V20 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.
Two such approaches are the **elementary flux modes** (Schuster et al. 1999, 2000) and **extreme pathways** (Schilling et al. 2000).

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.

MPA can be used to study e.g.

- routing + flexibility/redundancy of networks
- functionality of networks
- idenfication of futile cycles
- gives all (sub)optimal pathways with respect to product/biomass yield
- can be useful for calculability studies in MFA

Klamt et al. Bioinformatics 19, 261 (2003)

Metabolic Pathway Analysis: Elementary Flux Modes

The technique of **Elementary Flux Modes (EFM)** was developed prior to extreme pathways (**EP**) by Stephan Schuster, Thomas Dandekar and co-workers:

Pfeiffer et al. Bioinformatics, 15, 251 (1999)

Schuster et al. Nature Biotech. 18, 326 (2000)

The method is very similar to the „extreme pathway“ method to construct a basis for metabolic flux states based on methods from convex algebra.

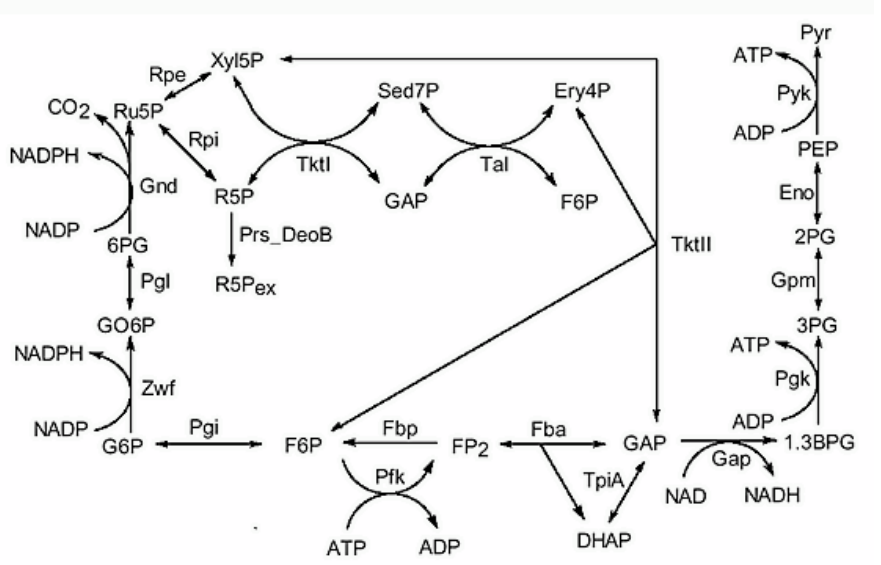
Extreme pathways are a subset of elementary modes, and for many systems, both methods coincide.

Are the subtle differences important?

Elementary Flux Modes

Start from list of reaction equations and a declaration of reversible and irreversible reactions and of internal and external metabolites.

E.g. reaction scheme of monosaccharide metabolism. It includes 15 internal metabolites, and 19 reactions.
→ **S** has dimension 15 × 19.



It is convenient to reduce this matrix by lumping those reactions that necessarily operate together.
→ {Gap,Pgk,Gpm,Eno,Pyk},
→ {Zwf,Pgl,Gnd}

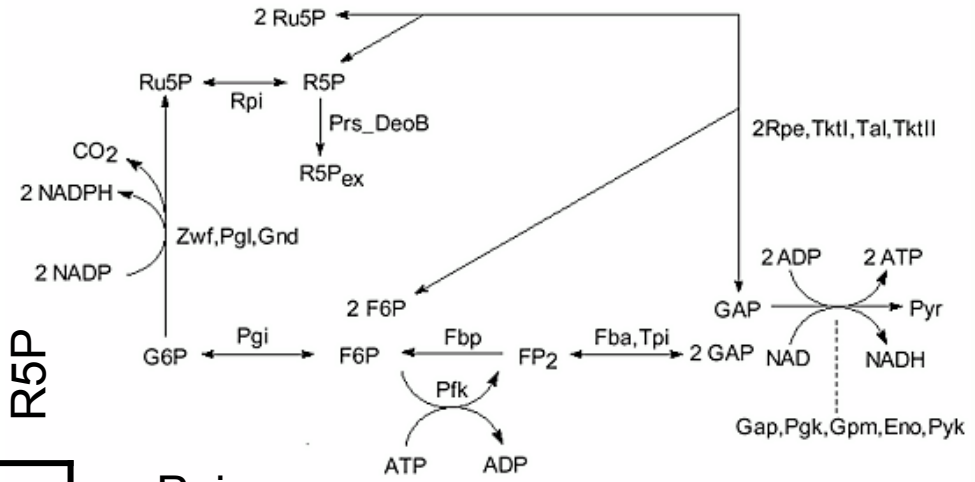
Such groups of enzymes can be detected automatically.
This reveals another two sequences {Fba,TpiA} and {2 Rpe,TktI,Tal,TktII}.

