V16 Metabolic Pathway Analysis (MPA)

Metabolic Pathway Analysis searches for meaningful structural and functional units in metabolic networks.

Today‘s most powerful methods are based on **convex analysis**.

2 such approaches are the **elementary flux modes** [1] and **extreme pathways**[2].

Both sets span the space of feasible steady-state flux distributions by non-decomposable routes, i.e. no subset of reactions involved in an EFM or EP can hold the network balanced using non-trivial fluxes.

Extreme pathways are a subset of elementary modes.
For many systems, both methods coincide.

Klamt et al. Bioinformatics 19, 261 (2003); Trinh et al. Appl. Microbiol Biotechnol. 81, 813-826 (2009)
[1] Schuster & Hilgetag J Biol Syst 2, 165-182 (1994), Pfeiffer et al. Bioinformatics, 15, 251 (1999), Schuster et al. Nature Biotech. 18, 326 (2000)
[2] Schilling et al. J Theor Biol 203, 229-248 (2000)

Applications of Metabolic Pathway Analysis (MPA)

MPA can be used to study e.g.

- metabolic network structure

- functionality of networks (including identification of futile cycles)

- robustness, fragility, flexibility/redundancy of networks

- to identiy all (sub-) optimal pathways with respect to product/biomass yield

- rational strain design

Klamt et al. Bioinformatics 19, 261 (2003) ; Trinh et al. Appl. Microbiol Biotechnol. 81, 813-826 (2009)

Definition of Elementary Flux Modes (EFMs)

A pathway P(**v**) is an **elementary flux mode** if it fulfills conditions C1 – C3.

(C1) **Pseudo steady-state**. **S** · **e** = 0. This ensures that none of the metabolites is consumed or produced in the overall stoichiometry.

(C2) **Feasibility**: rate $e_i \geq 0$ if reaction is irreversible. This demands that only thermodynamically realizable fluxes are contained in **e**.

(C3) **Non-decomposability***:* there is no vector **v** (except the null vector and **e**) fulfilling C1 and C2 and so that P(**v**) is a proper subset of P(**e**).
This is the core characteristics for EFMs and EPs and provides the decomposition of the network into smallest units that are able to hold the network in steady state.

C3 is often called „genetic independence“ because it implies that the enzymes in one EFM or EP are not a subset of the enzymes from another EFM or EP.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Definition of Extreme Pathways (Eps)

The pathway *P(**e**)* is an **extreme pathway** if it fulfills conditions C1 – C3 AND conditions C4 – C5.

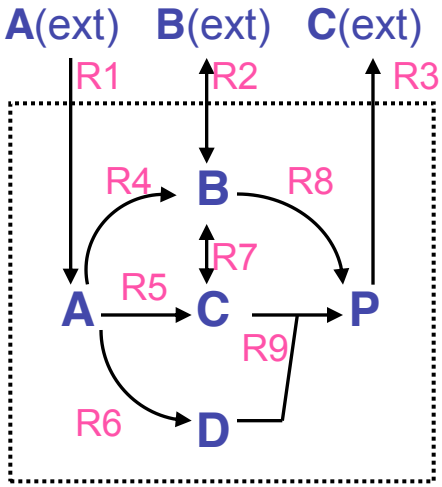
(C4) **Network reconfiguration**: Each reaction must be classified either as exchange flux or as internal reaction.
All **reversible** internal reactions must be **split up** into two separate, irreversible reactions (forward and backward reaction).

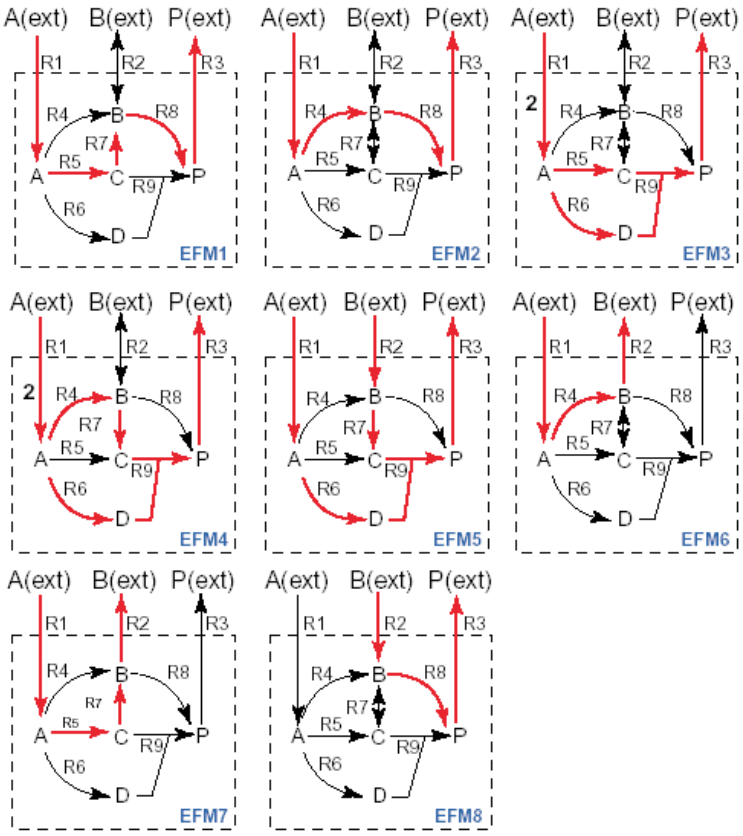
(C5) **Systemic independence**: the set of EPs in a network is the **minimal** set of EFMs that can describe all feasible steady-state flux distributions.

The algorithms for computing EPs and EFMs are quite similar.
We will not cover the algorithmic differences here.

Klamt & Stelling Trends Biotech 21, 64 (2003)

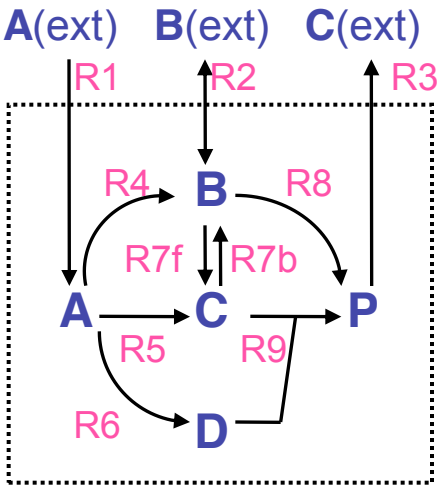
Comparison of EFMs and EPs

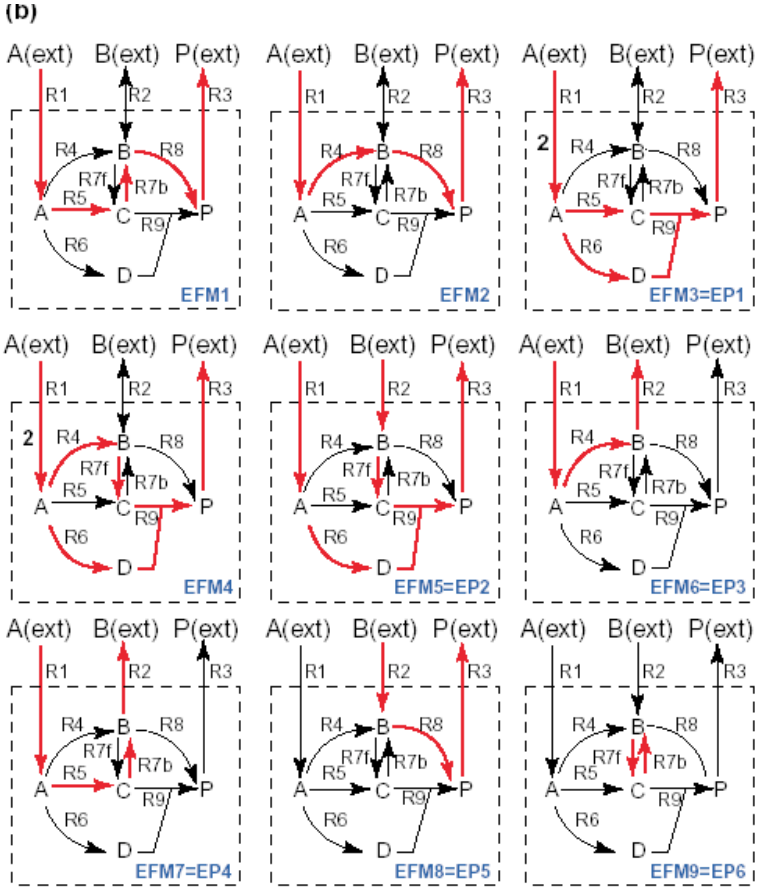




Klamt & Stelling Trends Biotech 21, 64 (2003)

Reconfigured Network: split up R7





3 EFMs are not systemically independent:
EFM1 = EP4 + EP5
EFM2 = EP3 + EP5
EFM4 = EP2 + EP3

Klamt & Stelling Trends Biotech 21, 64 (2003)

Property 1 of EFMs

The only difference in the set of EFMs emerging upon reconfiguration consists in the two-cycles that result from splitting up reversible reactions.
However, two-cycles are not considered as meaningful pathways.

Valid for any network: Property 1
Reconfiguring a network by splitting up reversible reactions leads to the same set of meaningful EFMs.

