Mini-Review

Elementary flux modes in a nutshell: Properties, calculation and applications

Jürgen Zanghellini[1,2], David E. Ruckerbauer[1,2], Michael Hanscho[1,2] and Christian Jungreuthmayer[1,2]

[1] Austrian Centre of Industrial Biotechnology, Vienna, Austria
[2] Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna, Austria

Elementary flux mode (EFM) analysis allows the unbiased decomposition of a metabolic network into minimal functional units, making it a powerful tool for metabolic engineering. While the use of EFM analysis (EFMA) is still limited by the size of the models it can handle, EFMA has been successfully applied to solve real-world metabolic engineering problems. Here we provide a user-oriented introduction to EFMA, provide examples of recent applications, analyze current research strategies to overcome the computational restrictions and give an overview over current approaches, which aim to identify and calculate only biologically relevant EFMs.

Received 07 DEC 2012
Revised 26 FEB 2013
Accepted 08 MAY 2013

Keywords: Computational methods · Metabolic engineering · Metabolic pathway analysis · Minimal cut set · Strain optimization

1 Introduction: An illustration of elementary flux mode analysis (EFMA)

Elementary flux mode analysis (EFMA) is a promising tool for metabolic engineering; however, the calculation of elementary flux modes (EFMs) is computationally demanding, which hinders widespread application. We outline the fundamental (mathematical) properties of EFMA in this review, and discuss recent applications in bioengineering and biotechnology, focusing on methodical developments in the computation of EFMs.

To provide a "real life" analogy of EFMA, we use the following subway map example. Suppose Anna and Bruno live in Vienna, Austria. She lives in *Ottakring* and he in *Heiligenstadt* (Fig. 1). Bruno wants to visit Anna at her place. She wants him to take a circuitous route and bring some ice cream from a particular shop at *Schwedenplatz*. Bruno does not want to make a detour and would like to get to her place as quickly as possible. How can Anna convince him to travel via *Schwedenplatz* and bring ice cream with him?

To solve this problem, Anna analyzes all possible routes Bruno may take and tries to find a way to eliminate those pathways that do not pass *Schwedenplatz*. This can be done with the help of EFMA. EFMs are minimal pathways through a network. They are minimal in the sense that if transit at any point of such an elementary route is disrupted, Bruno will not reach his destination. Importantly, every possible path through the transit network can be described as a combination of EFMs. Figure 1B and 1C illustrate what is and what is not considered an "EFM". The full list of EFMs connecting *Heiligenstadt* to *Ottakring* is shown in the Supporting information.

If Anna knows all EFMs, she will be able to describe every path Bruno may take as a combination of several EFMs (Fig. 1C). Anna finds Bruno has the option of 32 different elementary routes to travel between *Heiligenstadt* and *Ottakring*. After analyzing those 32 EFMs, she realizes that a service interruption in the brown line anywhere between *Spittelau* and *Westbahnhof* and, simultaneously, in the purple line anywhere between *Schottenring* and *Volkstheater*, would force all remaining admissi-

Correspondence: Dr. Jürgen Zanghellini, Austrian Centre of Industrial Biotechnology, Muthgasse 11, A-1190 Wien, Austria
E-mail: juergen.zanghellini@acib.at

Abbreviations: EFM, elementary flux model; **EFMA**, elementary flux mode analysis; **GSMM**, genome-scale metabolic model

Color online: refer to online PDF file for figure in color.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Biotechnol. J. 2013, 8, 1009–1016