Schuster et al. Nature Biotech 18, 326 (2000)

Elementary Flux Modes

Lumping the reactions in any one sequence gives the following reduced system:

Construct initial tableau by combining **S** with identity matrix:



	Ru5P	FP_2	F6P	GAP	R5P
	1	0	...	0	0	0	1	0	0
	0	1	...	0	0	-1	0	2	0
	0	0	...	0	-1	0	0	0	1
	0	0	...	0	-2	0	2	1	-1
$T^{(0)}$=	0	0	...	0	0	0	0	-1	0
	0	0	...	0	1	0	0	0	0
	0	0	...	0	0	1	-1	0	0
	0	0	...	0	0	-1	1	0	0
	0	0	...	1	0	0	0	0	-1

Pgi
{Fba,TpiA}
Rpi reversible
{2Rpe,TktI,Tal,TktII}
{Gap,Pgk,Gpm,Eno,Pyk}
{Zwf,Pgl,Gnd}
Pfk irreversible
Fbp
Prs_DeoB

Schuster et al. Nature Biotech 18, 326 (2000)

Elementary Flux Modes

Aim again: bring all entries of right part of matrix to 0.

E.g. 2*row3 - row4 gives „reversible“ row with 0 in column 10

New „irreversible“ rows with 0 entry in column 10 by row3 + row6 and by row4 + row7.

In general, linear combinations of 2 rows corresponding to the same type of directionality go into the part of the respective type in the tableau. Combinations by different types go into the „irreversible“ tableau because at least 1 reaction is irreversible. Irreversible reactions can only combined using positive coefficients.

$T^{(0)}=$

1									0	0	1	0	0
	1								0	-1	0	2	0
		1							-1	0	0	0	1
			1						-2	0	2	1	-1
				1					0	0	0	-1	0
					1				1	0	0	0	0
						1			0	1	-1	0	0
							1		0	-1	1	0	0
								1	0	0	0	0	-1

$T^{(1)}=$

1									0	0	1	0	0
	1								0	-1	0	2	0
		2	-1						0	0	-2	-1	3
				1					0	0	0	-1	0
						1			0	1	-1	0	0
							1		0	-1	1	0	0
								1	0	0	0	0	-1
		1			1				0	0	0	0	1
			1		2				0	0	2	1	-1

Schuster et al. Nature Biotech 18, 326 (2000)

Elementary Flux Modes

Aim: zero column 11.

Include all possible (direction-wise allowed) linear combinations of rows.

$T^{(1)}=$

1									0	0	1	0	0
	1								0	-1	0	2	0
		2	-1						0	0	-2	-1	3
				1					0	0	0	-1	0
						1			0	1	-1	0	0
							1		0	-1	1	0	0
								1	0	0	0	0	-1
		1			1				0	0	0	0	1
			1		2				0	0	2	1	-1

$T^{(2)}=$

1									0	0	1	0	0
		2	-1						0	0	-2	-1	3
				1					0	0	0	-1	0
								1	0	0	0	0	-1
		1			1				0	0	0	0	1
			1		2				0	0	2	1	-1
	1					1			0	0	-1	2	0
	-1						1		0	0	1	-2	0
						1	1		0	0	0	0	0

continue with columns 12-14.

Schuster et al. Nature Biotech 18, 326 (2000)

Elementary Flux Modes

In the course of the algorithm, one must avoid

- calculation of nonelementary modes (rows that contain fewer zeros than the row already present)
- duplicate modes (a pair of rows is only combined if it fulfills the condition $S(\mathbf{m}_i^{(j)}) \cap S(\mathbf{m}_k^{(j)}) \not\subseteq S(\mathbf{m}_l^{(j+1)})$ where $S(\mathbf{m}_l^{(j+1)})$ is the set of positions of 0 in this row.

- flux modes violating the sign restriction for the irreversible reactions.

Final tableau

$\mathbf{T^{(5)}} =$

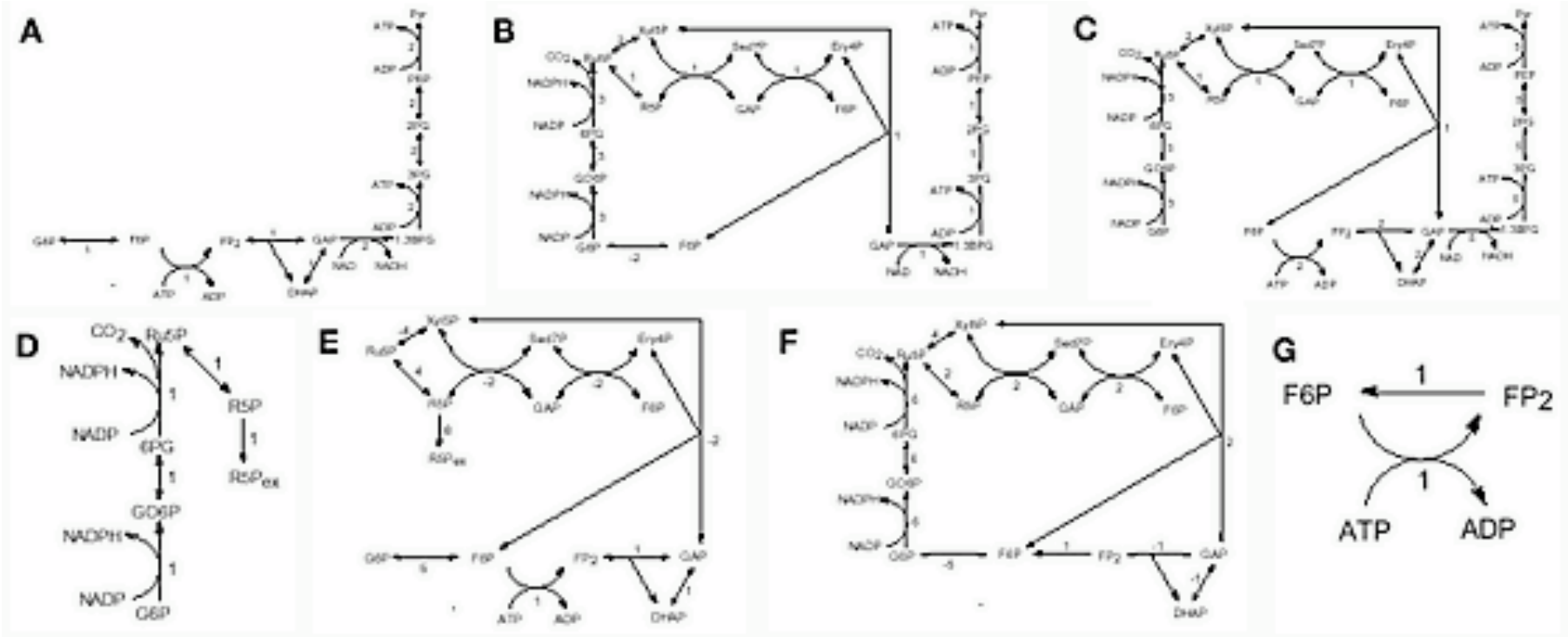
1	1	0	0	2	0	1	0	0	0	...		...	0
-2	0	1	1	1	3	0	0	0	...				...
0	2	1	1	5	3	2	0	0					
0	0	1	0	0	1	0	0	1					
5	1	4	-2	0	0	1	0	6					
-5	-1	2	2	0	6	0	1	0	...				...
0	0	0	0	0	0	1	1	0	0	...		...	0