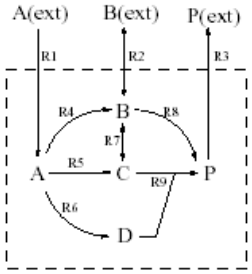
Klamt & Stelling Trends Biotech 21, 64 (2003)

EFMs vs. EPs

What is the consequence when all exchange fluxes (and hence all reactions in the network) are made irreversible?

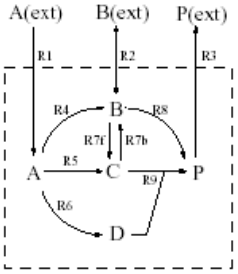
Table 1. Configurations of the example network (upper part N1 and N3; lower part N2 and N4), with corresponding elementary flux modes (EFM) and extreme pathways (EP) (see also Fig. 1)

N1 (R2 and R7 reversible) N3 (as N1 but R2 irreversible)



N1		N3								
EFMs		EFMs	R1	R2	R3	R4	R5	R6	R7	R8	R9
EFM1		×	1	0	1	0	1	0	−1	1	0
EFM2		×	1	0	1	1	0	0	0	1	0
EFM3		×	2	0	1	0	1	1	0	0	1
EFM4		×	2	0	1	1	0	1	1	0	1
EFM5		×	1	1	1	0	0	1	1	0	1
EFM6			1	−1	0	1	0	0	0	0	0
EFM7			1	−1	0	0	1	0	−1	0	0
EFM8		×	0	1	1	0	0	0	0	1	0

N2 (R2 reversible, R7 split up) N4 (as N2 but R2 irreversible)



N2		N4										
EFMs	EPs	EFMs	EPs	R1	R2	R3	R4	R5	R6	R7f	R8	R9	R7b
EFM1		×	EP1′	1	0	1	0	1	0	0	1	0	1
EFM2		×	EP2′	1	0	1	1	0	0	0	1	0	0
EFM3	EP1	×	EP3′	2	0	1	0	1	1	0	0	1	0
EFM4		×	EP4′	2	0	1	1	0	1	1	0	1	0
EFM5	EP2	×	EP5′	1	1	1	0	0	1	1	0	1	0
EFM6	EP3			1	−1	0	1	0	1	0	0	0	0
EFM7	EP4			1	−1	0	0	1	0	0	0	0	1
EFM8	EP5	×	EP6′	0	1	1	0	0	0	0	1	0	0
EFM9	EP6	×	EP7	0	0	0	0	0	0	1	0	0	1

Then EFMs and EPs always co-incide!

Klamt & Stelling Trends Biotech 21, 64 (2003)

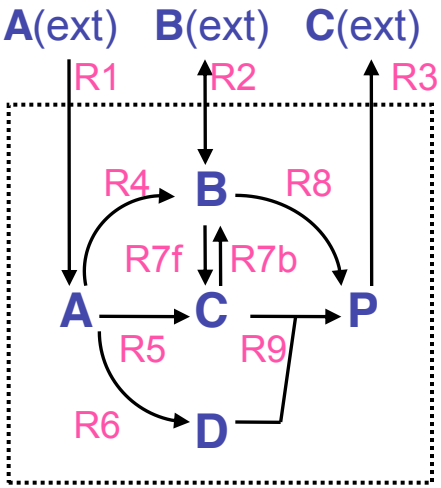
Property 2 of EFMs

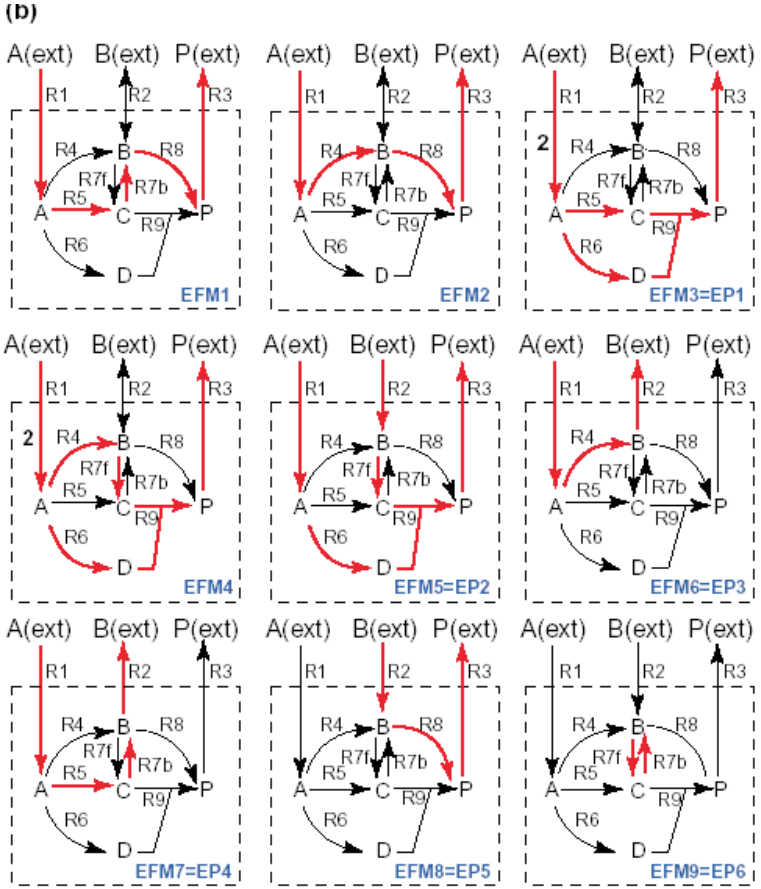
Property 2

If all exchange reactions in a network are irreversible then the sets of meaningful EFMs (both in the original and in the reconfigured network) and EPs coincide.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Reconfigured Network



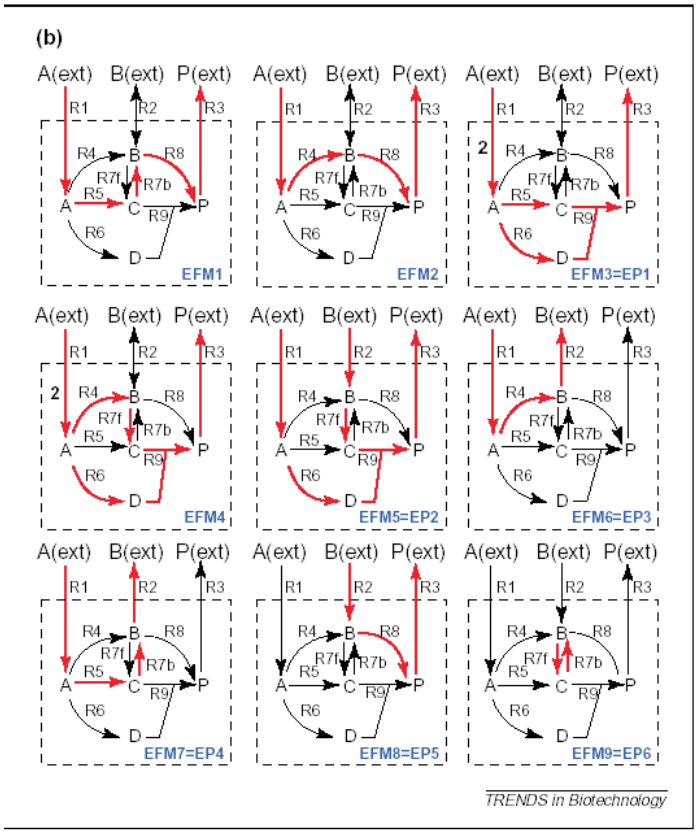


3 EFMs are not systemically independent:
EFM1 = EP4 + EP5
EFM2 = EP3 + EP5
EFM4 = EP2 + EP3

Klamt & Stelling Trends Biotech 21, 64 (2003)

Operational modes

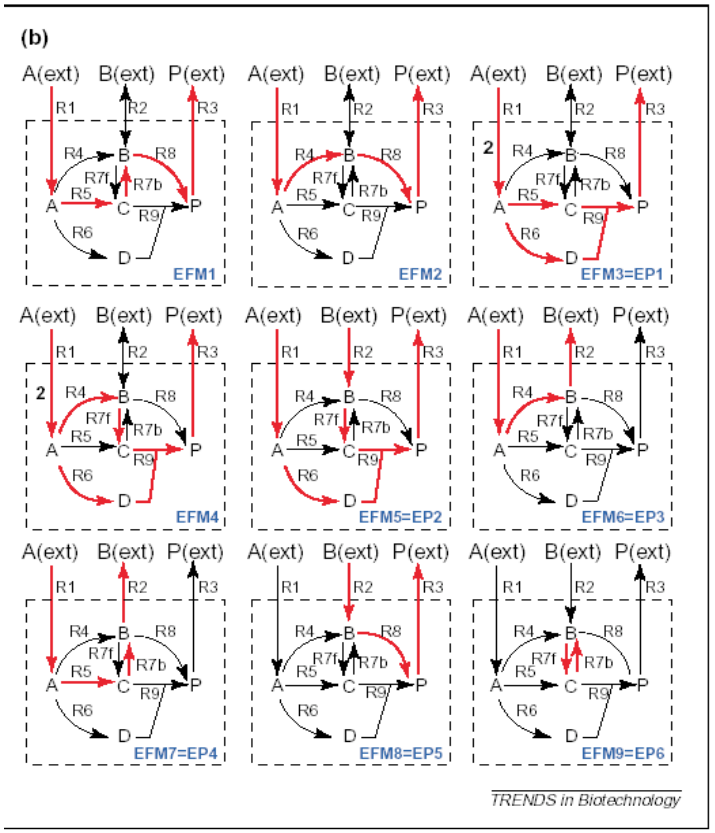
Problem	EFM (network N1)	EP (network N2)
Recognition of operational modes: routes for converting exclusively A to P.	4 genetically independent routes (EFM1-EFM4)	Set of EPs does not contain all genetically independent routes, only EP1. No EP leads directly from A to P via B.



