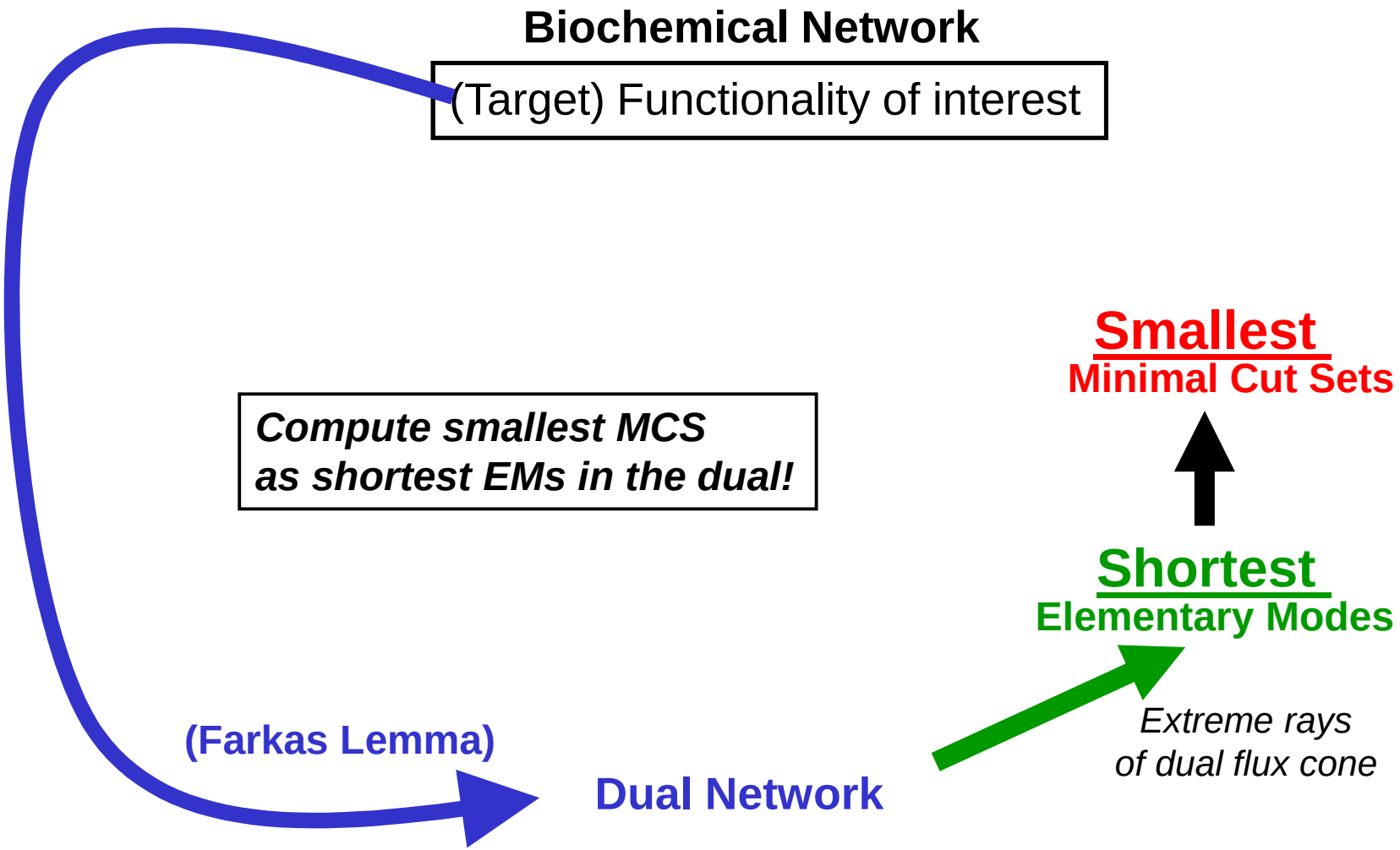
Received on May 11, 2009; revised on September 10, 2009; accepted on September 25, 2009