This shows that the number of rows may decrease or increase in the course of the algorithm. All constructed elementary modes are irreversible.

Schuster et al. Nature Biotech 18, 326 (2000)

Elementary Flux Modes

Graphical representation of the elementary flux modes of the monosaccharide metabolism. The numbers indicate the relative flux carried by the enzymes.



Schuster et al. Nature Biotech 18, 326 (2000)

Two approaches for Metabolic Pathway Analysis?

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** (unequal to the zero vector and to **e**) fulfilling C1 and C2 and that P(**v**) is a proper subset of P(**e**). This is the core characteristics for EFMs and EPs and supplies the decomposition of the network into smallest units (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)

Two approaches for Metabolic Pathway Analysis?

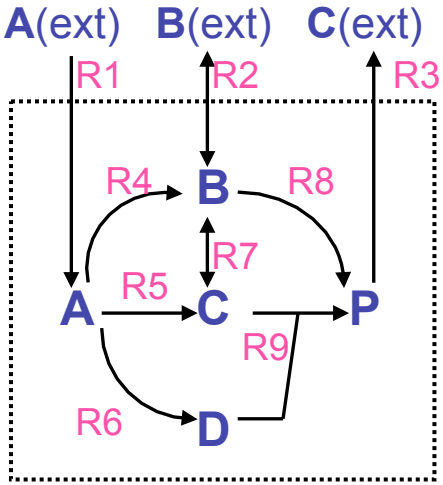
The pathway *P(**e**)* is an **extreme pathway** if it fulfills conditions C1 – C3 AND conditions C4 – C5.

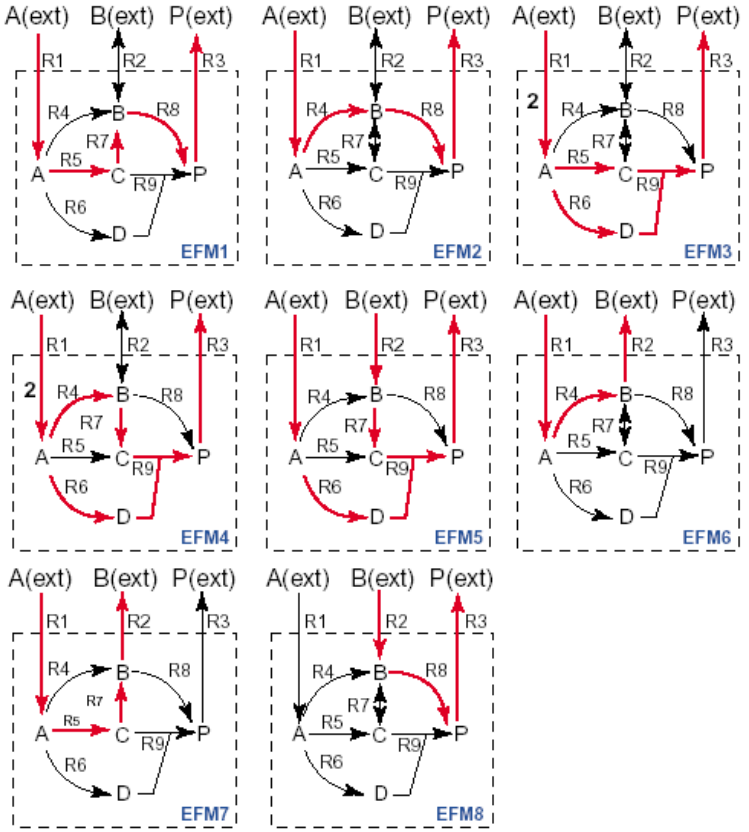
(C4) *Network reconfiguration*: Each reaction must be classified either as exchange flux or as internal reaction. All reversible internal reactions must be split up into two separate, irreversible reactions (forward and backward reaction).

(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.

