



FluxAnalyzer: exploring structure, pathways, and flux distributions in metabolic networks on interactive flux maps

Steffen Klamt*, Jörg Stelling, Martin Ginkel and Ernst Dieter Gilles

Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, D-39106 Magdeburg, Germany

Received on June 13, 2002; revised on July 30, 2002; accepted on August 2, 2002

ABSTRACT

Motivation: The analysis of structure, pathways and flux distributions in metabolic networks has become an important approach for understanding the functionality of metabolic systems. The need of a user-friendly platform for stoichiometric modeling of metabolic networks *in silico* is evident.

Results: The FluxAnalyzer is a package for MATLAB® and facilitates integrated pathway and flux analysis for metabolic networks within a graphical user interface. Arbitrary metabolic network models can be composed by instances of four types of network elements. The abstract network model is linked with network graphics leading to interactive flux maps which allow for user input and display of calculation results within a network visualization. Therein, a large and powerful collection of tools and algorithms can be applied interactively including metabolic flux analysis, flux optimization, detection of topological features and pathway analysis by elementary flux modes or extreme pathways. The FluxAnalyzer has been applied and tested for complex networks with more than 500 000 elementary modes. Some aspects of the combinatorial complexity of pathway analysis in metabolic networks are discussed.

Availability: Upon request from the corresponding author. Free for academic users (license agreement). Special contracts are available for industrial corporations.

Supplementary information: <http://www.mpi-magdeburg.mpg.de/projects/fluxanalyzer>

Contact: klamt@mpi-magdeburg.mpg.de

INTRODUCTION

One of the primary results derived from the huge amount of genomic and biochemical data currently produced is the

reconstruction of biochemical networks. Consequently, methods for analyzing functionality and regulation of metabolic networks are rapidly gaining in importance. Mathematical modeling and simulation of biological systems seems to be an approach capable of coping with the complexity of such networks and to study their behavior and capabilities *in silico* (Kremling *et al.*, 2000; Covert *et al.*, 2001). Modeling of cellular (sub)systems is not a new approach but it is becoming more attractive and the number of applications is growing rapidly. Virtual representations of cellular systems are not only useful for a system-level understanding of cellular processes (Kitano, 2002) but also for searching for promising targets of manipulations, e.g. in the pharmaceutical or biotechnological industry (Wiechert, 2002).

Diverse platforms have been developed for modeling cellular systems, including those for simulating metabolic networks on the basis of kinetic descriptions, for instance, GEPASI (Mendes, 1997) or JARNAC (Sauro, 2000), and those for whole cell modeling like E-CELL (Tomita *et al.*, 1999). However, analysis of the underlying stoichiometry of a metabolic network has been considered only to a minor extent.

Here, we present the FluxAnalyzer, a comprehensive and user-friendly graphical interface for analyzing metabolic networks at steady state. Studying the cellular metabolism by the use of quasi-stationary assumptions has frequently been used for quantifying metabolic fluxes (metabolic flux analysis) and for structural (topological) network analysis including pathway analysis. Applications can be found in 'pure' microbiology studies (e.g. Nuño *et al.*, 1997), in metabolic engineering and biotechnology (Stephanopoulos *et al.*, 1998; Wiechert, 2001; Schuster *et al.*, 2002) and for system-level analysis of biochemical networks (e.g. Edwards and Palsson, 2000; Schuster *et al.*, 2000; Stelling *et al.*, 2002; Papin *et al.*, 2002).

*To whom correspondence should be addressed.

PRINCIPLES OF STRUCTURE, PATHWAY, AND FLUX ANALYSIS IN METABOLIC NETWORKS

Metabolite balancing: the fundamental relation

For analyzing a biochemical or, more general, stoichiometric network, its structure has to be expressed by the stoichiometric matrix \mathbf{N} consisting of m rows corresponding to the substances (metabolites) and q columns corresponding to the stoichiometric coefficients of the metabolites in each reaction. Furthermore, a vector \mathbf{r} denotes the net reaction rates (mmol/(gDW h), DW = dry weight) and vector \mathbf{c} describes the metabolite concentrations (mmol/gDW).

If biomass synthesis is considered, then the stoichiometric matrix also contains a 'biosynthesis' column quantifying the cumulative efflux of metabolites into the biomass (mmol metabolite/gDW). The corresponding element of vector \mathbf{r} is then the growth rate μ .

Due to the high turnover of metabolite pools one often assumes pseudo-steady state ($\mathbf{c}(t)$ constant) leading to the fundamental *Metabolite Balancing Equation*:

$$\frac{d\mathbf{c}(t)}{dt} = \mathbf{0} = \mathbf{N}\mathbf{r} \quad (1)$$

Flux distributions \mathbf{r} satisfying (1) lie in the null space of \mathbf{N} (Heinrich and Schuster, 1996) and are able to balance all metabolites. They maintain homeostasis and are therefore meaningful for the long-term perspective of metabolism. The great advantage of Equation (1) is that the (mostly not available or uncertain) parameters and molecular mechanisms of the reactions are not involved.