Advance Access publication September 30, 2009

Associate Editor: Thomas Lengauer

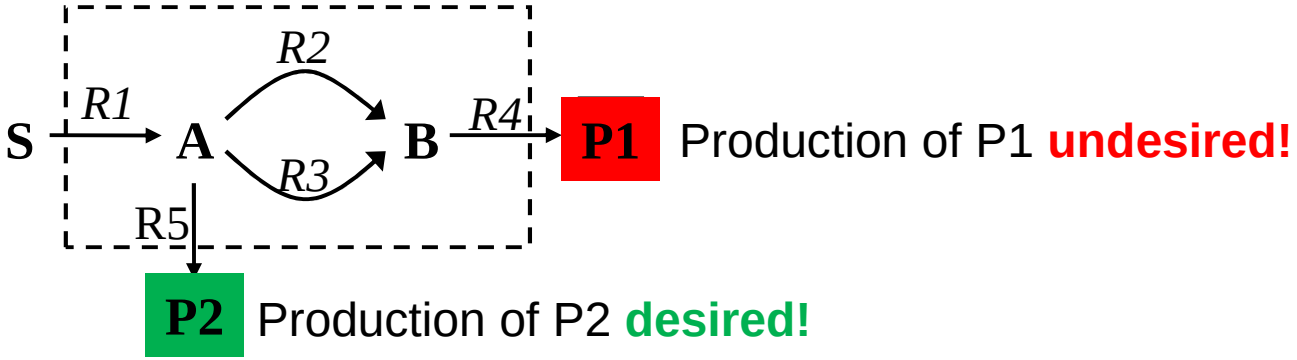
→ Enumerate smallest MCSs as shortest EMs in the dual network!

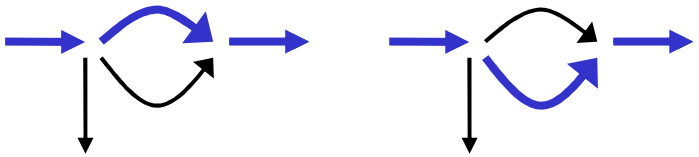
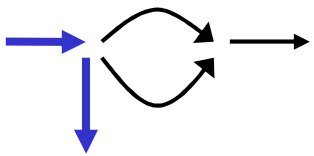
Elementary Modes and Minimal Cut Sets: Algorithms and Duality



Ballerstein K, von KampA, Klamt S and Haus UU. (2012) Minimal cut sets in a metabolic network are elementary modes in a dual network. Bioinformatics 18: 381-387.

Defining Undesired und Desired Metabolic Functions



Specification …	Undesired (targeted) function	Desired function
… via elementary flux modes	*Target / Undesired modes* 	*Desired modes* 
… via equalities and inequalities on fluxes	$r_4>0$ $\mathbf{T\ r \leq b}$	$r_5>0$ $(r_4=0)$ $\mathbf{D\ r \leq a}$

MCSEnumerator: Computation of Smallest (Constrained) MCSs in Genome-Scale Networks

1. Define metabolic network: $\mathbf{Nr} = \mathbf{0}$, $\mathbf{I}_{Irrev}\mathbf{r}_{Irrev} \geq \mathbf{0}$ — *To be provided by the user*

2. Define undesired $\mathbf{T}\,\mathbf{r} \leq \mathbf{b}$ and desired $\mathbf{D}\,\mathbf{r} \leq \mathbf{a}$ behaviors (flux vectors)

3. Dualization of undesired behaviors and formulation of MILP problem

$$\mathbf{Nr} = \mathbf{0}$$
$$\mathbf{I}_{Irrev}\mathbf{r}_{Irrev} \geq \mathbf{0}$$
$$\mathbf{Tr} \leq \mathbf{b}$$