Figure 1. (**A**) Map of Vienna's subway system. Anna lives in *Ottakring* "A", Bruno in *Heiligenstadt* "B". How can Anna convince her friend to travel past *Schwedenplatz* "S" on his way to her? If service interruptions were to occur at the indicated positions "X", all remaining pathways would pass *Schwedenplatz*. (**B**) Examples of EFMs (blue full and blue dashed lines). These modes are elementary, as Bruno will not get to his destination if a service interruption occurs anywhere along the pathways. (**C**) Examples of non-EFMs. The blue dashed line is not an EFM as it does not get Bruno to a terminal station. The blue full line is not an EFM, as it contains a loop and is therefore not indivisible. To illustrate: suppose a service interruption were to occur somewhere in the loop (e.g. at position X). By definition, an EFM will get blocked if it is interrupted anywhere along its path. However, by cutting a corner (traveling along the blue full line without entering the loop) Bruno gets to *Ottakring*. Thus, the blue full line cannot be an EFM. In fact, it represents the superposition of two EFMs, i.e. EFM1: *Heiligenstadt – Spittelau – Schwedenplatz – Westbahnhof – Ottakring* and EFM2: the loop *Stephansplatz – Landstraße – Karlsplatz – Stephansplatz*. (See text for further details.) The figure is a revision of work by "User:My Friend" licensed under Creative Commons Attribution-Share Alike 3.0 Unported, http://creativecommons.org/licenses/by-sa/3.0/legalcode. The original file may be found at http://commons.wikimedia.org/wiki/File: U-Bahn_Wien.png

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BIOtec visions
www.biotecvisions.com

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ble pathways to pass *Schwedenplatz*. Bruno would have no excuse not to stop there to pick up ice cream.

If the subway map is replaced by a metabolic pathway map; "Bruno" exchanged for a carbon source; "Anna" replaced with a product of interest and the necessity to pass *Schwedenplatz* substituted with a requirement for efficient production, a classical metabolic engineering problem is defined. That is, how can an organism be engineered to utilize a given substrate for the production of a product of interest as efficiently as possible?

A subway map is conceptually simpler than a metabolic network. While in the subway illustration, "reactions" only involve one reactant and one product (i.e. there are only pairwise connections between the stations), in a metabolic network, reactions will typically involve several products and several reactants. Furthermore, different metabolic reactions may be coupled, for example via cofactor regeneration. Even though metabolic networks appear more complicated, the fundamental principle remains the same.

2 Definition and properties of EFMs

Given a metabolic network, we are interested in analyzing its steady-state flux distributions. To this end, metabolites can be classified as internal or external. While internal metabolites are assumed to be in steady-state (i.e. their concentrations remain constant over time), the external metabolites are not. External metabolites act as sources or sinks for fluxes. In our example above, we designated the two terminal stops, *Heiligenstadt* and *Ottakring,* as "external". Thus, EFMs were either closed loops or connected two termini. Classification of a variable as internal or external is dependent on the problem at hand and up to the modeler. The number and structure of the resulting EFMs is strongly dependent on this choice.

The topology of a metabolic network is characterized by its $m \times n$ stoichiometric matrix, $\boldsymbol{S}$, where m and n correspond to the number of (internal) metabolites and reactions respectively. The value $S_{i,j}$ represents the stoichiometric coefficient of metabolite i in reaction j. $S_{i,j}$ is positive if metabolite i is produced in reaction j, and zero if metabolite i does not contribute to reaction j. If metabolite i is consumed in reaction j, then $S_{i,j}$ is negative.

We call any non-trivial flux vector $\boldsymbol{v}$ an admissible (flux) mode if $\boldsymbol{v} \neq \boldsymbol{0}$, $\boldsymbol{v} \in \mathrm{R}^n$, solves $\boldsymbol{Sv} = \boldsymbol{0}$ (steady-state condition) and obeys all (ir-)reversibility constraints of the reactions in question. Additionally, we define the support of a mode, $\mathbf{supp}(\boldsymbol{v}) = \{i | v_i \neq 0\}$ as the set of indices of non-zero elements in the flux mode $\boldsymbol{v}$. A mode is called an EFM, $\boldsymbol{e}$, if its support, $\mathbf{supp}(\boldsymbol{e})$, cannot be written as a proper superset of any other feasible mode $\boldsymbol{v}$, i.e. $\mathbf{supp}(\boldsymbol{e}) \supset \mathbf{supp}(\boldsymbol{v})$. In other words, an EFM is a minimal, unique set of flux-carrying reactions operating in steady-state. If any of its contributing reactions is deleted, the EFM can no longer operate in steady-state and the complete EFM is inactive [1–4].