Example network 'SMALLNET'

For illustration throughout this paper, we use a simple network 'SMALLNET' (Figure 1) involving 4 metabolites (A, B, C, D), 8 reactions (R1–R7 and cumulative biomass synthesis) and two macromolecular biomass components. The latter two are considered to be assembled from the metabolites:

Biomass component 1 : $\text{BC1[g]} = 2[\text{mmol}]A + 1[\text{mmol}]C$

Biomass component 2 : $\text{BC2[g]} = 1[\text{mmol}]C + 3[\text{mmol}]D$

The stoichiometric matrix of this network then reads:

$$\mathbf{N} = \begin{pmatrix} 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0.8 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & -1 & 1 & 1.8 \end{pmatrix} \quad (2)$$

The first seven columns correspond to reactions R1–R7. As indicated in Figure 1, reactions R1, R4, R6 and R7 are reversible. The last column of \mathbf{N} corresponds to biomass synthesis and results from the assumed biomass composition of 0.4 gBC1/gDW and 0.6 gBC2/gDW. Accordingly, changing the biomass composition would lead to an alteration of these entries.

Metabolic flux analysis

The aim of *Metabolic Flux Analysis* (MFA) is to determine preferably all components of the flux distribution \mathbf{r} in a metabolic network during a certain stationary growth experiment (scenario). Typically, some measured or known rates must be provided to calculate unknown rates. Accordingly, \mathbf{r} and \mathbf{N} are partitioned into the known (\mathbf{r}_b , \mathbf{N}_b) and unknown part (\mathbf{r}_n , \mathbf{N}_n). Rearranging (1) gives the central equation for MFA describing a flux scenario:

$$\mathbf{0} = \mathbf{N}\mathbf{r} = \mathbf{N}_n\mathbf{r}_n + \mathbf{N}_b\mathbf{r}_b \quad \mathbf{N}_n\mathbf{r}_n = -\mathbf{N}_b\mathbf{r}_b \quad (3)$$

A detailed description of how to further proceed with (3) is given elsewhere (van der Heijden *et al.*, 1994a,b; Klamt *et al.*, 2002) and involves the Penrose pseudo-inverse of $\mathbf{N}_n^\#$. The rank of \mathbf{N}_n determines whether scenario (3) is redundant and/or underdetermined. Redundant systems can be checked on inconsistencies. In underdetermined scenarios, only some elements of \mathbf{r}_n (or none at all) are uniquely calculable, which need to be found by an analysis of the null space of \mathbf{N}_n .

Figure 1 shows a simple non-redundant and underdetermined flux scenario in SMALLNET where \mathbf{r}_b consists of R1, R2, μ and \mathbf{r}_n of R3–R7, whereof R3, R4, R7 are calculable (discussed further below).

Structural network analysis

Whereas MFA focuses on a single flux distribution, techniques of *Structural (Stoichiometric, Topological) Network Analysis* (SNA) address general topological properties, overall capabilities, and the inherent pathway structure of a metabolic network. Basic topological properties are, for example, conserved moieties (Heinrich and Schuster, 1996). *Flux Balance Analysis* (FBA; Edwards and Palsson, 2000) searches for single *optimal* flux distributions (mostly with respect to the synthesis of biomass) fulfilling (1) and additionally reversibility and capacity restrictions for each reaction.

Metabolic Pathway Analysis (MPA) searches for meaningful structural and functional units in metabolic networks. The two most promising, very similar approaches are based on convex analysis and use the sets of elementary flux modes (EFMs; Schuster *et al.*, 1999, 2000) and extreme pathways (EPs, Schilling *et al.*, 2000), respectively. 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 (Papin *et al.*, 2002; Stellling *et al.*, 2002) and functionality of metabolic networks. It enables the identification of futile cycles and all (sub)optimal pathways with respect to product/biomass yield (Schuster *et al.*, 2000). EFMs are also useful for calculability studies in metabolic flux analysis (Klamt *et al.*, 2002).