Klamt & Stelling Trends Biotech 21, 64 (2003)

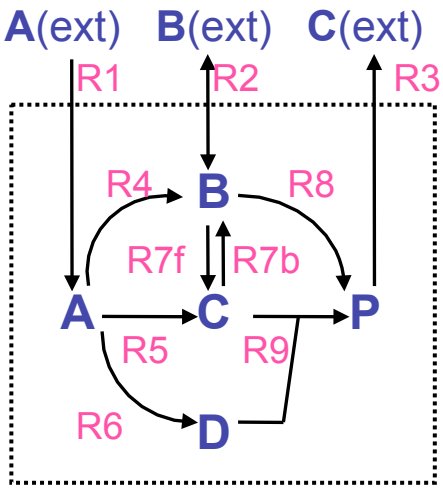
Two approaches for Metabolic Pathway Analysis?

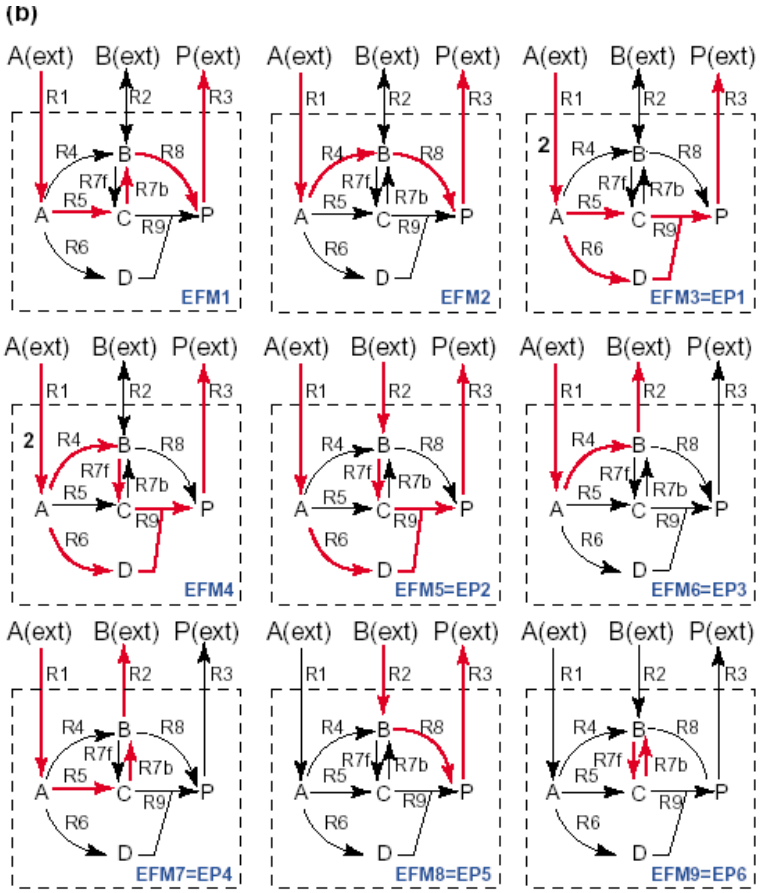




Klamt & Stelling Trends Biotech 21, 64 (2003)

Reconfigured Network





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

Klamt & Stelling Trends Biotech 21, 64 (2003)

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.

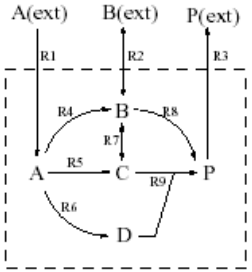
Valid for any network: Property 1
Reconfiguring a network by splitting up reversible reactions leads to the same set of meaningful EFMs.

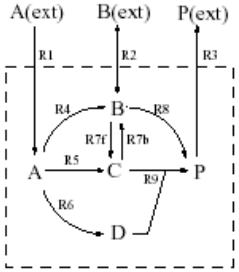
Klamt & Stelling Trends Biotech 21, 64 (2003)

EFMs vs. EPs

What is the consequence of when all exchange fluxes (and hence all reactions in the network) are 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		Reactions								
	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		Reactions									
	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!

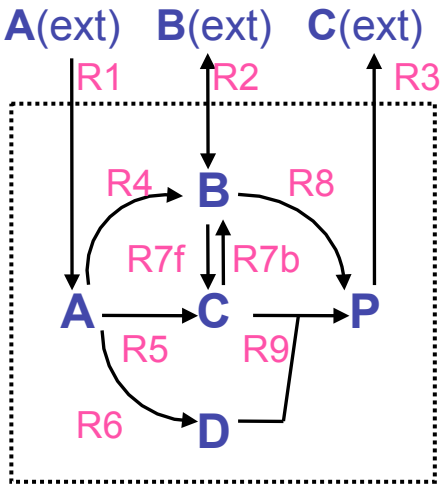
Klamt & Stelling Trends Biotech 21, 64 (2003)