N1 and (b) the EFMs and extreme pathways in network N2 (see also Table 1).

Finding optimal routes

Problem	EFM (network N1)	EP (network N2)
Finding all the optimal routes: optimal pathways for synthesizing P during growth on A alone.	EFM1 and EFM2 are optimal because they yield one mole P per mole substrate A (i.e. R3/R1 = 1), whereas EFM3 and EFM4 are only sub-optimal (R3/R1 = 0.5).	One would only find the suboptimal EP1, not the optimal routes EFM1 and EFM2.

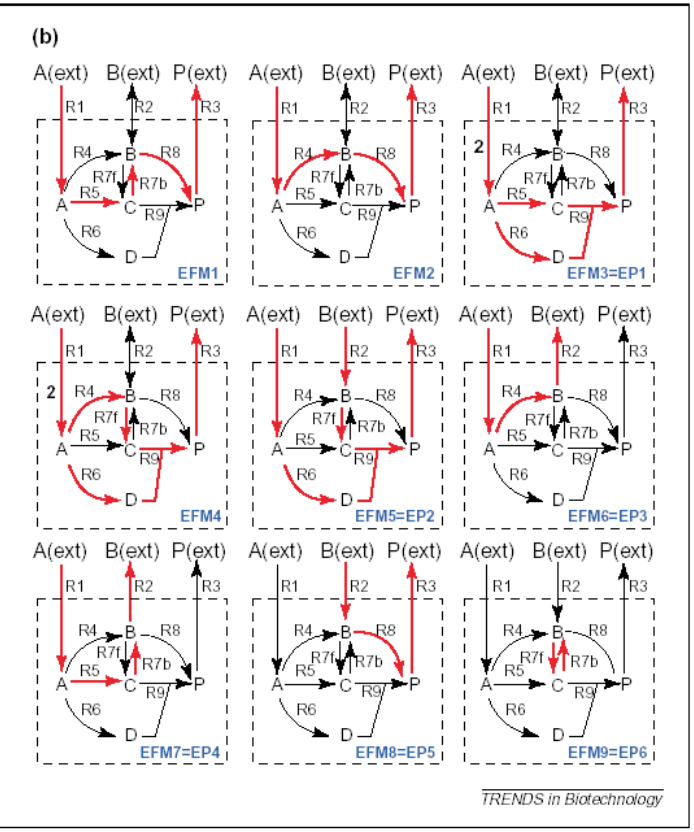


l and (b) the EFMs and extreme pathways in network N2 (see also Table 1).

Network flexibility (structural robustness, redundancy)

Problem

Analysis of network flexibility: relative robustness of exclusive growth on A or B.



N1 and (b) the EFMs and extreme pathways in network N2 (see also Table 1).

EFM (network N1)

4 pathways convert A to P (EFM1-EFM4), whereas for B only one route (EFM8) exists.

When one of the internal reactions (R4-R9) fails, 2 pathways will always „survive“ for production of P from A.
By contrast, removing reaction R8 already stops the production of P from B alone.

EP (network N2)

Only 1 EP exists for producing P by substrate A alone (EP1), and 1 EP for synthesizing P by (only) substrate B (EP5).

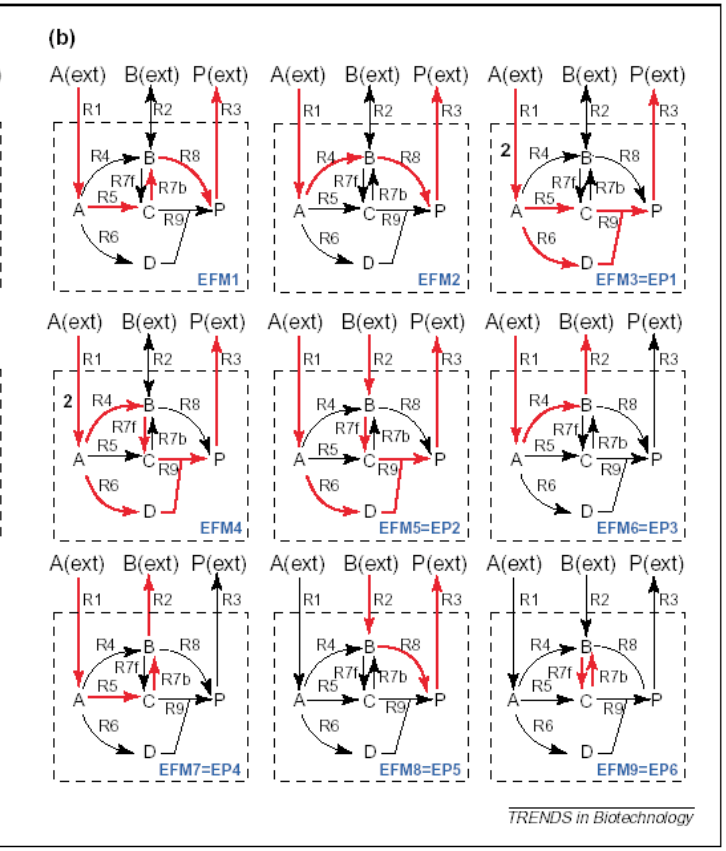
This suggests that both substrates possess the same redundancy of pathways, but as shown by EFM analysis, growth on substrate A is much more flexible than on B.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Relative importance of single reactions

Problem

Relative importance of single reactions:
relative importance of reaction R8.



k N1 and (b) the EFMs and extreme pathways in network N2 (see also Table 1).

EFM (network N1)

R8 is essential for producing P by substrate B (EFM8), whereas for A there is no structurally „favored“ reaction (R4-R9 all occur twice in EFM1-EFM4).

However, considering the optimal modes EFM1, EFM2, one recognizes the importance of R8 also for growth on A.

EP (network N2)

Consider again biosynthesis of P from substrate A (EP1 only).

Because R8 is not involved in EP1 one might think that this reaction is not important for synthesizing P from A.

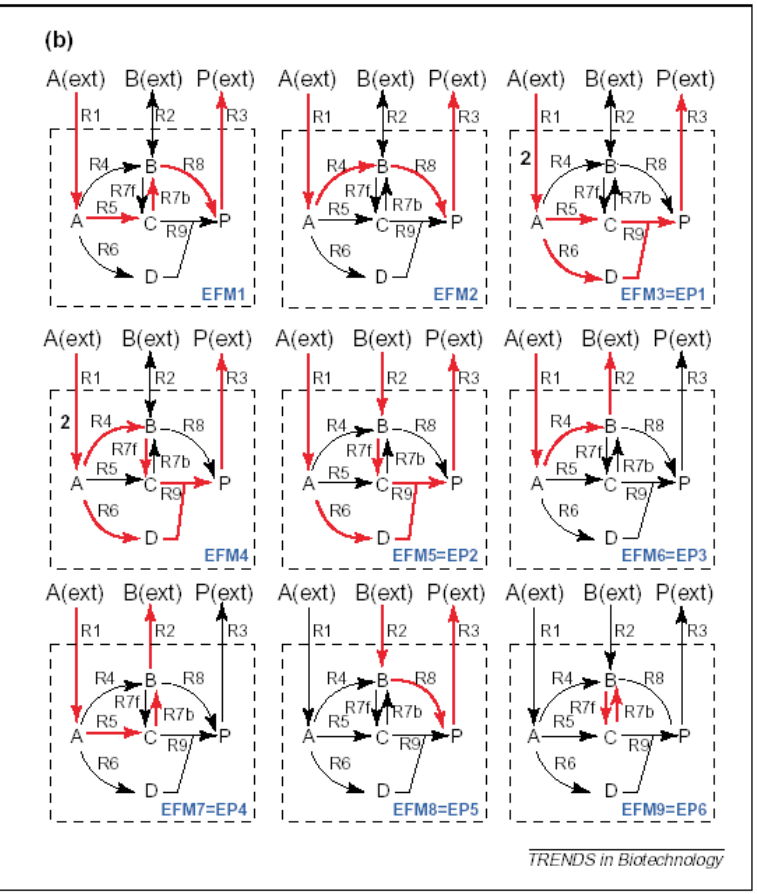
However, without this reaction, it is impossible to obtain optimal yields (1 P per A; EFM1 and EFM2).

(2003) Klamt & Stelling Trends Biotech 21, 64 (2003)

Enzyme subsets and excluding reaction pairs

Problem

Enzyme subsets and excluding reaction pairs: suggest regulatory structures or rules.



N1 and (b) the EFMs and extreme pathways in network N2 (see also Table 1).

EFM (network N1)

R6 and R9 are an **enzyme subset**.

By contrast, R6 and R9 never occur together with R8 in an EFM.

Thus (R6,R8) and (R8,R9) are excluding reaction pairs.
(In an arbitrary composable steady-state flux distribution they might occur together.)