Dualization →

$$\mathbf{N}_{dual}\mathbf{r}_{dual} = \begin{pmatrix}\mathbf{N}^T & -\mathbf{I}_{Irrev}^T & \mathbf{T}^T & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{u}\\ \mathbf{h}\\ \mathbf{w}\\ \mathbf{v}\end{pmatrix} = \mathbf{0}$$

$$\mathbf{b}^T\mathbf{w} \leq -c$$

$$\mathbf{u} \in \Re^m, \mathbf{v} \in \Re^n, \mathbf{h} \in \Re^{|Irrev|}, \mathbf{w} \in \Re^t, \mathbf{h} \geq \mathbf{0}, \mathbf{w} \geq \mathbf{0},\ c > 0$$

MILP problem for shortest EMs
Integer variables: z_i=0 ↔ v_i=0; z_i=1 ↔ $v_i \neq 0$
Objective function: *minimize* $\sum z_i$

von KampA and Klamt S (2014): Enumeration of smallest intervention in genome-scale metabolic networks. PLOS Comp Biol, 10: e1003378.

MCSEnumerator: Computation of Smallest (Constrained) MCSs in Genome-Scale Networks

4.) Enumerate shortest EMs in the dual by solving several MILP problems.
Normally: iteration *k* delivers *k*-th shortest EMs.
MCSEnumerator: iteration *k* delivers all EMs of size *k*.

↓

5.) The found **shortest EMs in the dual** correspond to the **smallest MCSs** for the (primal) undesired behaviors.

↓

6.) Filter ***constrained*** MCSs maintaining desired behavior

→ Solving one LP per MCS (low computational demand).

↓

Shortest constrained MCSs.

von KampA and Klamt S (2014): Enumeration of smallest intervention in genome-scale metabolic networks. PLOS Comp Biol, 10: e1003378.

Application Examples

of constrained MCSs in large-scale metabolic networks

Example 1:

Enumeration of Synthetic Reaction Lethals in *E.coli* model iAF1260

Synthetic reaction lethals (SRLs) = MCSs that block growth

Undesired behavior: $\mu \geq \mu_{min}$
Desired behavior: { }
Model: iAF1260 (2077 reactions; 1387 metabolites)

Original study: *Suthers et al. (2009, Mol Syst Biol)*
→ ***SL Finder*** (bi-level optimization): full enumeration of SRLs up to 3 knockouts

MCS size	number of MCSs	runtime* *SL Finder*	runtime* *MCSEnumerator*
1	277	11.1 s	11.1 s
2	96	1.5 h	39.1 s
3	247	74.1 h	16.8 min
4	402	---	18.5 h
5	1464	---	410.4 h

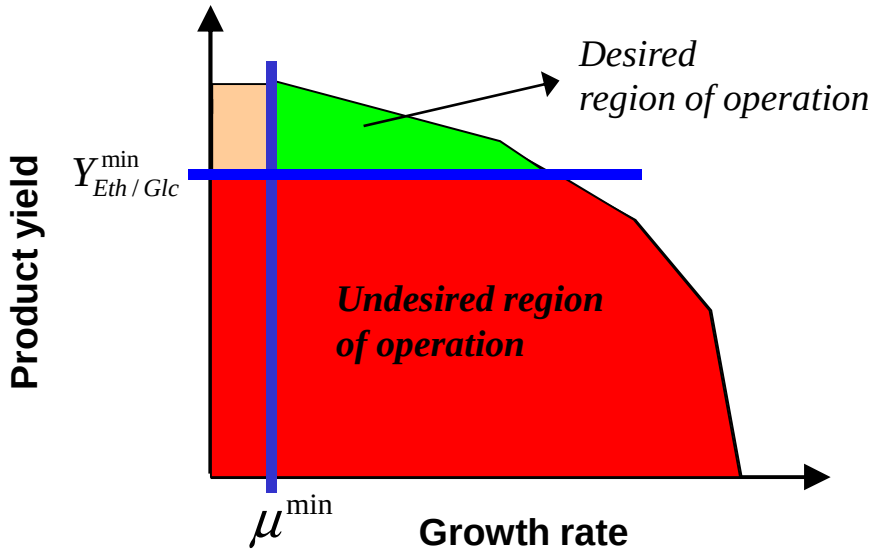
* Computed on mini-cluster with 2 Intel Xeon DP X5650 (6 cores each)
(runtime on typical quadcore is ~1.5 higher)

→ all SRLs up to size 5 can now be computed
→ >100 times faster than SL finder

von KampA and Klamt S (2014): Enumeration of smallest intervention in genome-scale metabolic networks. PLOS Comp Biol, 10: e1003378.