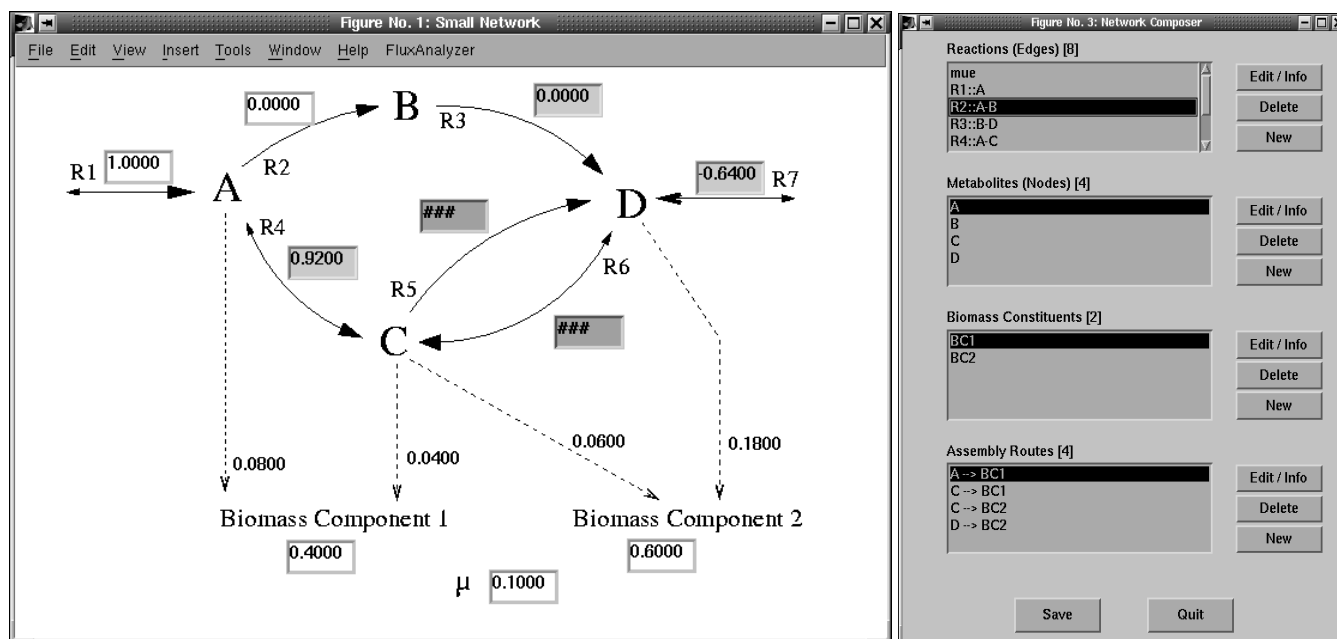


Fig. 1. The network project of 'SMALLNET' constructed by the FluxAnalyzer. Left: interactive flux map displaying a flux scenario (unknown rates are denoted by '###'). Right: network composer.

BASIC SETUP OF THE FLUXANALYZER: NETWORK PROJECTS AND INTERACTIVE FLUX MAPS

The development of the FluxAnalyzer was motivated by the fact that, although flux analysis and, to a slightly lesser extent, structure and pathway analysis are well-known frameworks for analyzing metabolic networks, there is a lack of a software tool which integrates these techniques in a comprehensive and user-friendly graphical interface. The FluxAnalyzer is a package for the commercial program MATLAB® (Mathworks Inc.; www.mathworks.com), a widely-used platform for complex computations.

As the structural setup shows (Figure 2), the FluxAnalyzer provides a toolbox for studying user-created *network projects*. Each network project contains at first an abstract (symbolic) network model constructed by the interactive declaration of network elements. Furthermore, network graphics visualizing the metabolic network (metabolic maps) must be provided by the user. Automatic layout of networks, particularly metabolic, using a symbolic network description is complex and does not always lead to a representation as desired. Therefore, the FluxAnalyzer gives the user the option to design and annotate his own network graphic(s) by external graphic programs or to use network representations such as provided by KEGG (<http://www.genome.ad.jp/kegg/kegg2.html>) or BioPath (<http://biopath.fmi.uni-passau.de>). Arbitrary

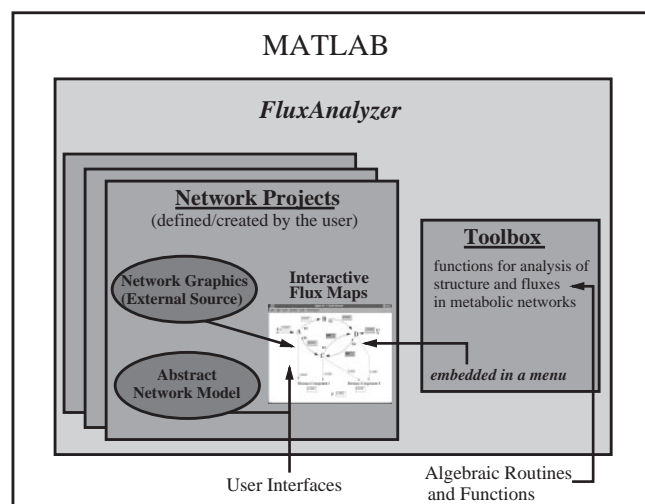


Fig. 2. Structural setup of the FluxAnalyzer.

numbers of metabolic maps in a variety of graphic file formats can be incorporated in a network project. After loading the network project into the FluxAnalyzer, each network graphic serves as a background in a MATLAB figure. Thereon, text boxes can easily be placed, e.g. at the associated pathway, referring to abstract network elements. These text boxes facilitate intuitive display and

user input of reaction rates and biomass composition in a network visualization. For the linkage of network graphics with the abstract network model by user interfaces we introduce the term *interactive flux maps*. Figure 1 depicts the network project of SMALLNET assembled by the FluxAnalyzer.

For analyzing network projects the toolbox of the FluxAnalyzer comprises various algorithms and functions which can conveniently be started by a pull-down menu within the flux maps.