EP (network N2)

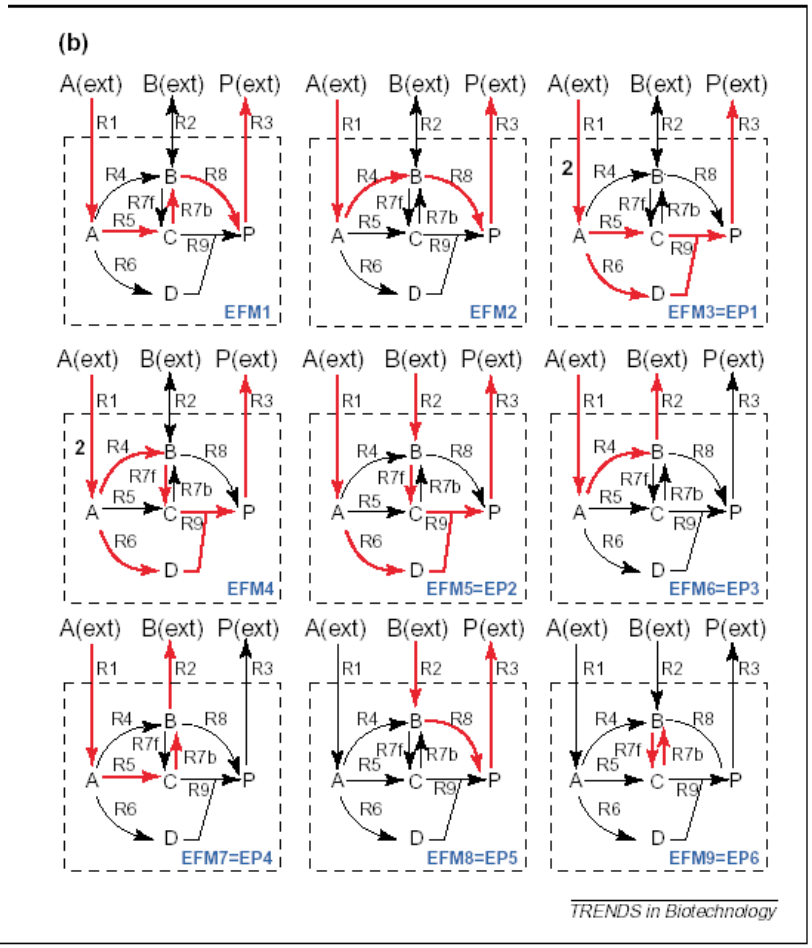
The EPs pretend R4 and R8 to be an excluding reaction pair – but they are not (EFM2).

The enzyme subsets would be correctly identified in this case. However, one can construct simple examples where the EPs would also pretend wrong enzyme subsets (not shown).

Klamt & Stelling Trends Biotech 21, 64 (2003)

Pathway length

Problem	EFM (network N1)	EP (network N2)
Pathway length: shortest/longest pathway for production of P from A.	The shortest pathway from A to P needs 2 internal reactions (EFM2), the longest 4 (EFM4).	Both the shortest (EFM2) and the longest (EFM4) pathway from A to P are not contained in the set of EPs.

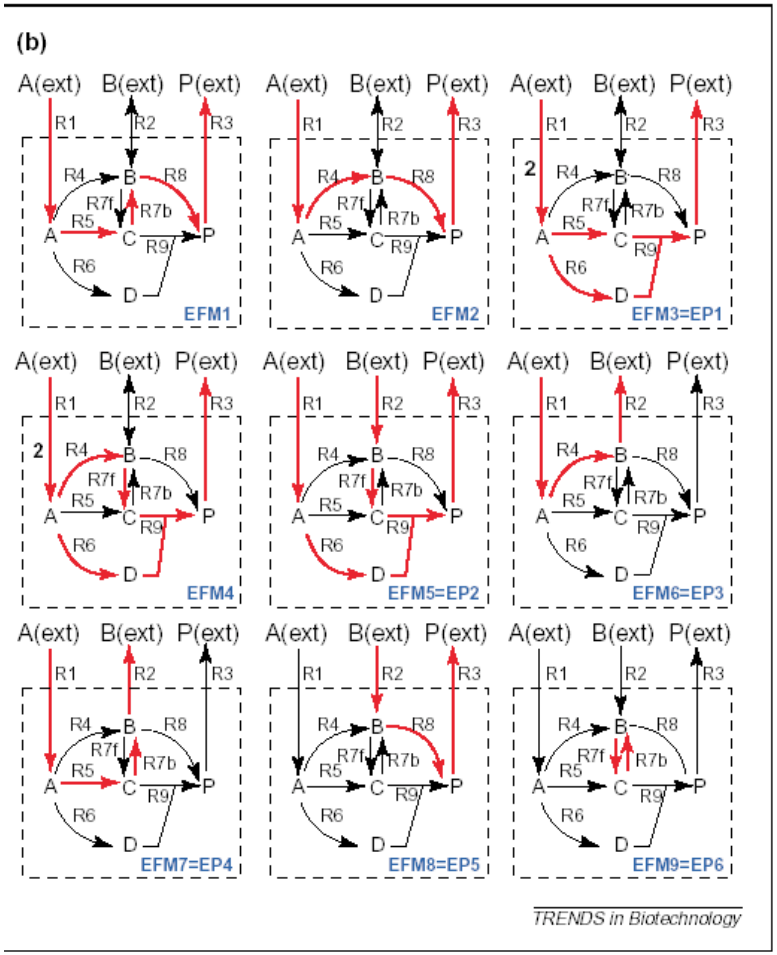


Klamt & Stelling Trends Biotech 21, 64 (2003)

Removing a reaction and mutation studies

Problem

Removing a reaction and mutation studies: effect of deleting R7.



and (b) the EFMs and extreme pathways in network N2 (see also Table 1).

EFM (network N1)

All EFMs not involving the specific reactions build up the complete set of EFMs in the new (smaller) sub-network.

If R7 is deleted, EFMs 2,3,6,8 „survive“. Hence the mutant is viable.

EP (network N2)

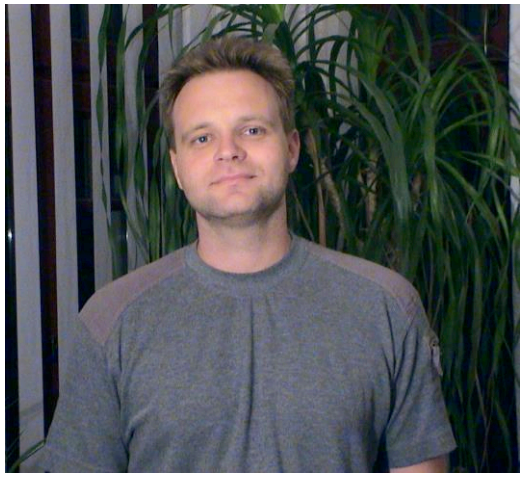
Analyzing a subnetwork implies that the EPs must be newly computed.

E.g. when deleting R2, EFM2 would become an EP.

For this reason, mutation studies cannot be performed easily.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Software: FluxAnalyzer, based on Matlab



Steffen Klamt.

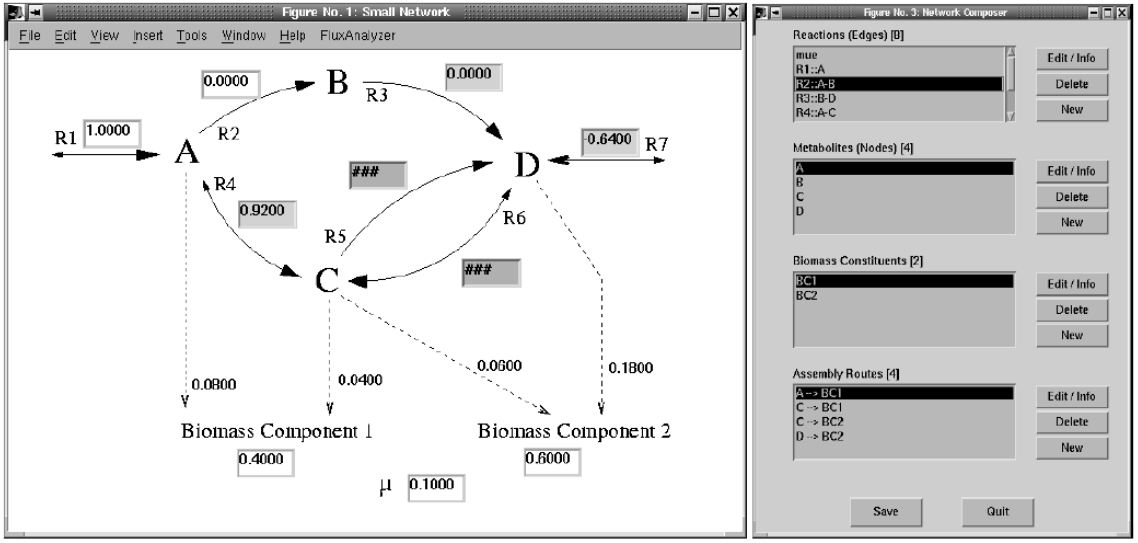


Fig. 1. The network project of 'SMALLNET' constructed by the FluxAnalyzer. Left: interactive flux map displaying a flux scenario (unknown rates are denoted by '###'). Right: network composer.