Example 2: constrained MCSs for coupling anaerobic growth and ethanol synthesis in *E.coli* (model: iAF1260)

Undesired behavior	**Desired behavior**
*maintATP = 8.39 mmol/(gDW*h)* *oxygen uptake = 0* *max. glucose uptake: 18.5 mmol/(gDW*h)*	
ethanol yield ≤ 1.4 / 1.8	ethanol yield > 1.4 / 1.8 min. growth rate μ_{min} ≥= 0.001 h^{-1}



→ high product yield guaranteed under any growth rate
→ minimal growth is feasible

von KampA and Klamt S (2014): Enumeration of smallest intervention in genome-scale metabolic networks. PLOS Comp Biol, 10: e1003378.

Example 2: constrained MCSs for coupling anaerobic growth and ethanol synthesis in *E.coli* (model: iAF1260)

Scenario	$Y_{Eth/Glc}^{min}$	# MCSs	# cMCSs	cMCSs size							runtime [h]
				1	2	3	4	5	6	7	
1	1.4	156477	**8819**	0	0	2	98	533	1737	6449	**16.6**
2	1.8	138675	**4618**	0	0	2	70	509	917	3120	**20.9**

Some results / conclusions:

→ Full enumeration of cMCSs up to size 7 within one day.

→ More than 4000/8000 intervention strategies found.
Ony 3-6% of the computed MCSs (>100,000) are valid cMCSs.

→ Many cMCSs correspond to well-known strategies.

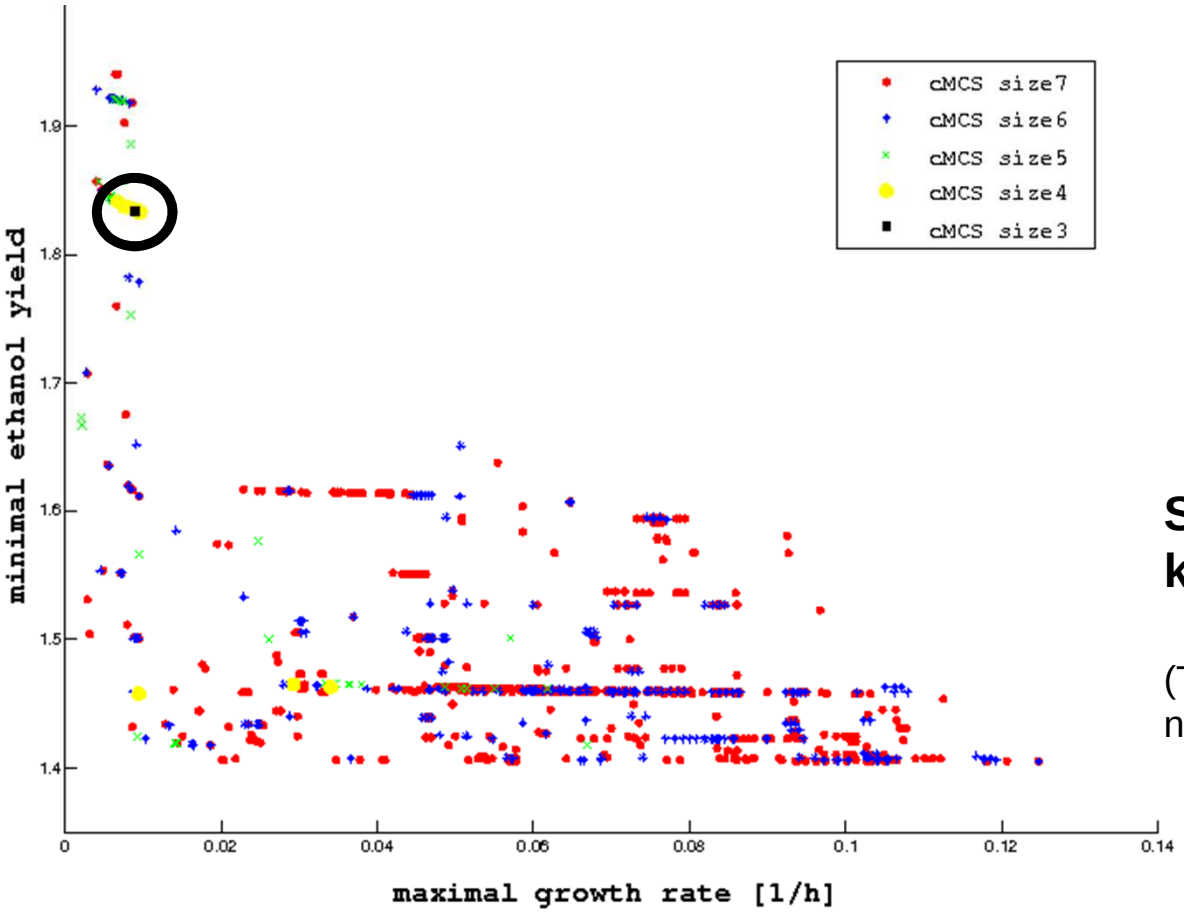
→ Some cMCSs with only 3 knockouts; even for high-yield strategy.
→ Experimental validation ongoing.