BUILDING NETWORK PROJECTS

For composing an abstract network model of an arbitrary (bio)chemical reaction system, the FluxAnalyzer provides four types of network elements: *metabolites*, *reactions*, *biomass constituents* and *assembly routes*. Each type owns a set of properties (Table 1), which has to be defined upon declaration of a new instance of this type. The set of properties of all types except *metabolite* comprises also variables defining style (editable/non-editable) and position (conveniently defined via a crosshair) of the element's text box occurring in one of the flux maps. A text box will always correspond to its assigned network element and allows user input as well as output of calculated results directly on the network graphics.

The network structure can be managed, edited and stored by the *network composer* (Figure 1, right). Type-specific input masks (see web-site) allow to define, modify or delete network elements any time during a session of a network project. The user-defined *symbolic reaction equations* internally generate the stoichiometric matrix \mathbf{N} .

Actually, the element types *metabolite* and *reaction* allow one to compose arbitrary stoichiometric networks. Biomass synthesis can be considered as a special 'reaction' (in the FluxAnalyzer denoted by '*mue*' for growth rate μ). The stoichiometry of '*mue*' specifies the cumulative efflux of metabolites into biomass. It depends on biomass composition and can therefore vary even for the same organism. The element type *biomass constituent* (BC) enables the convenient determination of the overall stoichiometry of '*mue*'. The properties of each BC (like protein, DNA, RNA, or lipids) allows to include the cumulative stoichiometry of consumption of metabolites for its synthesis (mmol metabolite/gBC; cf. Neidhardt et al., 1990). Together with the user-defined biomass composition the current stoichiometry of '*mue*' can be computed internally.

Assembly route is an auxiliary element type allowing one to display the consumption of a certain metabolite for the biosynthesis of one BC. In SMALLNET, for example, metabolite C is needed for synthesis of BC1 and BC2 and participates therefore in two *assembly routes* (Figure 1, dashed lines). Note, as the assembly *rates* are implicitly

determined by growth rate and biomass composition, they are not part of the reaction vector in the balancing Equation (1) and their associated text boxes have 'non-editable' style.

TOOLBOX FOR ANALYZING STRUCTURE, PATHWAYS AND FLUXES IN METABOLIC NETWORKS

After creating a network project, the user can start to study it by a variety of functions provided in the FluxAnalyzer's toolbox covering a broad spectrum of techniques for flux and structural analysis in metabolic networks. The procedures are implemented in the MATLAB programming language (executable *m*-files), often taking advantage of a large set of predefined and optimized matrix functions. However, some special procedures of the FluxAnalyzer had to be fully self-implemented.

All functions can be started from a pull-down menu automatically installed in the first flux map (Figure 1: menu-item 'FluxAnalyzer'). Using the interactive flux maps and the menu-controlled functions, the user does not need to be aware of the mathematical details of calculations.

Usually, starting an action from the menu is accompanied by a readout of the text box values. Together with the internal abstract network model, they are used for the respective calculations. In some cases, the user is requested for further specifications. After the calculation has been finished, results are shown on the flux maps and/or in separate windows. An overview of the most important functions provided by the menu is given in the following paragraphs (the complete menu is shown on the web-site).

Analyzing basic topological network properties

The FluxAnalyzer facilitates the extraction of some basic features of the network's topology, e.g. to detect construction errors or to remove redundant constraints caused by conservation relations:

- detection of *dead-end metabolites* (participating only in one reaction) and those *never participating*;
- *strictly detailed balanced reactions*: reactions whose rates are *per se* determined to be zero, for example, when involving a dead-end metabolite;
- *enzyme subsets*: reactions always operating together to keep the network balanced (Pfeiffer et al., 1999);
- if the *rank* of \mathbf{N} is smaller than the number of metabolites, then linear dependencies between the rows occur (*conservation relations*). In this case, all elementary conservation relations are determined from which the non-negative ones are useful for detecting conserved moieties (Heinrich and Schuster, 1996).

Table 1. Network element types in the FluxAnalyzer and their properties

Element type	Metabolite (Substance)	Reaction	Biomass constituent	Assembly route
	Network node	(Bio)chemical conversion	Substantial macromolecule	Efflux of a metabolite for synthesis of a biomass constituent
Text box	None	Reaction rate ([mmol]/(gDW h))	Relative biomass concentration ([g/gDW])	Rate of metabolite consumption ([mmol]/(gDW h))
Properties	<ul style="list-style-type: none"> • Full name • Symbolic name • External-flag (External metabolites are not considered to be in pseudo-steady state and therefore not balanced by Equation (1))	<ul style="list-style-type: none"> • Symbolic name • Reaction equation • Default rate • Rate minimum • Rate maximum • Coefficient in linear objective function • Variance of measurements • Text box parameters 	<ul style="list-style-type: none"> • Full name • Symbolic name • Default concentration • Cumulative synthesis equation • Text box parameters 	<ul style="list-style-type: none"> • Symbolic name • Metabolite • Biomass constituent • Text box parameters