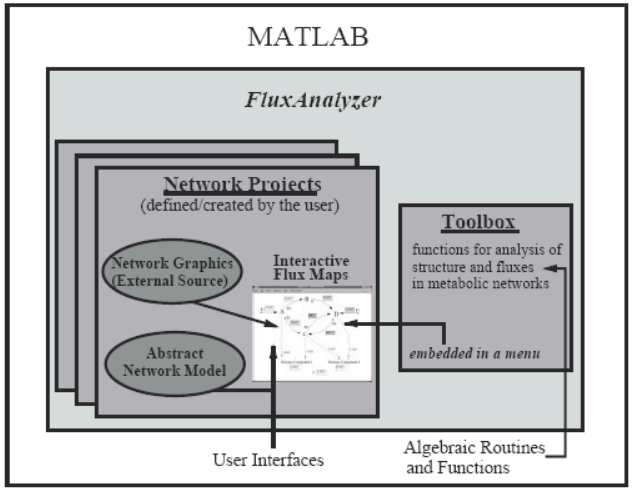


Fig. 2. Structural setup of the FluxAnalyzer.

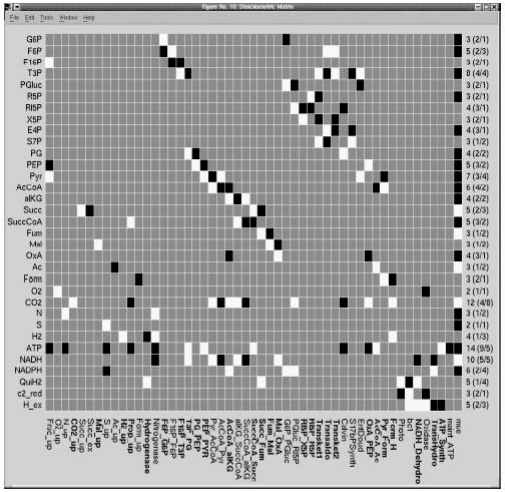


Fig. 3. Concise graphical representation of the stoichiometric matrix (here: catabolic part of the network studied in Klamt *et al.*, 2002)

`FluxAnalyzer` has both EPs and EFMs implemented.

Allows convenient studies of metabolicsystems.

Klamt et al.
Bioinformatics 19, 261 (2003)

Strain optimization based on EFM-analysis

Metabolic Engineering 12 (2010) 112–122

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid

Pornkamol Unrean, Cong T. Trinh, Friedrich Srienc*

Department of Chemical Engineering and Materials Science, and BioTechnology Institute, University of Minnesota, 240 Gortner Laboratory, 1479 Gortner Ave, St. Paul, MN 55108, USA

Carotenoids (e.g. DPL and DPA) are light-harvesting pigments, UV-protecting compounds, regulators of membrane fluidity, and antioxidants.

They are used as nutrient supplements, pharmaceuticals, and food colorants.

Aim: increase carotenoid synthesis in *E.coli*

Unrean et al. Metabol Eng 12, 112-122 (2010)

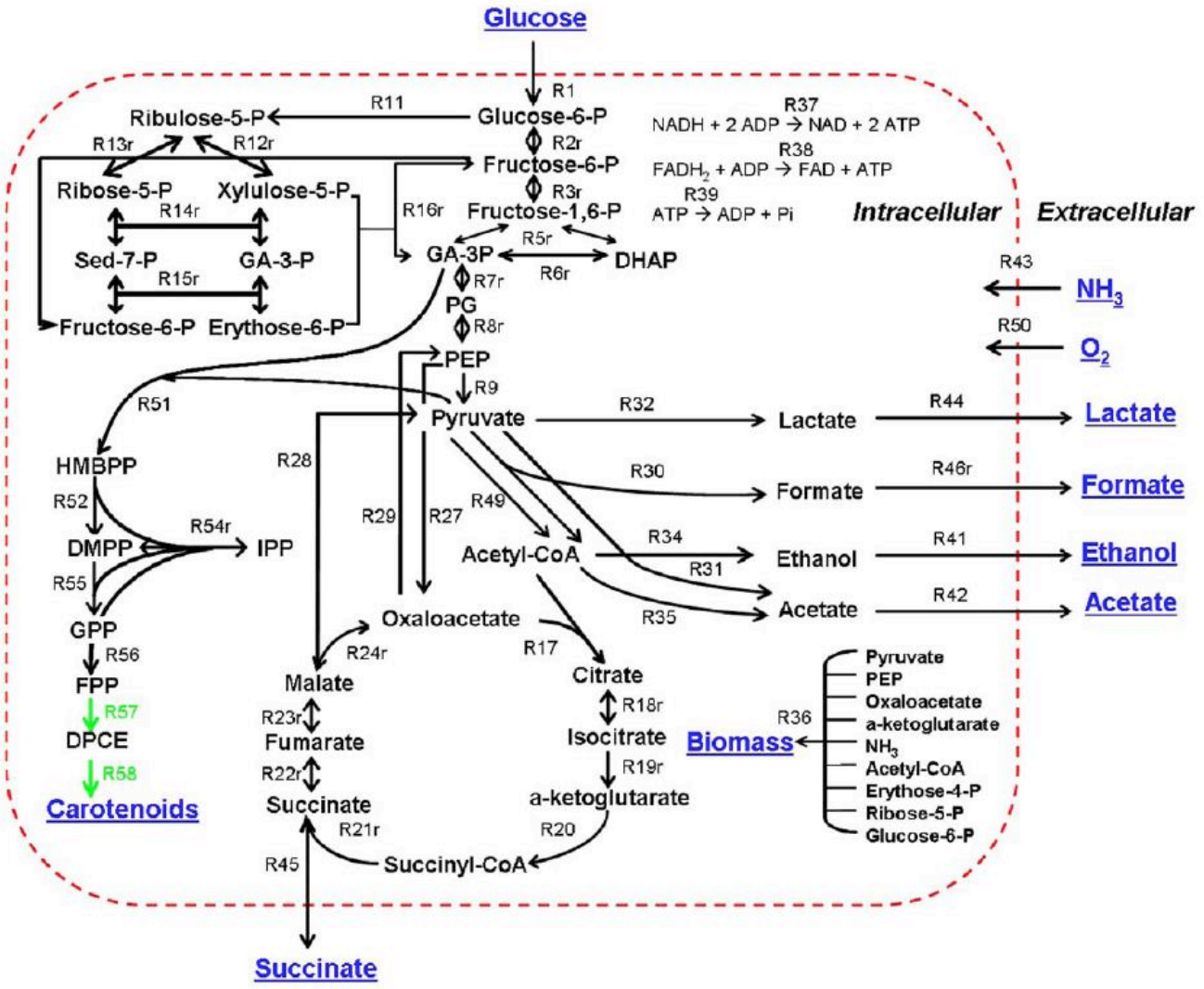
Metabolic network of recombinant *E.coli*

58 metabolic reactions
22 reversible
36 irreversible

57 metabolites

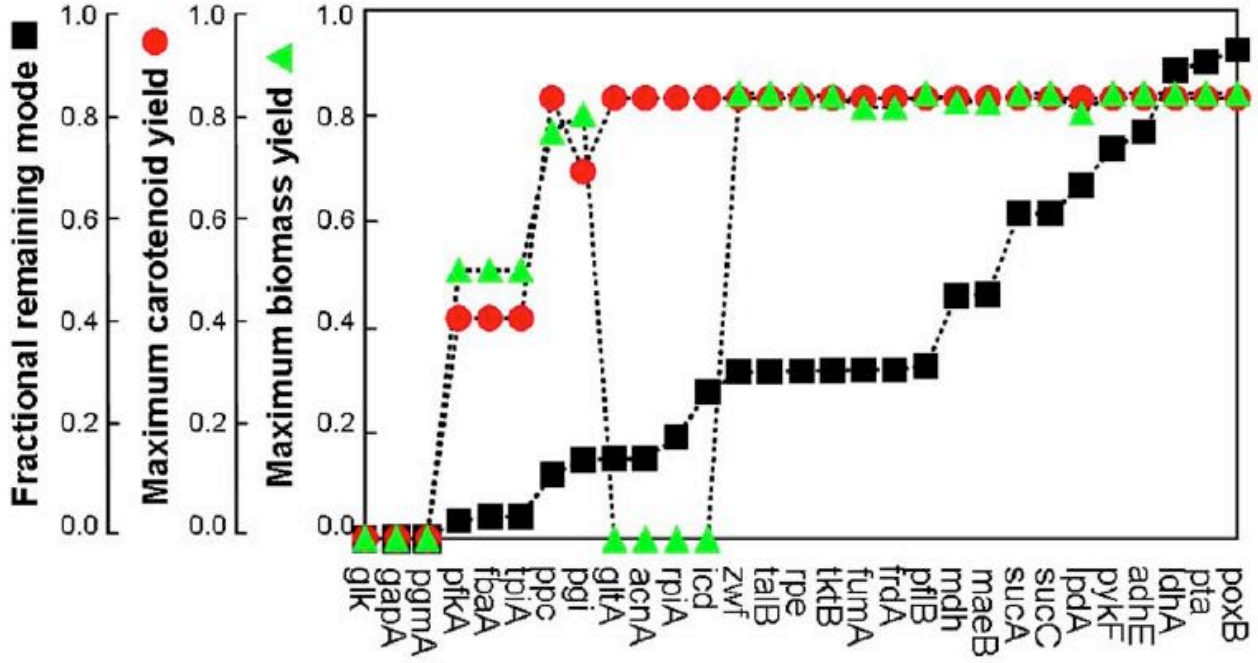
29532 EFMs

In 5923 EFMs, the production of biomass and DPA are coupled.