EFMs can be found by removing one reaction at a time and solving the steady-state condition until it is no longer possible to remove a reaction while still obtaining an admissible flux distribution [5]. Also, it can be easily checked whether a given flux mode is elementary [6, 7], as only for EFMs does the nullity $[\boldsymbol{S}_{\mathrm{red}}(\boldsymbol{v})] = 1$. Here $\boldsymbol{S}_{\mathrm{red}}(\boldsymbol{v}) = [S_{i,j}]$, $i \in \{1, \ldots, m\}$, $j \in \mathbf{supp}(\boldsymbol{v})$ denotes the reduced stoichiometric matrix. $\boldsymbol{S}_{\mathrm{red}}$ depends on $\boldsymbol{v}$ and contains only those columns of $\boldsymbol{S}$ for which $v_i \neq 0$. Unfortunately this is an *ex post* verification. As it turns out, the calculation of all EFMs for an entire network is computationally demanding and a longstanding, unsolved mathematical question [8]. The challenge sits in the combinatorial explosion of the number of EFMs with increasing size of the network [9]. For instance, the medium-scale model for the core metabolism of *Escherichia coli* (containing about 100 reactions) [10] consists of around 272 million EFMs [11].

3 Significance of EFMs

EFMs are non-decomposable steady-state pathways through a network. The non-decomposability criterion implies that if any of its contributing reactions is deleted, an EFM will not be able to carry a steady-state flux. Its function will no longer be available to the network. EFMs therefore are minimal functional building blocks. We demonstrated in the introductory example that a complete set of EFMs fully defines a network. Every steady-state flux distribution can be represented as a weighted superposition of EFMs with non-negative weights.

The single most important advantage of EFMs is that every flux distribution can be decomposed into fundamental functional units without "cancelation" [1]. That is, if a flux through a reaction is zero, then each contributing EFM will be required to have a zero flux value in that reaction. This is in contrast to the closely related concept of extreme pathways [12], where fluxes can cancel out. Note that all extreme pathways are a subset of the EFMs and are therefore EFMs as well.

Similar to a Cartesian coordinate system, EFMs are direction vectors in a metabolic network. The weights represent the usage-coefficient of a direction vector, or – in analogy to the Cartesian coordinate system – are the coordinates. Unfortunately, such an expansion in EFMs is not unique. This is reminiscent of an organism's structural robustness, as cells typically developed several redundant pathways to preserve essential functionalities. In fact, EFMA has been used to quantify cellular robustness [13, 14].

While any steady-state flux distribution linearly decomposes into EFMs, tracer-based labeling patterns (typically observed in fluxomic-approaches [15]) in gener-

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

al do not. EFMA identifies independent pathways and does not track individual atoms. However, recently elementary carbon modes have been introduced [16], which, in analogy to EFMs, allow for all labeling patterns to be analyzed to reveal their "generating" metabolic pathways.

4 EFMA applied to metabolic engineering

The ability to express the full metabolic capabilities of a cell in terms of elementary functional units makes EFMA a powerful tool for metabolic engineering [17, 18]. EFMA can be harnessed to consider whether a cell is able to produce a metabolite of interest from a given substrate. If there is at least one EFM (or a combination of EFMs) which connects the substrate and the product of interest, the production strategy is possible. The (stoichiometric) efficiency of conversion is easily computable by following the chain of reactions in an EFM. By doing so, the mode with the maximum yield can be identified.

The aim of metabolic engineering is to channel available resources towards the engineering objective. EFMA identifies those desirable network functions without bias. Identifying desirable EFMs is only the first step in the metabolic engineering of an organism, and needs to be followed by the elimination of counterproductive network functions. When the EFMs are known, disabling a particular function or EFM is achieved by deleting a contributing reaction. Since an EFM is an indivisible, minimal steady-state pathway, by definition knocking out a single reaction is sufficient to remove the EFM and its steady-state functionality from the network. However, since a reaction typically does not only contribute to one but multiple EFMs, other functions will also be disabled upon deletion. Thus, the question arises whether it is possible to delete all unwanted network properties, while retaining all desirable states.