→ Pick optimal intervention strategy via performance measures …

von KampA and Klamt S (2014): Enumeration of smallest intervention in genome-scale metabolic networks. PLOS Comp Biol, 10: e1003378.

Picking Optimal Knockout Strategies via Performance Parameters

Min-Ethanol-Yield vs. Max-Growth-Rate
for all cMCSs of scenario 1 ($Y_{Eth/Glc}^{\min} = 1.4$)

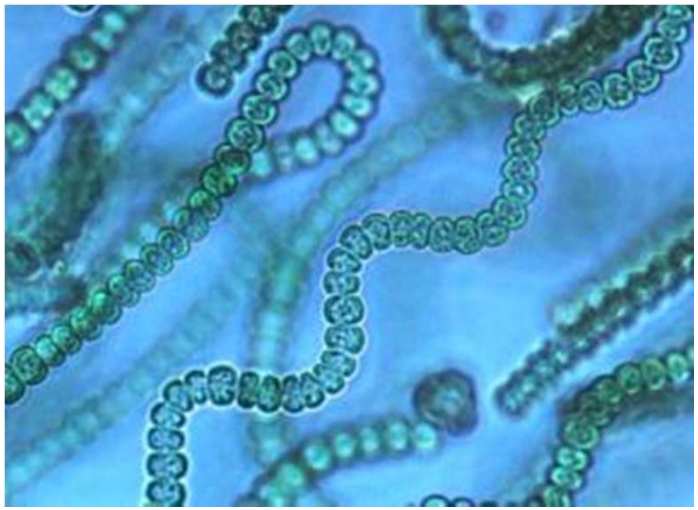


Select the optimal knockout strategy!

(The „smallest“ are not always the best.)

von KampA and Klamt S (2014): Enumeration of smallest intervention in genome-scale metabolic networks. PLOS Comp Biol, 10: e1003378.