Unrean et al. Metabol Eng 12, 112-122 (2010)

Effect of single gene deletions



Results of virtual gene knockout calculations (counting number of EFMs and computing their yield from reaction stochiometries).

Select target genes where knockouts still maintain a maximum possible yield of carotenoid production, a reasonable yield of biomass while the largest number of EFMs is eliminated.

Unrean et al. Metabol Eng 12, 112-122 (2010)

Effect of single gene deletions

Strain	Total modes	Aerobic modes	Anaerobic modes	Predicted CRT yield[a]
Wild-type	29,532	24,155	5377	0.0-426
Δ*ldhA*	15,662	13,405	2257	0.0-426
Δ*ldhA*Δ*frdA*	8573	7810	763	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*	7541	6861	680	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*	6171	5600	571	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*	4099	4099	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*	2573	2573	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*	375	375	0	0.0-426
Δ*ldhA*Δ*frdA*Δ*poxB*Δ*pta*Δ*adhE*Δ*pykF*Δ*zwf*Δ*maeB*	5	5	0	0.4-426

[a] Yield is in mg-diapolycopendioic acid/g–glucose.

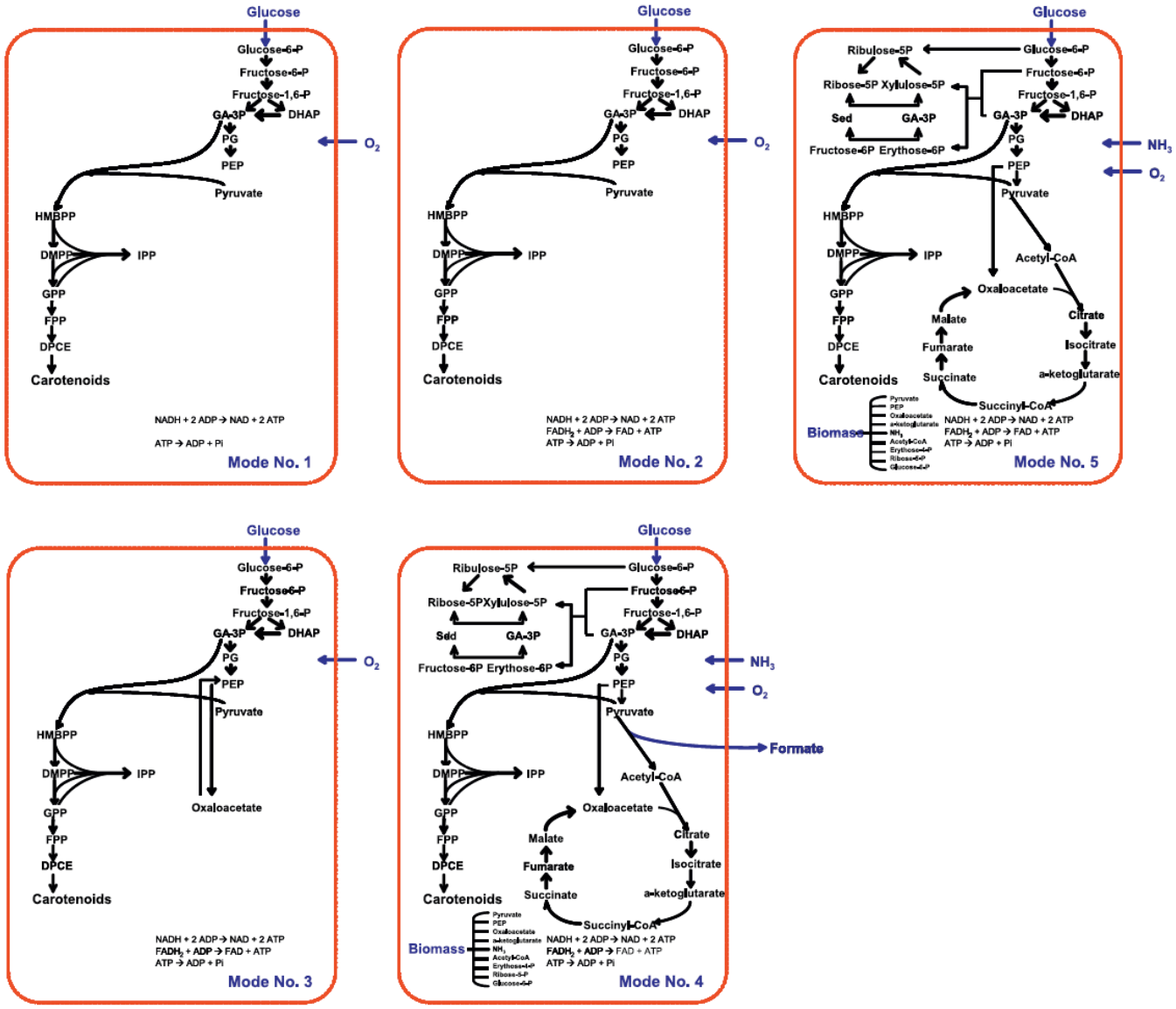
Deleted Reaction	Corresponding gene	Enzyme	Pathway
R9	*pykF*	Pyruvate kinase	Glycolysis
R11	*zwf*	Glucose-6-phosphate-1-dehydrogenase	Pentose phosphate
R22	*frdA*	Fumarate reductase	Fermentation
R28	*maeB*	Malate dehydrogenase	Anapleurotic
R31	*poxB*	Pyruvate oxidase	Fermentation
R32	*ldhA*	Lactate dehydrogenase	Fermentation
R34	*adhE*	Alcohol dehydrogenase	Fermentation
R35	*pta*	Phosphate acetyltransferase	Fermentation

Optimal: 8 gene knockouts lead to predicted over-production of DPL and DPA.

Only 5 EFMs remain.

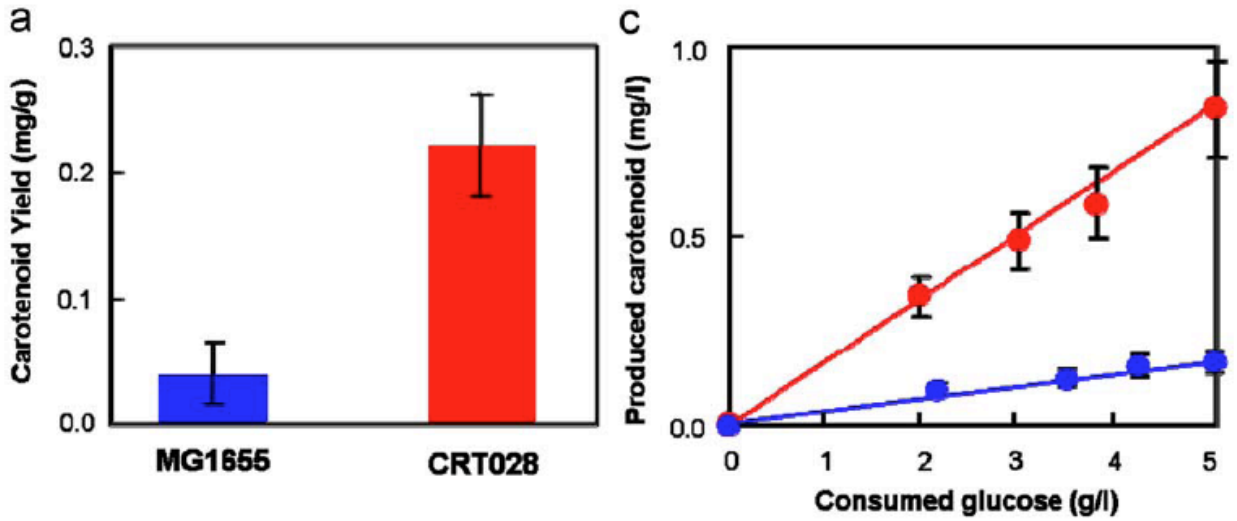
Unrean et al. Metabol Eng 12, 112-122 (2010)

Remaining EFMs



Unrean et al. Metabol Eng 12, 112-122 (2010)

Experimental verification: increased carotenoid yield



	MG1655/ pACMNOx	CRT028/ pACMNOx
Growth rate (/h)	0.17 ± 0.02	0.13 ± 0.01
Carotenoid production (mg/l)	0.19 ± 0.02	0.83 ± 0.20
Carotenoid yield (mg carotenoid/g glucose)	0.04 ± 0.00	0.17 ± 0.04
Specific production (mg carotenoid/g cell dry weight-h)	0.01 ± 0.00	0.10 ± 0.02

Mutant grows slower, but CRT production is increased 4 times.

Unrean et al. Metabol Eng 12, 112-122 (2010)

Complexity of finding and enumerating EFMs

Theorem: Given a stochiometric matrix *S*, an elementary mode can be found in polynomial time.

Theorem: In case all reactions in a metabolic network are reversible, the elementary modes can be enumerated in polynomial time.

The enumeration task becomes dramatically more difficult if the reactions are irreversible.
In this case, the modes of the network form a cone, and the elementary modes are the rays of the cone.

Theorem: Given a flux cone and two coordinates *i* and *j*, deciding if there exists and extreme ray of the cone that has both r_i and r_j in its support is NP-complete.

Theorem: Given a matrix *S* and a number *k*, deciding the existence of an elementary mode with at most *k* reactions in its support is NP-complete.