Carlson and Srienc [19, 20] were the first to develop an iterative procedure which identified a series of gene deletions to turn an organism into a network of minimal functionality, containing only desirable network properties. Carlson and Srienc [19, 20] quantified the impact of a single deletion on the number of unaffected EFMs and chose the deletion which removed the maximum number of EFMs without influencing production efficiency of their product of interest. The procedure was repeated until no further deletion was possible. The metabolic engineering strategy so designed was successfully tested experimentally [21]. Since then, this approach has proven successful in several applications [22–25]. However, the procedure does not guarantee that the final design is obtained with the least possible effort [26].

Recently, constrained minimal cut sets have been introduced to rigorously address the problem of finding minimal deletion strategies that achieve a given engineering objective [26]. This approach is not only able to predict the minimal number of necessary metabolic interventions, but also exhaustively enumerates all possible combinations of deletions which result in networks with identical functionalities. Minimal cut sets are (minimal) sets of deletions, which block undesirable network functionality, such as the secretion of unwanted by-products. Constrained minimal cut sets allow desirable network properties to be retained while simultaneously disabling unwanted functionalities [26]. The calculation of these cut sets may also be formulated as an optimization problem, which allows including regulatory constraints [11].

Over-expressions are a most important tool in the arsenal of cellular engineers. Melzer et al. [27] developed an EFM-based method, called FluxDesign, which identifies deletion, knockdown and over-expression targets. After calculating all EFMs of the network, the reactions are screened for statistically significant correlation with a predefined target reaction, such as the production of a metabolite of interest. If the flux values of all EFMs for one reaction positively correlate with the flux values of the target reaction, then the corresponding gene should be over-expressed; conversely, this gene should be knocked-down or deleted in case of a negative correlation. The approach was validated experimentally by investigating lysine production in *Corynebacterium glutamicum* [28]. FluxDesign is a heuristic approach, and has been shown to identify many known metabolic engineering targets [27].

EFMA has been successfully applied to relatively small metabolic reconstructions (up to 100 reactions). For larger systems, the calculation of the EFMs represents a major challenge to this approach and hinders its wider application. In the following section we therefore review current calculation approaches and their applicability to larger, genome-scale metabolic models (GSMM).

5 Calculating EFMs

5.1 Standard approaches

Several tools are available to calculate EFMs, either as standalone software (for example "efmtool" [29] and Metatool [30]) or integrated in various "general purpose" mathematical software [31–34]. These methods are all variants of the double description method [35]. This algorithm successively generates EFM-candidates by pairwise combination of existing EFMs followed by verification that each EFM candidate has not been previously identified. This verification procedure is the major bottleneck of the double description method. Several algorithmic improvements have been published, including a binary representation to relax the memory requirements [6] and accelerating the testing of EFM-candidates via matrix rank [36] and (recursive) bit pattern trees [29, 37]. Efmtool [29] has so far been shown to be the best performing tool on a single processor machine [17, 38].

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The most critical infrastructure requirements for all these implementations are computers with high amounts of primary data storage available for the computation process. For instance, we were only able to run an EFMA of the full *E. coli* -core model [10] using a machine with at least 153 GB of RAM. Efmtool [29] does not support distributed memory parallelization and hence relies heavily on machines that are equipped with large amounts of RAM. To circumvent this problem, a distributed memory parallelization has been introduced recently, with promising results [38]. However, even the power of high performance computing does not solve the essential problem of EFMA, which is the combinatorial explosion of the number of EFMs with network size. Currently, EFMA is restricted to medium-scale reconstructions with several hundred million EFMs. Thus, a complete EFMA of current GSMMs, containing typically a few thousand reactions, is unrealistic using standard methods.