Example 3: constrained MCSs for coupling photosynthetic growth and ethanol synthesis in cyanobacteria



- joint work with Steuer group (HU Berlin)
- genome-scale model of *Synechocystis* sp. PCC 6803 (Knoop et al. (2013))
- **undesired behavior:**
  → low yield of ethanol (per mmol photons)
- **desired behavior:**
  → high yield of ethanol and growth possible

Key results:

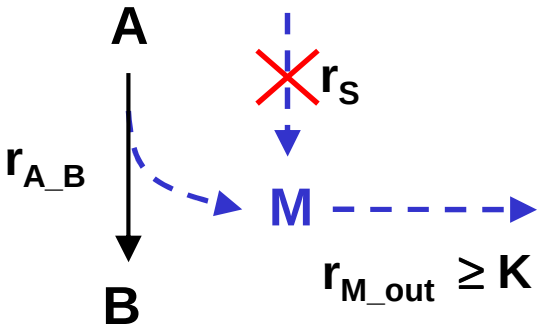
- **coupling** of biomass and ethanol synthesis under photoautotropic conditions **much harder** than in heterotrophic organisms: ≥ 14 knockouts required
- all cMCSs of **size 14, 15 and 16** enumerated within one day (1564 cMCSs)
  → reveal principles of suitable coupling strategies
  → selected strategies **under experimental validation** (with *Algenol Biofuels Germany*)

Erdrich P, Knoop H, Steuer R, Klamt S: Cyanobacterial biofuels: new insights and engineering strategies revealed by computational modeling. Accepted

Recent Work: Extension to Regulatory cMCSs

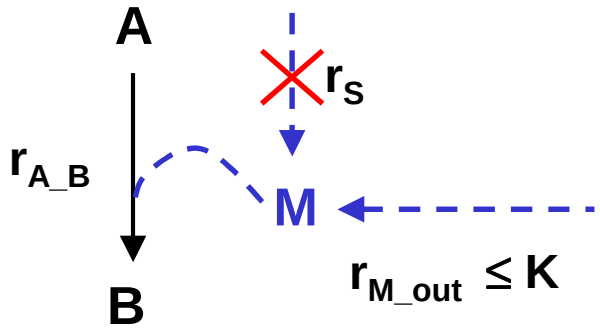
- joint work with **Radhakrishnan Mahadevan (Toronto)**
- extend concept of cMCSs to allow also **up / downregulation** of metabolic fluxes
- idea: represent up / downregulation as „cuts“:

Representing upregulation as a reaction cut



Knockout of r_S induces upregulation of flux r_{A_B} ($\geq$ **K**)

Representing downregulation as a reaction cut



Knockout of r_S induces downregulation of flux r_{A_B} ($\leq$ **K**)

Mahadevan R, von Kamp A and Klamt S: Geneome-scale strain design based on regulatory minimal cut sets. Submitted