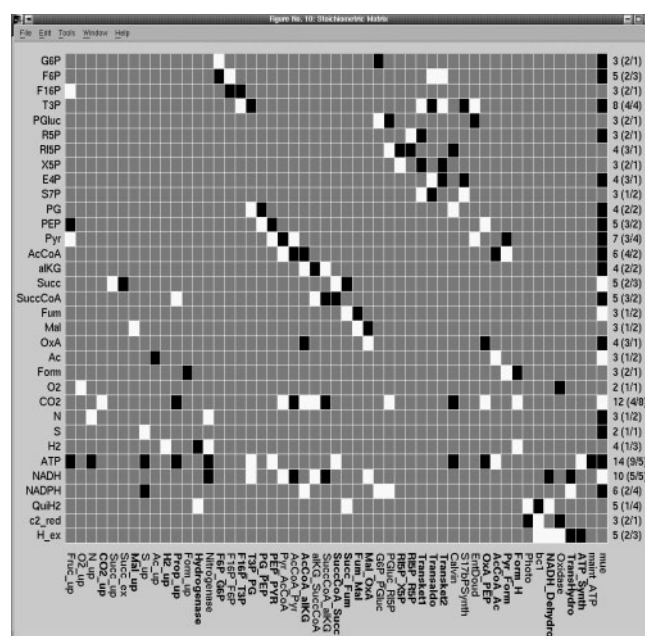
For obtaining a more detailed, but still concise overview of the complete stoichiometric system, the FluxAnalyzer provides an intuitively comprehensible graphical display of the stoichiometric matrix (Figure 3). The rows correspond to metabolites (names shown left), the columns represent the reactions (names shown on the bottom, reversible reactions are bold). A cell n_{ij} is black if the metabolite i is consumed in reaction j , white if produced or gray if not involved. At the end of each row, the metabolite's connectivity number, and the number of consuming and producing reactions are shown. This condensed representation provides one example for how network analysis, especially for large networks, is supported in the FluxAnalyzer by graphical displays.

Computing the set of elementary modes (extreme pathways/convex bases)

One of the most powerful facilities of the FluxAnalyzer is pathway analysis, i.e. the calculation, display and evaluation of elementary flux modes, extreme pathways, and convex bases directly assisted by the interactive flux maps. The capability to calculate and handle huge sets of these structural elements also enables one to apply these tools to complex networks.

Determination of elementary modes is a highly combinatorial problem. This is reflected by the upper bound S_{\max} for the number S of elementary modes (assuming full rank of \mathbf{N} and number of metabolites $m <$ number of reactions q):

$$S \leq S_{\max} = \binom{q}{m+1} = \frac{q!}{(q-m-1)!(m+1)!} = \frac{q(q-1)(q-2) \dots (q-m)}{(m+1)!}. \quad (4)$$

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

If, for example, one reaction is added to a network, the boundary increases by a factor of $(q+1)/(q-m)$. Fortunately, at least in biochemical networks, Equation (4) seems to be a very pessimistic estimation (for SMALL-NET: $S = 12$, $S_{\max} = 56$; see also the large example below). S_{\max} would only be reached if all possible subsets containing m columns of \mathbf{N} were linearly independent and if all reactions were reversible. The number of modes is greatly reduced by small pathways, irreversible reactions,

the relatively low connectivity of many metabolites and enzyme subsets (Klamt and Stelling, 2002). For example, an enzyme subset can be considered as one lumped reaction without changing the modes in the network, thus, lowering q in Equation (4) and accordingly S_{max} . The upper bound represents rough estimates of the problem complexity, since e.g. the number of modes even for networks having the same number of metabolites and reactions can differ by magnitudes (Klamt and Stelling, 2002). This emphasizes the inherent structural information reflected by the elementary modes.

The calculation procedure for elementary modes is based on the iterative algorithm described in Schuster *et al.* (2000). A similar algorithm is used for computing convex bases and extreme pathways, providing the network has been configured according to Schilling *et al.* (2000). To cope with the combinatorial complexity of large networks, we optimized the algorithm with respect to speed, memory requirement and numerical stability. Preprocessing steps map the original network to a smaller one yielding the same final modes by lumping enzyme subsets into one overall reaction (Schuster *et al.*, 2000), and by removal of nonparticipating metabolites and strictly detailed balanced reactions. Moreover, the tableau is ordered ascending with respect to metabolite connectivity to avoid a rapidly growing tableau already during the first iterations. In SMALLNET, R2 and R3 would be lumped into one reaction $A \rightarrow D$, metabolite B would be removed and the metabolite order would be A, C, D.

During the iterative algorithm, preliminary modes are compared to each other to ensure that no duplicate and only *elementary* modes are calculated (Schuster *et al.*, 2000). These subset checks turned out to be the most time-consuming part of the computation (50% in a larger problem). The participating reactions of each elementary mode are therefore stored bit-wise to save memory and to allow the use of fast bit operations increasing the speed drastically. Parts of the algorithm were implemented in external C functions directly accessible to MATLAB (via its MEX interface), which proved very favorable under the conditions of large memory throughput and relatively simple loops.