The question whether all elementary modes of a general network can be enumerated in polynomial time is an open question.

Acuna et al. BioSystems 99, 210-214 (2010); BioSystems 95, 51-60 (2009)

Summary EFMs

EFMs are a robust method that offers great opportunities for studying functional and structural properties in metabolic networks.

The **decomposition** of a particular flux distribution (e.g. determined by experiment) as a linear combination of EFMs is **not unique**.

Klamt & Stelling suggest that the term „elementary flux modes“ should be used whenever the sets of EFMs and EPs are identical.
In cases where they don‘t, EPs are a subset of EFMs.

It remains to be understood more thoroughly how much valuable information about the pathway structure is lost by using EPs.

Ongoing Challenges:
- study really large metabolic systems by subdividing them into sub-systems
- combine metabolic model with model of cellular regulation.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Minimal cut sets in biochemical reaction networks

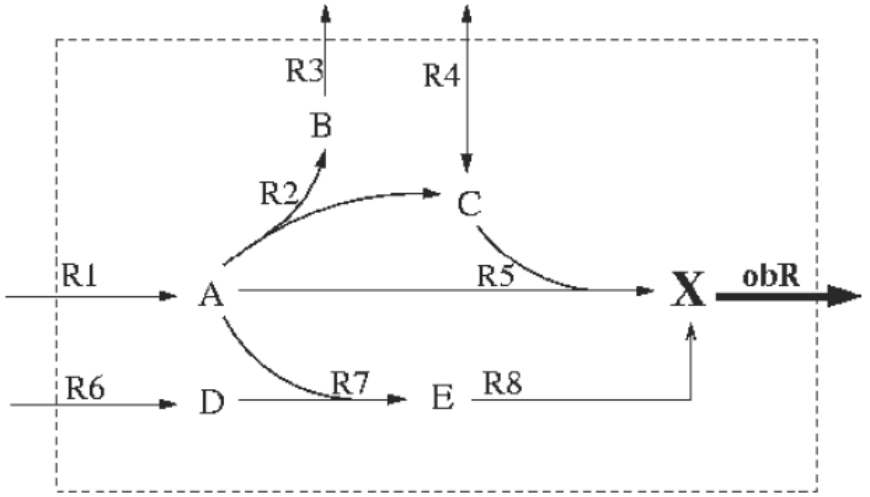
Concept of minimal cut sets (MCSs): smallest „failure modes“ in the network that render the correct functioning of a cellular reaction impossible.

Right: fictitious reaction network NetEx.

The only reversible reaction is R4.

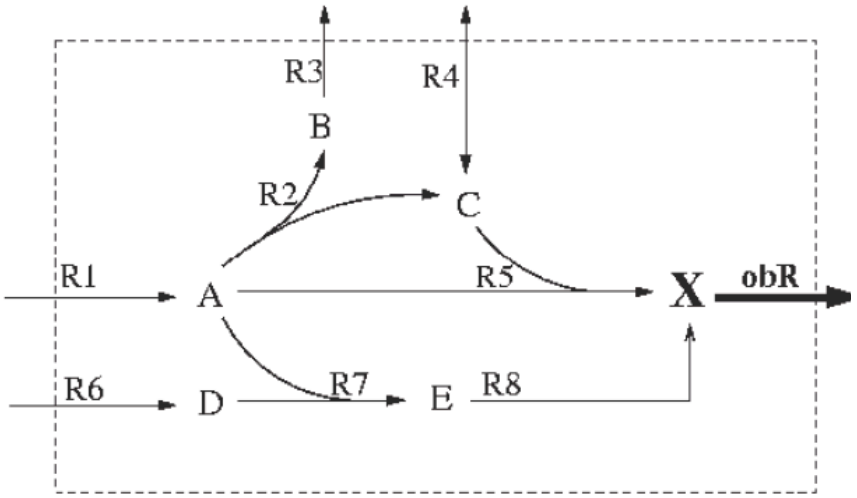
We are particularly interested in the flux `obR` exporting synthesized metabolite X.

→ Characterize solution space by computing elementary flux modes.



Klamt & Gilles, Bioinformatics 20, 226 (2004)

Elementary flux modes of NetEx



	R1	R2	R3	R4	R5	R6	R7	R8	obR
Elementary modes									
EM1	1	1	1	−1	0	0	0	0	0
EM2	1	0	0	0	0	1	1	1	1
EM3	2	1	1	0	1	0	0	0	1
EM4	1	0	0	1	1	0	0	0	1

One finds 4 elementary flux modes for NetEx.

3 of them (shaded) allow the production of metabolite X.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Cut set

Now we want to prevent the production of metabolite X.

→ demand that there is no balanced flux distribution possible which involves `obR`.

Definition. A set of reactions is termed a **cut set** (with respect to a defined objective reaction)
if after the removal of these reactions from the network no feasible balanced flux distribution involves the objective reaction.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Cut set

A trivial cut set is the reaction itself: C0 = {obR}.

Another extreme case is the removal of all reactions except obR .. not efficient!

Desirable solutions:

- From an engineering point of view, it might be desirable to cut reactions at the beginning of a pathway.

- The production of biomass is usually not coupled to a single gene or enzyme, and can therefore not be directly inactivated.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

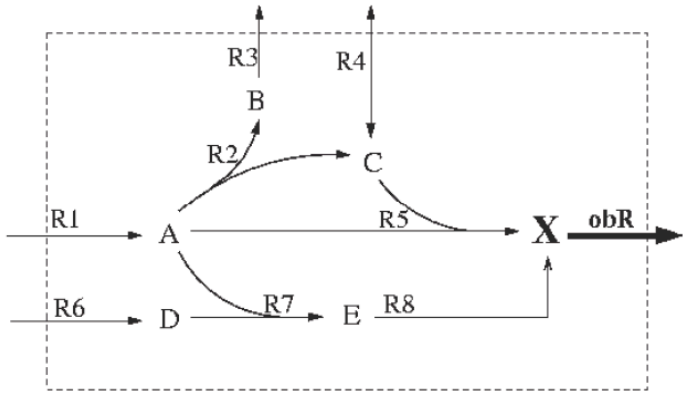
Cut set

C1 = {R5,R8} is a cut set already sufficient for preventing the production of X.

Removing R5 or R8 alone is not sufficient.

Definition. A cut set C (related to a defined objective reaction) is a **minimal cut set** (MCS) if no proper subset of C is a cut set.

→ C1 is a minimal cut set



	R1	R2	R3	R4	R5	R6	R7	R8	obR
Elementary modes									
EM1	1	1	1	−1	0	0	0	0	0
EM2	1	0	0	0	0	1	1	1	1
EM3	2	1	1	0	1	0	0	0	1
EM4	1	0	0	1	1	0	0	0	1
Minimal cut sets (objective reaction: obR)									
MCS0									×
MCS1	×								
MCS2					×	×			
MCS3					×		×		
MCS4					×			×	
MCS5		×		×		×			
MCS6			×	×		×			
MCS7		×		×			×		
MCS8			×	×			×		
MCS9		×		×				×	
MCS10			×	×				×	

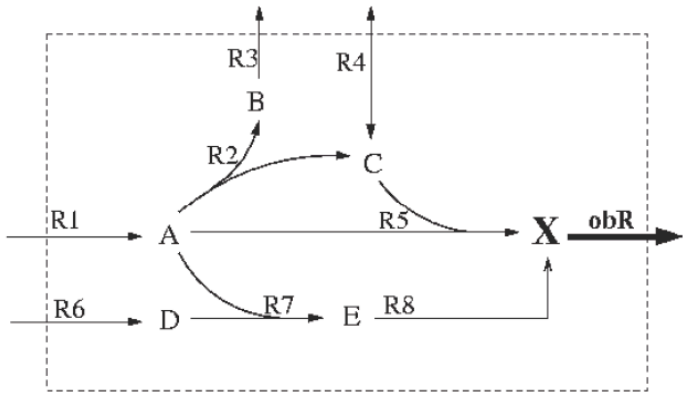
Klamt & Gilles, Bioinformatics 20, 226 (2004)

Remarks

(1) An MCS always guarantees dysfunction as long as the assumed network structure is currect. However, additional regulatory circuits or capacity restrictions may allow that even a proper subset of a MCS is a cut set.
The MCS analysis should always be seen from a purely structural point of view.

(2) After removing a complete MCS from the network, other pathways producing other metabolites may still be active.

(3) MCS4 = {R5,R8} clearly stops production of X.



What about MCS6 = {R3,R4,R6}?

Cannot X be still be produced via R1, R2, and R5?
However, this would lead to accumulation of B and is therefore physiologically impossible.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Similar concepts

Graph theory:
we previously introduced a similar definition of minimal cut sets where they ensure a disconnectivity of a given graph.

However, these graph-theoretical concepts do not fit into the definition of MCSs as defined here and would, in general, lead to other results!

The reason is that metabolic networks use an explicit consideration of the hypergraphical nature of metabolic networks.

Hypergraphs: generalized graphs, where an edge (reaction) can link *k* nodes (reactants) with *l* nodes (products), whereas in graphs only 1:1 relations are allowed.