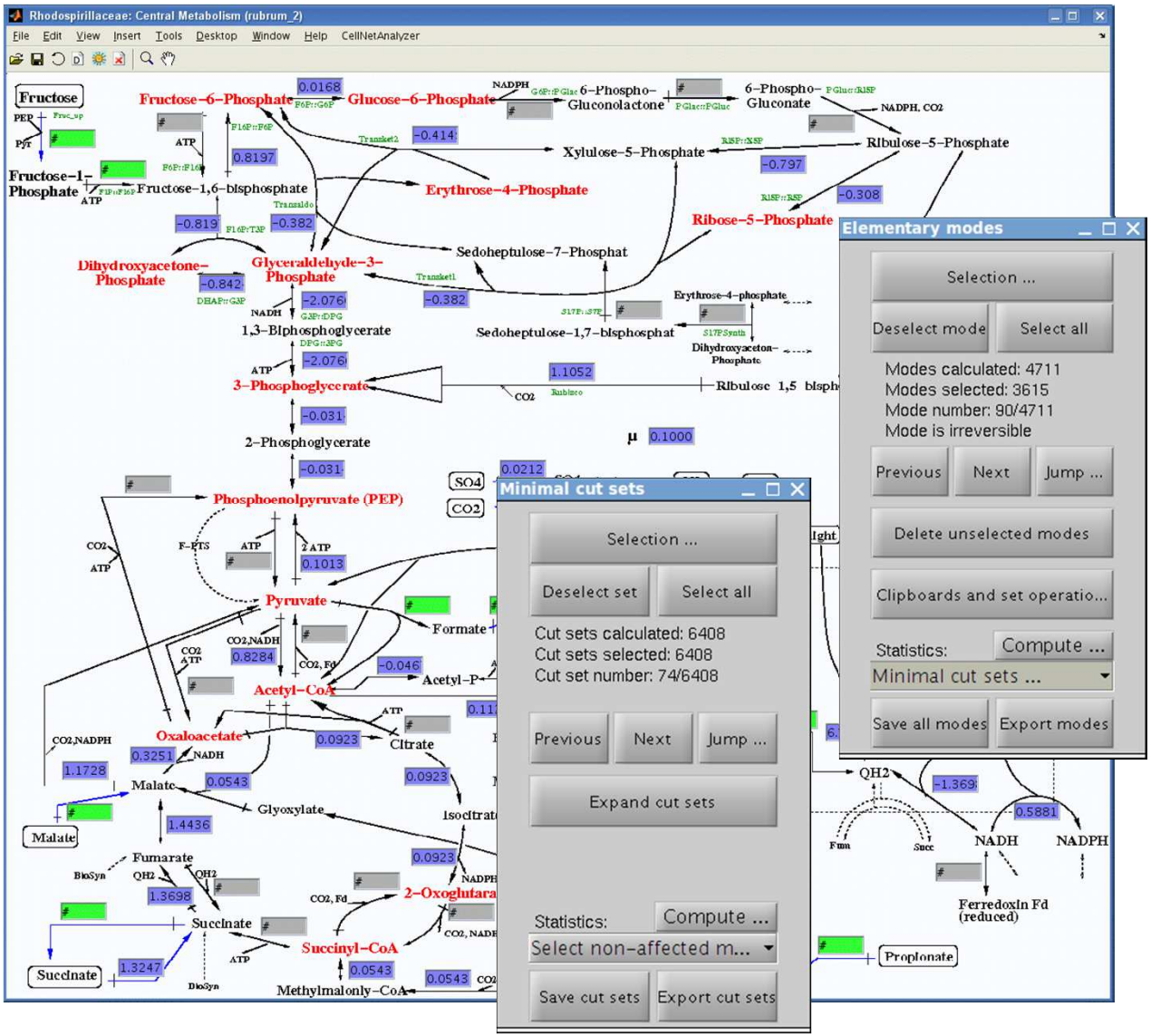
CellNetAnalyzer

MAX-PLANCK-GESELLSCHAFT

MATLAB toolbox for Biological Network Analysis.

Large number of methods for stoichiometric and metabolic network analysis including computation and analysis of EMs and (c)MCSs).

Free for academic use:
http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html



Conclusions

- **Stoichiometric Network Analysis:** an established framework for exploring (large-scale) metabolic networks.
- **Elementary Modes**: for studying functional properties of metabolic networks.
- **Constrained Minimal Cut Sets**: a universal and flexible approach for formulating and solving intervention problems in reaction networks.
- EMs and MCSs are **dual representations** of network functions.
- **Duality** between EMs and MCSs offers new *computational* and *conceptual* perspectives for exploring functions and design strategies in metabolic networks.
- **New and effective algorithm** for enumerating smallest (c)MCSs in genome-scale networks (via shortest EMs in the dual system)

Applications of EMs, MCSs, … to Macrobiology?

Related topics in ecological/ecosystems modeling:

- stoichiometric models of microbial communities (work in progress in our lab)
- material flow analysis in ecosystems (e.g., in food webs, biogeochemical models)
- „ecological stoichiometry“

Acknowledgements

Constrained MCS
•Oliver Hädicke
(MPI Magdeburg)

Implementation
• Axel von Kamp
(MPI Magdeburg)

Duality
• Kathrin Ballerstein (ETH Zurich)
• Utz-Uwe Haus (ETH Zurich)

Cyanobacteria biofuels
• Philipp Erdrich (MPI Magdeburg)
• Ralf Steuer (HU Berlin)
• Henning Knoop (HU Berlin)

THANK YOU!