Since the algorithm uses finite-precision real numbers (in contrast to integer arithmetic in METATOOL; Pfeiffer *et al.*, 1999) the user can control the numerical precision by defining *epsilon* (smallest number greater than zero). To avoid badly scaled row operations, preliminary modes are normalized after each iteration.

As a ‘benchmark test’ we calculated the elementary modes in the central metabolism in *Escherichia coli* comprising 89 metabolites and 110 reactions including simultaneous uptake of four substrates, excretion of 5 products and biomass synthesis (Klamt and Stelling, 2002). According to Equation (4), S_{max} is 4.39×10^{21} ,

which drops to 4.85×10^{13} after preprocessing. Due to aforementioned reasons, ‘only’ 507 632 elementary modes were calculated by the FluxAnalyzer, which, however, far exceeds previously reported set sizes of approximately 10,000. In general, the computation time (here: about 50 h, Intel Pentium IV, 1.0 GHz, 4 GB RAM) correlates well with the square of the final number of elementary modes (Klamt and Stelling, 2002).

Although the algorithm has been optimized, it seems to be impossible or would at least require parallel algorithms to calculate all elementary modes in genome-wide networks (cf. Edwards and Palsson, 2000: 436 metabolites, 720 reaction for *E. coli*). Nevertheless, applications of pathway analysis demonstrate that it is also worth and feasible to study the modes in (still large) subnetworks. For this purpose, the FluxAnalyzer supports user-defined exclusions of reactions before calculating the modes.

Analyzing the set of elementary modes

Once the calculation of the elementary modes has been finished, a control panel comes up with several functions for studying—and saving—the set of these topological elements:

Display of elementary modes: Each elementary flux mode can be displayed separately in the flux maps in an intuitive way. The participating reactions (and the fluxes they carry) are identified by an emphasized coloration of their boxes.

Selections: In large networks it would be a tedious task to step through the complete set of modes. The user might be rather interested in subsets of modes fulfilling certain conditions, e.g. ‘reaction r_x or metabolite M is / is not involved’ or ‘metabolite Y is excreted’ or ‘pathway length (number of involved reactions) is less/larger than or equal to z ’. The ‘Selection’ tool allows one to build up subsets of modes by specifying such properties. Different subsets can be combined through ‘subset clipboards’ and ‘UNION’, ‘INTERSECTION’, ... operators. When a selection has been specified, all functions work only on these subsets, which is especially useful for statistical evaluations.

Statistics: A valuable statistical method is to determine how often each reaction is applied in the current selection. For SMALLNET, we would obtain information like: 58% (7/12) of the modes allow growth; besides substrate uptake, the most important reaction for growth is R4 participating in 71% (5/7) of these modes; biomass production on *exclusively* substrate A (four modes) shows higher flexibility than on substrate D (one mode). In our eyes, such ‘combinatorial pathway analyses’—by studying *subsets* of elementary modes—provides deeper insights into routing constraints of different growth regimes in metabolic networks.

Pathway engineering often relies on the detection of pathways with maximal product or growth yield for a

certain substrate. In contrast to linear optimization, by screening the elementary modes *all* optimal routes are found and displayed by the FluxAnalyzer. Displaying a histogram of *pathway lengths* and the determination of all *enzyme subsets* occurring in the selected modes round off the statistical tools.

Tools for flux analysis

Since the functions related to flux analysis focus only on single (and not all feasible) flux distributions they are computationally much less elaborate than pathway analysis. In general, performing flux analysis in the FluxAnalyzer requires the definition of a (*flux*) *scenario* by entering the biomass composition and some measured or known rates. The following procedures are available for flux analysis (for termini see **Introduction**):

‘Feasibility check’: This function checks whether for a given scenario any flux distribution exists at all, that complies with the currently defined rates, the mass balances, and constraints on reaction rates including reaction reversibilities. For example, setting $R_4 = R_6 = 0$ and $R_7 = 0.1$ in SMALLNET results in an underdetermined scenario where no other rate could be determined. However, as one can simply verify, there is no feasible flux distribution applying these rates as R_3 and R_5 are irreversible. In this way, one can detect inconsistencies or check whether a ‘virtual’ mutant (by setting the respective reaction rate to zero) is still able to grow or not.

‘Classify rates’: Checks whether the scenario is redundant and/or determined and examines which of the unknown rates are calculable (observable) and which of the known rates cause redundancies, i.e. are balanceable. For planning experiments, for instance, one might check whether a certain set of measurements would be sufficient to calculate all or at least some rates.

‘Determine calculable rates’: This function computes and displays all calculable rates for nonredundant scenarios. For a better distinction, the text boxes of defined (here: white), calculated (gray) and noncalculable (dark gray) rates as well as the biomass composition (white) are colored differently. Thus, the example scenario in Figure 1 can be fully reconstructed.