5.2 Recent ideas and developments

Several approaches have been developed to deal with the combinatorial explosion generated by GSMMs. Recently we used Boolean transcriptional regulatory networks to restrict the number of biologically feasible EFMs [39]. The same approach has been shown to reduce the solution space and improve the accuracy of flux balance analysis predictions [40, 41]. EFMA is solely based on the (steady-state) topology of a metabolic network and disregards any other biologically relevant constraints, such as transcriptional regulation. For instance, in *E. coli* the glyoxylate shunt is repressed during growth on glucose [42, 43]. Although representing a feasible steady-state solution, any EFM in which both glucose uptake and the glyoxylate shunt are active is not biologically feasible. Thus, by including Boolean transcriptional regulatory networks, millions of EFMs can be eliminated on the basis of biological regulatory contraints, thereby considerably reducing the computational costs. In *E. coli* we showed that 99% of 272 million EFMs are not biologically feasible due to transcriptional regulatory constraints. However, Boolean transcriptional networks are currently only readily available for *E. coli* and *Saccharomyces cerevisiae*. Additionally, it is not yet clear whether the constraints set due to Boolean regulation might be too tight, leading to the suppression of too many EFMs from the analysis.

Jol et al. [44] made the observation that 54% of the EFMs in their yeast model were thermodynamically infeasible. They coupled network embedded thermodynamics [45] with metabolomics and EFMA to derive this result. Their classification was done *ex post facto*, i.e. after all EFMs, whether thermodynamically feasible or not, were calculated. It is not clear if their method might also help to avoid the calculation of thermodynamically infeasible modes altogether. Nonetheless, both examples show that the number of biologically feasible modes may be dramatically smaller than expected by ordinary EFMA. Similar findings were reported in other constraint-based approaches like flux balance analysis as well [46–48].

An alternative idea is to isolate subsystems from the parent network [49, 50] and calculate EFMs only for the subsystems of interest [4]. However, simply cutting out subnetworks adulterates the principle of EFMA for two reasons: (i) in most cases, isolating subsystems from larger models is biased by the (modeler's) choice of the (sub-)system's boundaries; and (ii) due to possible dependencies outside of the isolated subsystem, EFMs predicted for the subnetwork may not necessarily be part of any EFM of the full system, and vice versa. Kaleta et. al [51] resolved these problems by proposing the concept of EFM patterns.

An EFM pattern is the Boolean representation of all reactions of an EFM, which sit inside the subsystem of interest. These patterns can be calculated and analyzed even without fully calculating all EFMs of the lager model. Each EFM pattern corresponds to at least one EFM of the full system. EFMs are specific flux distributions. EFM patterns on the other hand, are their binary representation in the region of interest. Therefore, many different EFMs can be mapped onto the same EFM pattern. It is possible to calculate all EFMs for a given EFM pattern– although computationally very demanding to do so.

EFM patterns are calculated by optimizing a mixed-integer linear problem. A similar mathematical framework can be used to successively calculate EFMs in increasing order of the number of participating reactions [52]. This method – although less efficient than the double description method – may also be used to enumerate all EFMs for any model. However, a mixed-integer approach is in itself computationally demanding. Additionally, due to the strict sequential nature of this approach, it does not qualify for simple parallelization.

From a technical standpoint, knowing the shortest pathway to a product of interest is advantageous, as this potentially minimizes the (genetic) engineering effort. However, it has been argued that evolutionarily optimized metabolic systems naturally evolve towards the state of maximum entropy production [53–55]. In this state, the contribution of an EFM to the total flux is weighted by its reaction entropy distributed according to the Boltzmann distribution law. Whether or not the number of participating reactions of an EFM correlates with its entropy is yet to be shown.

Due to the large number of EFMs in a GSMM it becomes more difficult to analyze and identify the biological function of an EFM. To deal with this complexity, an unsupervised clustering method has been developed [56] which assigns the group membership of an EFM according to the set of common reactions. However, depending on the application it may not always be necessary to calculate all EFMs in order to sufficiently characterize a network.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Kaleta et al. [57] sampled EFMs with the help of a genetic algorithm. To find EFMs, they knocked out reactions (i.e. set zeros in the flux vector) and calculated EFMs, which are consistent with the given set of deletions using linear programing. In principle, this would allow the calculation of the complete set of EFMs if all possible combinations of deletions were tested, but for real world systems the number of combinations is prohibitively large. Therefore, the authors chose to randomly initialize a population of flux vectors with zeros and used a genetic algorithm to generate offspring-flux vectors for which associated EFMs could be calculated even in a GSMM [57]. Computationally this approach allows for a very easy parallelization. However, all EFMs have to be filtered in order to detect duplications, an approach which progressively slows down the calculation of new EFMs.