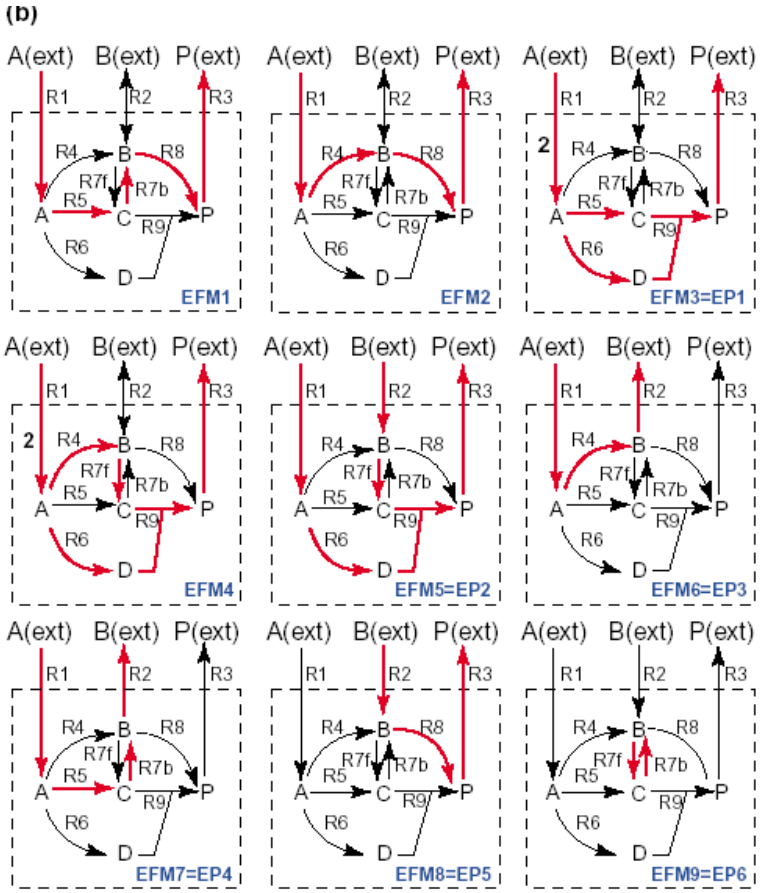
Property 2 of EFMs

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)

Comparison of EFMs and EPs

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. Searching for EPs leading from A to P via B, no pathway would be found.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Comparison of EFMs and EPs

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.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Comparison of EFMs and EPs

Problem	EFM (network N1)	EP (network N2)
Analysis of network flexibility (structural robustness, redundancy): relative robustness of exclusive growth on A or B.	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, for production of P from A 2 pathways will always „survive“. By contrast, removing reaction R8 already stops the production of P from B alone.	Only 1 EP exists for producing P by substrate A alone, and 1 EP for synthesizing P by (only) substrate B. One might suggest 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)

Comparison of EFMs and EPs

Problem	EFM (network N1)	EP (network N2)
Relative importance of single reactions: relative importance of reaction R8.	R8 is essential for producing P by substrate B, 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.	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).

Klamt & Stelling Trends Biotech 21, 64 (2003)

Comparison of EFMs and EPs

Problem	EFM (network N1)	EP (network N2)
Enzyme subsets and excluding reaction pairs: suggest regulatory structures or rules.	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.)	The EPs pretend R4 and R8 to be an excluding reaction pair – but they are not (EFM2). The enzyme subsets would be correctly identified. However, one can construct simple examples where the EPs would also pretend wrong enzyme subsets (not shown).

Klamt & Stelling Trends Biotech 21, 64 (2003)

Comparison of EFMs and EPs

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)

Comparison of EFMs and EPs

Problem	EFM (network N1)	EP (network N2)
Removing a reaction and mutation studies: effect of deleting R7.	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.	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)

Comparison of EFMs and EPs

Problem	EFM (network N1)	EP (network N2)
Constraining reaction reversibility: effect of R7 limited to B → C.	For the case of R7, all EFMs but EFM1 and EFM7 „survive“ because the latter ones utilize R7 with negative rate.	In general, the set of EPs must be recalculated: compare the EPs in network N2 (R2 reversible) and N4 (R2 irreversible).

Klamt & Stelling Trends Biotech 21, 64 (2003)

Software: FluxAnalyzer, based on Matlab

Steffen Klamt.

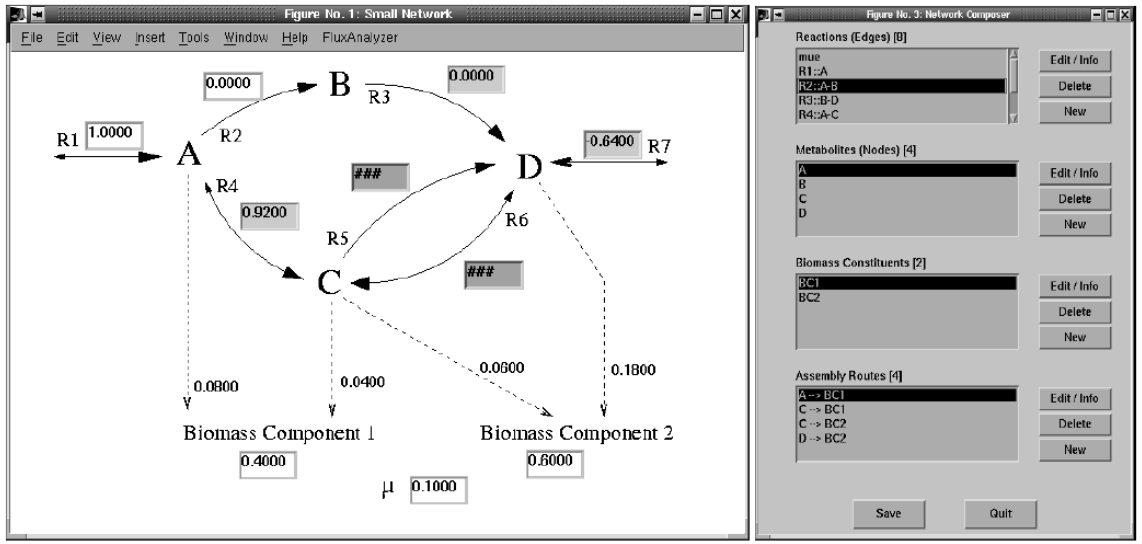


Fig. 1. The network project of 'SMALLNET' constructed by the FluxAnalyzer. Left: interactive flux map displaying a flux scenario (unknown rates are denoted by '###'). Right: network composer.