‘Procedures for redundant systems’: For treating redundant scenarios, three methods are provided for calculating a least square solution (‘Simple least squares’, ‘Variance-weighted least square’, ‘Gross error detection’). The latter two enable to incorporate the variances of measured rates (reaction property, Table 1) to detect modeling/measurement errors (Stephanopoulos *et al.*, 1998) or to find the reaction rate that causes the highest deviation from a consistent system (cf. van der Heijden *et al.*, 1994a).

‘Sensitivity analysis’: Applying this tool, the user can check the sensitivity dr_c/dr_m (partial derivative) of a

calculated rate, r_c , with respect to a measured or given rate, r_m .

‘Optimization’: This procedure minimizes the linear objective function

$$f = c_1 r_1 + c_2 r_2 + \dots c_n r_n = \min!$$

whereby c_1, \dots, c_n are the respective coefficients for each reaction (see Table 1) and r_1, \dots, r_n the reaction rates. As additional constraints, the lower and upper boundaries of each reaction rate are considered as well as the linear relation (3) arising by user-defined rates. Thus, arbitrary optimization problems for the network can be defined and handled (see example below).

Flux clipboard

The provided flux clipboard serves like a common clipboard: the current set of reaction rates can be copied from and later pasted back into the corresponding text boxes. This is in particular useful for combining two different flux distributions arithmetically (‘+’, ‘-’, ‘*’, ‘\’, normalization by a scalar) and, thereby, comparing them. This is illustrated by the following example:

Assume a user wants to compare two flux distributions optimized with respect to growth rate and ATP production, respectively. A first step would be to optimize a scenario with respect to the growth rate and to copy the resulting flux distribution to the flux clipboard. For optimizing the network with respect to ATP production, the objective function’s coefficients for the growth rate must be changed in the network composer from -1 to 0 and for ‘ATP production’ from 0 to -1 (the optimization routine minimizes). Then, the rates defined prior to the first optimization can be reset and the network is optimized again, which results in the calculation and display of the optimized flux distribution with respect to ATP production. Using the arithmetic operations one can calculate and display the differences between this solution and the growth-optimal solution stored in the flux clipboard. Optionally, the results can be printed and stored or be further analyzed in a bar chart allowing for quick identification of the high and low values.

DISCUSSION AND CONCLUSIONS

In the growing interdisciplinary field of systems and *in silico* biology, one of the most challenging tasks is to develop user-friendly software tools enabling reproducible network and systems analyses, also for users not so familiar with the underlying algorithms. Accordingly, the central focus during the development of the FluxAnalyzer was not only to create powerful algorithms for studying metabolic networks but also to embed them into a comprehensive graphical user interface. The core concept of the FluxAnalyzer are the interactive flux maps. They

enable a fully interactive and menu-controlled network analysis. The opportunity to incorporate externally created graphics preserves a high degree of freedom for the layout of the flux maps and the option to use sophisticated tools for generating them. The conceptual framework for composing arbitrary stoichiometric network models also ensures an adequate consideration of biomass composition and biomass synthesis in metabolic networks.

Due to the combination of visualization, interactivity and broad functionality the FluxAnalyzer represents a valuable tool for a holistic analysis of *in silico* metabolic networks based on their underlying stoichiometry. It complements other tools such as Jarnac (Sauro, 2000) or GEPASI (Mendes, 1997), which focus more on the dynamic simulation of metabolic networks or Metabolic Control Analysis (MCA). In these software packages, structural network analysis is considered only to a supplemental extent, e.g. both programs contain the routine from METATOOL for calculating elementary modes. METATOOL (Pfeiffer *et al.*, 1999) is a frequently used tool for pure structural network analysis. However, none of these programs is embedded in a graphical interface such that user input as well as output of results can be realized intuitively in a network visualization. Nevertheless, for extended analysis using external programs, it is possible to export the stoichiometric matrix from the FluxAnalyzer in ASCII format or to generate an input file for METATOOL.

Besides flux analysis, one of the main facilities of the FluxAnalyzer are tools for exploring the pathway structure of a network. An efficient algorithm has been developed which calculates elementary flux modes or extreme pathways even in networks of higher complexity. The interactive flux maps allow a convenient display of the extracted structural features. 'Pathway subsets' having certain properties can be defined representing different growth regimes or strategies. Some tools have been introduced for studying such subsets providing further insights in the functionality of a metabolic network, e.g. for evaluating the importance of a reaction with respect to a selected subset.

Due to space limitations it is impossible to demonstrate the application of the FluxAnalyzer in larger and more realistic networks. The user should refer to Klamt *et al.* (2002) dealing with observability of reaction rates and detection of structural constraints in the metabolism of purple non-sulfur bacteria, or to our web-site. Moreover, the FluxAnalyzer was used for elementary-mode analysis of the central metabolism in *E. coli* (Klamt and Stelling, 2002) to predict mutant phenotypes and gene expression ratios to a certain extent (Stelling *et al.*, 2002). An industrial corporation uses the FluxAnalyzer for online calculation and visualization of metabolic fluxes during fermentation processes.