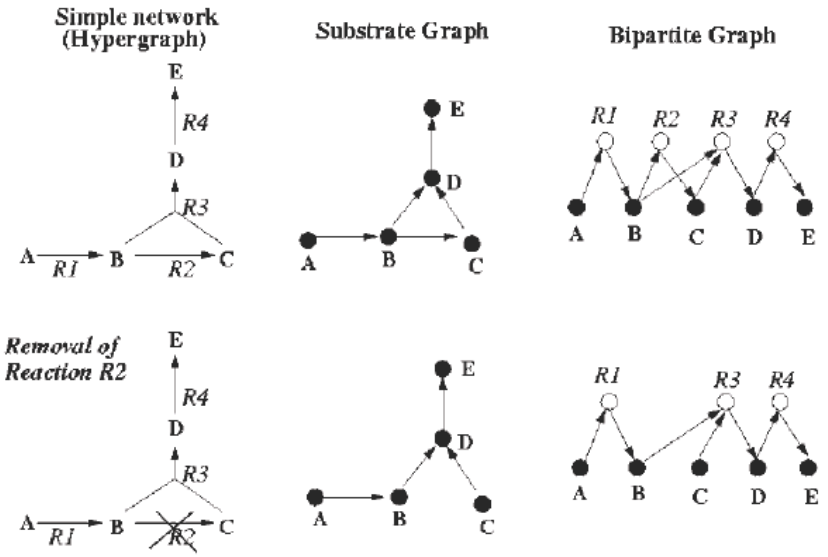
Klamt & Gilles, Bioinformatics 20, 226 (2004)

Comparison with graph theory

Example: we are interested in inhibiting the production of E.
Thus, R4 is our objective reaction.

If R2 is removed from the network, E can no longer be produced because C is required for driving reaction R3.

However, R2 would not be an MCS in terms of graph theory, neither in the substrate or in the bipartite graph representation because all metabolites are still connected after R2 is removed.



Klamt & Gilles, Bioinformatics 20, 226 (2004)

Algorithm for computing MCSs

The MCSs for a given network and objective reaction are members of the power set of the set of reaction indices and are uniquely determined.

A systematic computation must ensure that the calculated MCSs are:

(1) cut sets („destroying“ all possible balanced flux distributions involving the objective reaction), and

(2) that the MCSs are really minimal, and

(3) that all MCSs are found.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Algorithm for computing MCSs

(1) cut sets („destroying“ all possible balanced flux distributions involving the objective reaction),

→ any feasible steady-state flux distribution in a given network – expressed as vector **r** of the *q* net reaction rates – can be represented by a non-negative linear combination of the *N* elementary modes:

$$\mathrm{r} = \sum_{i=1}^{N} \alpha_i EM_i, \qquad \alpha_i \geq 0$$

To ensure that the rate r_k of the objective reaction is 0 in all **r**, each EM must contain 0 at the *k*-th place.

→ If C is a proper cut set the following cut set condition must hold:
For each EM involving the objective reaction (with a non-zero value),
there is at least one reaction in C also involved in this EM.

This guarantees that all EMs, in which the objective reaction participates,
will vanish when the reactions in the cut set are removed from the network.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Algorithm

ALGORITHM:

(1) Calculate the EMs in the given network

(2) Define the objective reaction obR

(3) Choose all EMs where reaction obR is non-zero and store it in the binary array *em_obR* (*em_obR*[*i*][*j*]==1 means that reaction *j* is involved in EM *i*)

(4) Initialize arrays *mcs* and *precutsets* as follows (each array contains sets of reaction indices): append {*j*} to *mcs* if reaction *j* is essential (*em_obR*[*i*][*j*]=1 for each EM *i*), otherwise to *precutsets*

(5) FOR i=2 TO MAX_CUTSETSIZE

(5.1) *new_precutsets*=[ ];

(5.2) FOR *j* = 1 TO *q* (*q*: number of reactions)

(5.2.1) Remove all sets from *precutsets* where reaction *j* participates

(5.2.2) Find all sets of reactions in *precutsets* that do not cover at least one EM in *em_obR* where reaction *j* participates; combine each of these sets with reaction *j* and store the new preliminary cut sets in *temp_precutsets*

(5.2.3) Drop all *temp_precutsets* which are a superset of any of the already determined minimal cut sets stored in *mcs*

(5.2.4) Find all retained *temp_precutsets* which do now cover all EMs and append them to *mcs*; append all others to *new_precutsets*

ENDFOR

(5.3) If isempty(*new_precutsets*)

(5.3.1) Break

ELSE

(5.3.2) *precutsets*=*new_precutsets*

ENDIF

ENDFOR

(6) result: *mcs* contains the MCSs

Klamt & Gilles, Bioinformatics 20, 226 (2004)

According to Acuna (2009) this algorithm is often very inefficient.

Applications of MCSs

Target identification and repression of cellular functions

A screening of all MCSs allows for the identification of the best suitable manipulation. For practical reasons, the following conditions should be fulfilled:

- usually, a small number of interventions is desirable (small size of MCS)

- other pathways in the network should only be weakly affected

- some of the cellular functions might be difficult to shut down genetically or by inhibition, e.g. if many isozymes exist for a reaction.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Applications of MCSs

Network verification and mutant phenotype predictions

We expect that cutting away an MCS from the network is definitely intolerable for the cell with respect to certain cellular reactions/processes.

Such predictions, derived purely from network structure, are a useful strategy for verification of hypothetical or reconstructed networks.

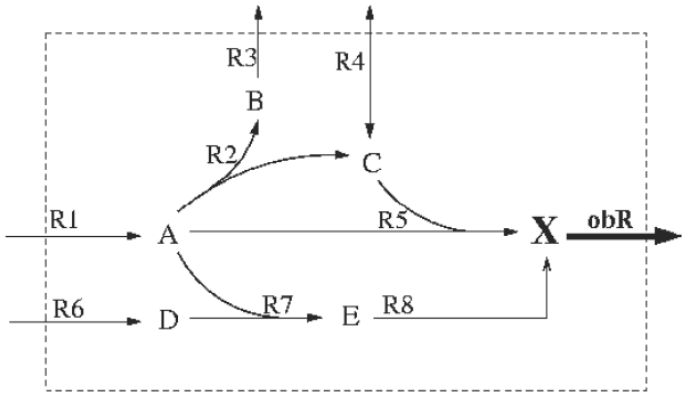
If the outcome of prediction and experiments differ, this often indicates an incorrect or incomplete network structure.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Structural fragility and robustness

If we assume that each reaction in a metabolic network has the same probability to fail, small MCSs are most probable to be responsible for a failing objective function.

Define a **fragility coefficient** F_i as the reciprocal of the average size of all MCSs in which reaction *i* participates.



	R1	R2	R3	R4	R5	R6	R7	R8	obR
Elementary modes									
EM1	1	1	1	−1	0	0	0	0	0
EM2	1	0	0	0	0	1	1	1	1
EM3	2	1	1	0	1	0	0	0	1
EM4	1	0	0	1	1	0	0	0	1
Minimal cut sets (objective reaction: obR)									
MCS0									×
MCS1	×								
MCS2					×	×			
MCS3					×		×		
MCS4					×			×	
MCS5		×		×		×			
MCS6			×	×		×			
MCS7		×		×			×		
MCS8			×	×			×		
MCS9		×		×				×	
MCS10			×	×				×	
F_i	1	1/3	1/3	1/3	1/2	3/8	3/8	3/8	1

Besides the essential reaction R1, reaction R5 is most crucial for the objective reaction.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Example: MCSs in the central metabolism of *E.coli*

objective reaction
„biomass synthesis“
Network: 110 reactions,
89 metabolites,
see Stelling et al. (2002)

Table 2. Overview on computed MCSs in the central metabolism of *E.coli* for growth on four different substrates

	Acetate	Succinate	Glycerol	Glucose
No. of EMs with growth	363	3421	9479	21 592
No. of MCSs (objective reaction: growth)	245	1255	2970	4225
Maximal number of preliminary MCSs (during computation)	3563	69 628	344 196	902 769
Computation time (Intel Pentium, 1 MHZ; 4 GB RAM)	7 s	20 min	5.42 h	29.67 h
F_i values (in parentheses: size of the smallest MCS in which the reaction occurs)				
F16P-bisphosphatase	1 (1)	1 (1)	1 (1)	0.102 (6)
ATP-synthase	1 (1)	0.325 (3)	0.141 (3)	0.149 (3)
SuccCoA-synthetase	0.207 (2)	0.145 (2)	0.125 (2)	0.131 (2)
PEP-carboxylase	0.128 (2)	0.117 (2)	0.120 (2)	0.143 (2)
Malic enzyme	0.5 (2)	0.5 (2)	0.114 (2)	0.123 (2)
R15P-X5P (epimerase)	0.198 (2)	0.135 (2)	0.128 (2)	0.148 (2)
F	0.783	0.718	0.699	0.643

The computation time does not involve the time needed for computing the elementary modes. F_i: fragility coefficient of reaction i; **F**: network (overall) fragility coefficient.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Conclusion

An MCS is a irreducible combination of network elements whose simultaneous inactivation leads to a guaranteed dysfunction of certain cellular reactions or processes.

Theorem: Determining a reaction cut of minimum cardinality is NP-hard.

MCSs are inherent and uniquely determined structural features of metabolic networks similar to EMs.

The computation of MCSs and EMs becomes challenging in large networks.

Analyzing the MCSs gives deeper insights in the structural fragility of a given metabolic network and is useful for identifying target sets for an intended repression of network functions.

Klamt & Gilles, Bioinformatics 20, 226 (2004)
Acuna et al. BioSystems 95, 51-60 (2009)