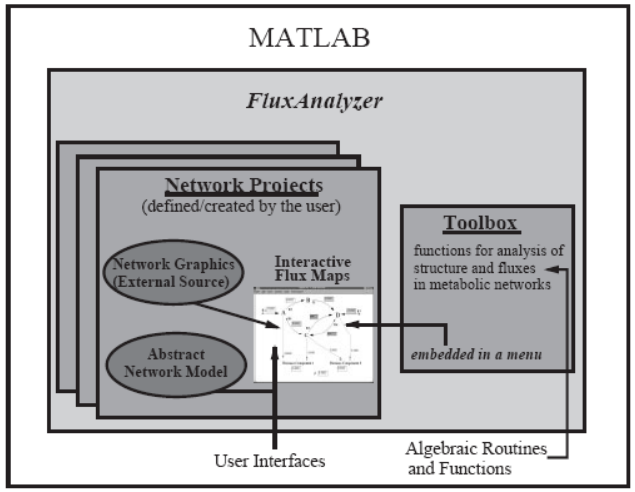


Fig. 2. Structural setup of the FluxAnalyzer.

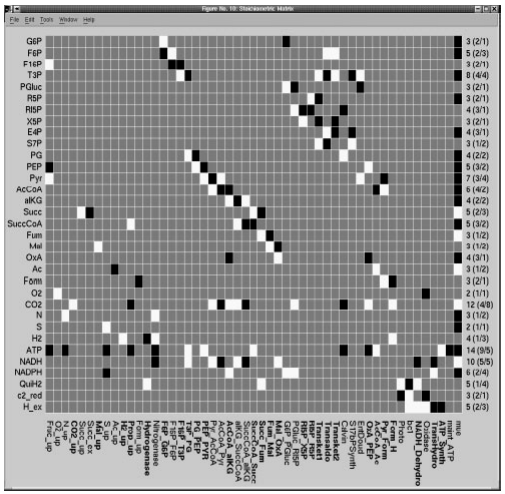


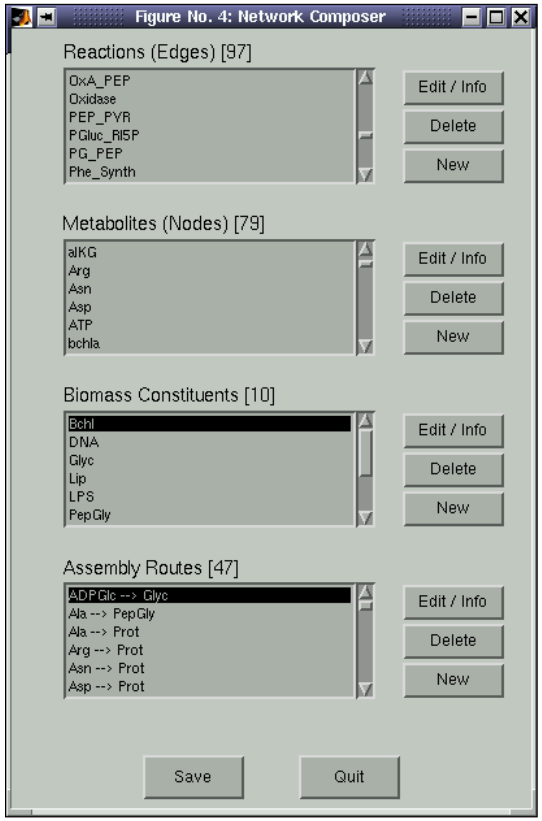
Fig. 3. Concise graphical representation of the stoichiometric matrix (here: catabolic part of the network studied in Klamt *et al.*, 2002)

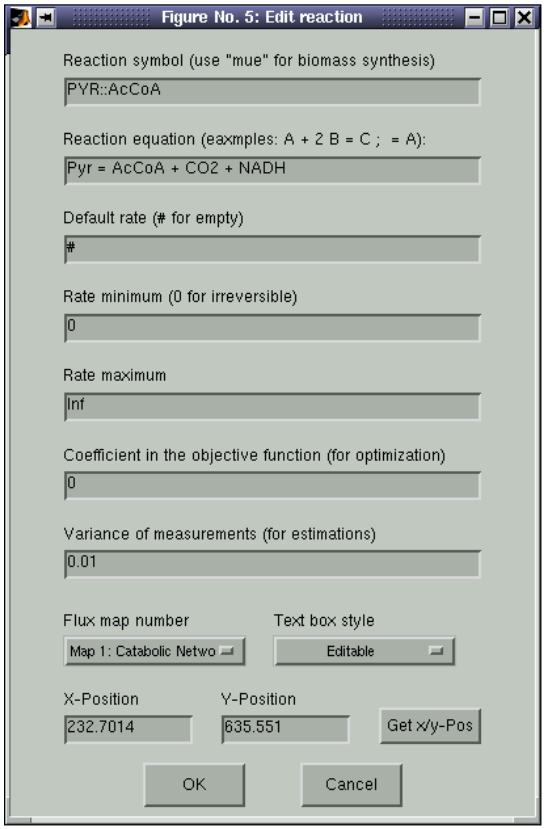
`FluxAnalyzer` has both EPs and EFMs implemented.