The FluxAnalyzer is still under construction and it is straightforward to implement and embed new features based on operations on the stoichiometric matrix. Possible extensions or applications for the FluxAnalyzer could also be the display of results obtained by the more complicated ^{13}C flux analysis (Wiechert, 2001), or by dynamic simulations of metabolic networks. Furthermore, the FluxAnalyzer supports the analysis of any kind of stoichiometric network, for example chemical reaction networks as studied in (Happel and Sellers, 1989).

REFERENCES

- Covert, M.W., Schilling, C.H., Famili, I., Edwards, J.S., Goryanin, I.I., Selkov, E. and Palsson, B.O. (2001) Metabolic modeling of microbial strains *in silico*. *Trends Biochem. Sci.*, **26**, 179–186.
- Edwards, J.S. and Palsson, B.O. (2000) The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl Acad. Sci. USA*, **97**, 5528–5533.
- Happel, J. and Sellers, P.H. (1989) The characterization of complex systems of chemical reactions. *Chem. Eng. Comm.*, **83**, 221–240.
- van der Heijden, R.T.J.M., Heijnen, J.J., Hellinga, C., Romein, B. and Luyben, K.Ch.A.M. (1994a) Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. *Biotechnol. Bioeng.*, **43**, 3–10.
- van der Heijden, R.T.J.M., Romein, B., Heijnen, J.J., Hellinga, C. and Luyben, K.Ch.A.M. (1994b) Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotechnol. Bioeng.*, **43**, 11–20.
- Heinrich, R. and Schuster, S. (1996) *The Regulation of Cellular Systems*. Chapman and Hall, New York.
- Kitano, H. (2002) Systems biology: a brief overview. *Science*, **295**, 1662–1664.
- Klamt, S., Schuster, S. and Gilles, E.D. (2002) Calculability analysis in underdetermined metabolic networks illustrated by a model of the central metabolism in purple nonsulfur bacteria. *Biotechnol. Bioeng.*, **77**, 734–751.
- Klamt, S. and Stelling, J. (2002) Combinatorial complexity of pathway analysis in metabolic networks. *Mol. Biol. Rep.*, **29**, 233–236.
- Kremling, A., Jahreis, K., Lengeler, J.W. and Gilles, E.D. (2000) The organization of metabolic reaction networks: a signal-oriented approach to cellular models. *Metab. Eng.*, **2**, 190–200.
- Mendes, P. (1997) Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem. Sci.*, **22**, 361–363.
- Neidhardt, F.C., Ingraham, J.L. and Schaechter, M. (1990) *Physiology of the Bacterial Cell: A Molecular Approach*. Sinauer, Sunderland, MA.
- Nuño, J.C., Sánchez-Valdenebro, I., Pérez-Iratxeta, C., Meléndez-Hevia, E. and Montero, F. (1997) Network organization of cell metabolism: monosaccharide interconversion. *Biochem. J.*, **324**, 104–111.
- Papin, J.A., Price, N.D., Edwards, J.E. and Palsson, B.O. (2002) The genome scale metabolic extreme pathway structure in

- Haemophilus influenzae* shows significant network redundancy. *J. Theor. Biol.*, **215**, 67–82.
- Pfeiffer,T., Sánchez-Valdenebro,I., Nuño,J.C., Montero,F. and Schuster,S. (1999) METATOOL: for studying metabolic networks. *Bioinformatics*, **15**, 251–257.
- Sauro,H.M. (2000) Jarnac: a system for interactive metabolic analysis. In Hofmeyr,J.-H.S., Rohwer,J.M. and Snoep,J.L. (eds), *Animating the cellular Map: Proceedings of the 9th International Meeting on BioThermoKinetics*. Stellenbosch University Press.
- Schilling,C.H., Letscher,D. and Palsson,B.O. (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J. Theor. Biol.*, **203**, 229–248.
- Schuster,S., Dandekar,T. and Fell,D.A. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol.*, **17**, 53–60.
- Schuster,S., Fell,D. and Dandekar,T. (2000) A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.*, **18**, 326–332.
- Schuster,S., Klamt,S., Weckwerth,W., Moldenhauer,F. and Pfeiffer,T. (2002) Use of network analysis of metabolic systems in bioengineering. *Bioproc. Biosyst. Eng.*, **24**, 363–372.
- Stephanopoulos,G.N., Aristidou,A.A. and Nielsen,J. (1998) *Metabolic Engineering*. Academic Press, San Diego.
- Stelling,J., Klamt,S., Bettenbrock,K., Schuster,S. and Gilles,E.D. (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature*, **420**, 190–193.
- Tomita,M., Hashimoto,K., Takahashi,K., Shimizu,T.S., Matsuzaki,Y., Miyoshi,F., Saito,K., Tanida,S., Yugi,K., Venter,J.C. and Hutchison,3rd,C.A. (1999) E-CELL: software environment for whole-cell simulation. *Bioinformatics*, **15**, 72–84.
- Wiechert,W. (2001) ¹³C metabolic flux analysis. *Metab. Eng.*, **3**, 195–206.
- Wiechert,W. (2002) Modeling and simulation: tools for metabolic engineering. *J. Biotechnol.*, **94**, 37–64.