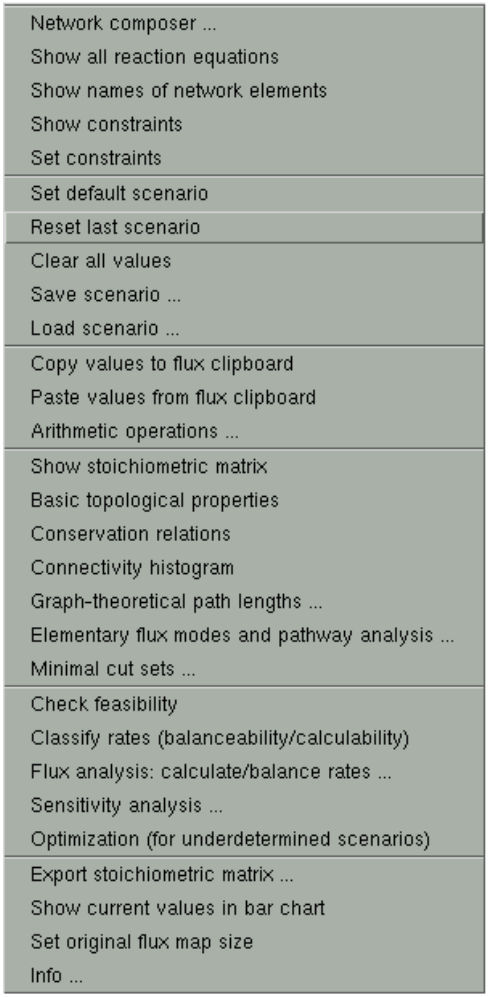
Allows convenient studies of metabolicsystems.

Klamt et al.
Bioinformatics 19, 261 (2003)

Software: FluxAnalyzer, based on Matlab







Left: Network composer of the *FluxAnalyzer* facilitating the definition of the network structure.

Middle: Input mask for defining a new network element of type *reaction*

Right: Pull-down menu of the *FluxAnalyzer* providing an interactive use of the various functions of the toolbox.

http://pinguin.biologie.uni-jena.de/bioinformatik/

Summary EM

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

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
- combine metabolic model with model of cellular regulation.

Klamt & Stelling Trends Biotech 21, 64 (2003)

Minimal cut sets in biochemical reaction networks

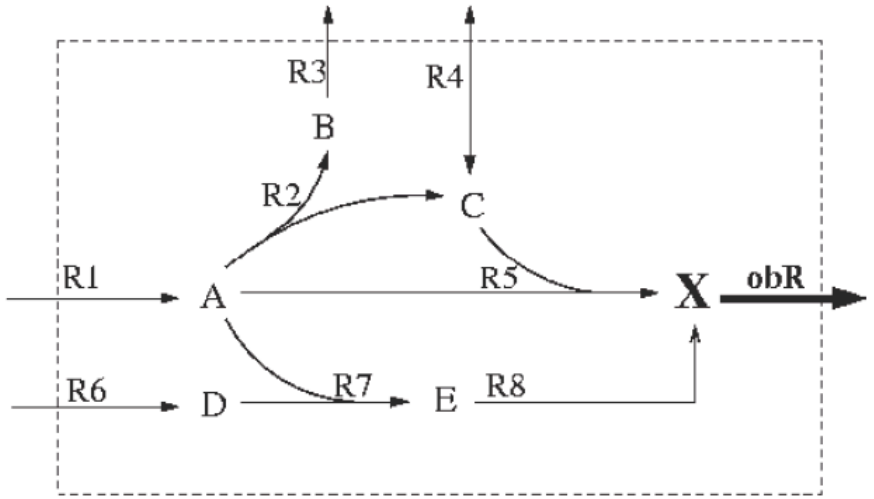
Concept of minimal cut sets (MCSs): smallest „failure modes“ in the network that render the correct functioning of a cellular reaction impossible.

Right: fictitious reaction network NetEx.

The only reversible reaction is R4.

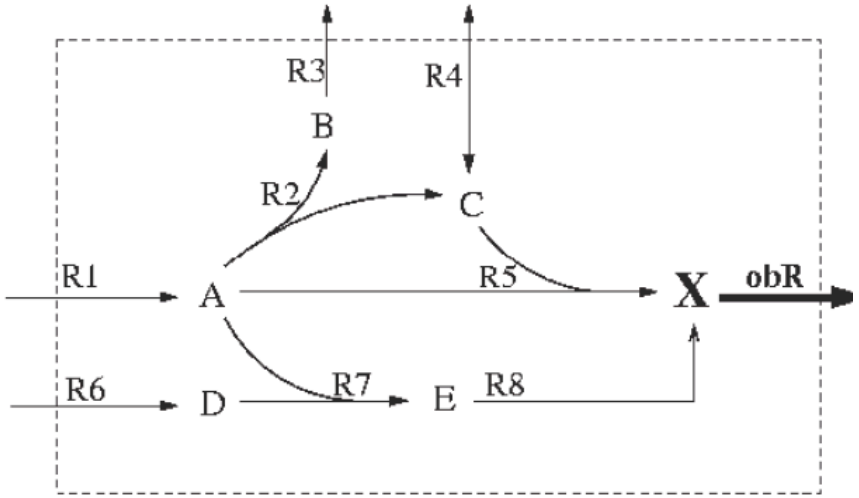
We are particularly interested in the flux `obR` exporting synthesized metabolite X.

→ Characterize solution space by computing elementary modes.



Klamt & Gilles, Bioinformatics 20, 226 (2004)

Elementary 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 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. We call a set of reactions 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.

A trivial cut set if the reaction itself: C0 = {`obR`}. Why should we be interested in other solutions as well?

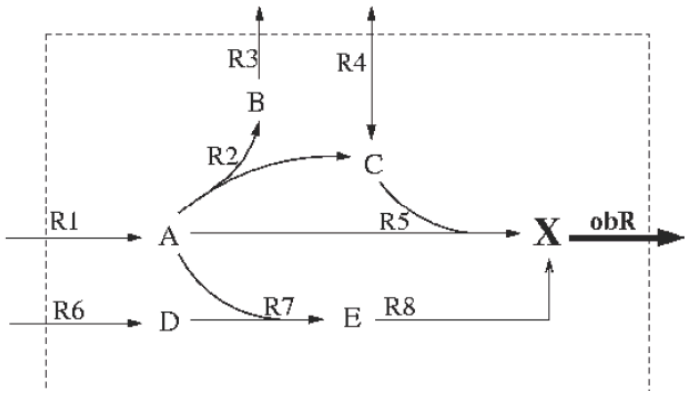
- From an engineering point of view, it might be desirable to cut reactions at the beginning of a pathway.
- The production of biomass is usually not coupled to a single gene or enzyme, and can therefore not be directly inactivated.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Cut set

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

E.g. C1 = {R5,R8} is a cut set already sufficient for preventing the production of X. Removing R5 or R8 alone is not sufficient.
→ C1 is a minimal cut set



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.

	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

Risk assessment:
a very similar definition of MCSs exists for „fault trees“ studied in reliability and risk assessment of industrial systems.

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.

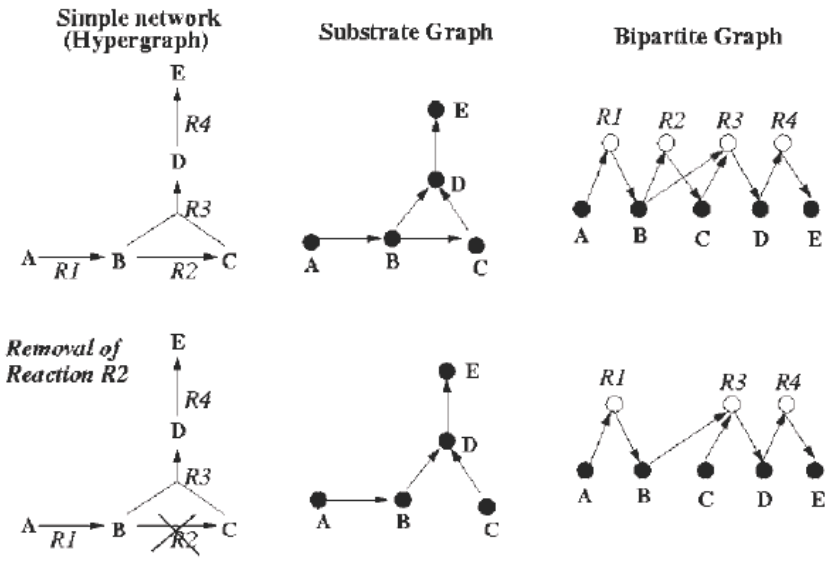
Klamt & Gilles, Bioinformatics 20, 226 (2004)

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),

→ use the fact that 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:

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

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 (CSC) 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)

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

objective reaction „biomass synthesis“

Network: 110 reactions, 89 metabolites, see Stelling et al. (2002) covered in V18

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)

Applications of MCSs

Target identification and repressing 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 run off 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

Cutting away an MCS from the network can be predicted to be definitely intolerable with respect to certain cellular reactions/processes.
These predictions, derived purely from network structure, might be suitable for verification of hypothetical or reconstructed networks.

If a set of gene deletions completely contains an MCS, then
(case 1) it should lead to a non-viable phenotype, otherwise
(case 2) growth is structurally possible.

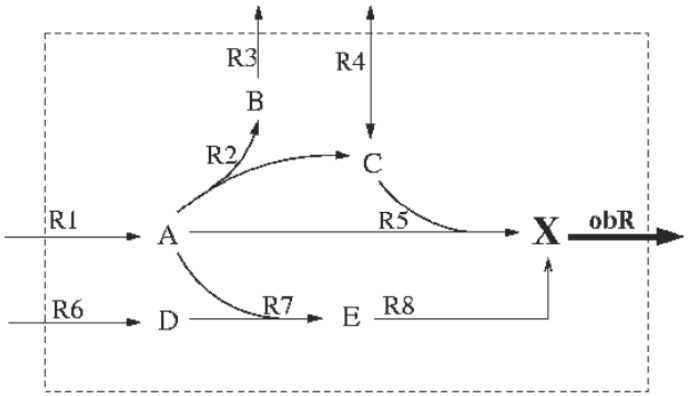
A wrong prediction for case 1 would be a false negative prediction and is a proof for an incorrect or incomplete network structure.
A wrong prediction for case 2 would be a false positive prediction and is a clue (but not necessarily proof) of a false assumption in the network structure.

Klamt & Gilles, Bioinformatics 20, 226 (2004)

Applications of MCSs

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)

Conclusion

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

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)