Very recently an alternative sampling procedure was published by Machado et al. [58], who used the standard double description approach for EFM enumeration. However, rather than evaluating all EFM candidates, Machado et al. [58], proceeded with randomly selected EFM candidates to avoid explosion in testable combinations. The selection probability can be controlled with a single parameter. If the selection probability tends to one, the full system is recovered. Provided that this parameter is sufficiently large, the authors were able to show that the EFM spectrum was well sampled [58]. However, it remains to be seen if this sampling approach allows the unbiased analysis of GSMMs as well, since a high selection probability may once again lead to a combinatorial explosion in the number of calculated EFMs.

If the aim of an analysis is to understand particular steady-state flux distributions in terms of their functional building blocks, it can be argued that calculation of all EFMs is a computational overkill [59]. Ip et al. [59] developed a quick decomposition which only computes EFMs needed to represent the given flux distribution. Ip et al. [59] calculated an EFM which contained the reaction with the highest flux. Next, the obtained EFM is subtracted from the flux distribution and the procedure repeated until no further EFM can be found. This procedure will always terminate and is applicable to GSMMs, since typically only a small fraction of the total number of EFMs will be found. Chan and Ji [60] developed a similar method. However, decomposition into EFMs is not unique [61, 62]. In fact, the expansion is strongly dependent on the particular choices of EFM made in the iteration procedure, which calls into question the biological significance of the resulting decomposition [61]. Despite this limitation, this methodology allows significant additional functional insight into a given flux distribution. This was demonstrated by analyzing an acetate over-producing *E. coli* strain [59] and the study of the growth of *E. coli* in a complex medium [60].

6 Concluding remarks

EFMA is a valuable tool for metabolic engineering. However, in order to fully utilize the potential of EFMA, methodological improvements are essential: to date EFMA has only been applied to small and medium-scale networks. Two approaches for further method advances are currently investigated: (i) complete enumeration of biologically feasible EFMs ; and (ii) sampling the available EFM solution space. The first approach aims to constrain the available solution space by including additional thermodynamic and/or biological data. The second approach does not aim to fully enumerate the solution space but calculate only those EFMs with particular structural properties. The latter approach has already been shown to work with GSMMs [63]. However, the biological relevance of the calculated EFMs must be further investigated. It remains to be seen whether the knowledge of some important EFMs is already sufficient for targeted, rational metabolic engineering.

If all EFMs are available, optimal engineering strategies can be developed based on the concept of networks of minimal functionalities and constrained minimal cut sets. However, while there is sufficient evidence that standard EFMA dramatically overestimates biologically accessible network states, a clear and computationally manageable strategy to suppress these surplus modes is yet to be developed.

The Austrian BMWFJ, BMVIT, SFG, Standortagentur Tirol and ZIT through the Austrian FFG-COMET-Funding Program supported this work.

The authors declare no conflict of interest.

7 References

[1] Schuster, S., Hilgetag, C., Woods, J. H., Fell, D. A., Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J Math Biol.* 2002, *45*, 153–181.

[2] Schuster, S., Fell, D. A., Dandekar, T., A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat. Biotechnol.* 2000, *18*, 326–332.

[3] Schuster, S., Dandekar, T., Fell, D. A., Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol.* 1999, *17*, 53–60.

[4] Schuster, S., Hilgetag, C., On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* 1994, *2*, 165–182.

[5] Acuña, V., Chierichetti, F., Lacroix, V., Marchetti-Spaccamela, A. et al., Modes and cuts in metabolic networks: Complexity and algorithms. *Biosystems*, 2009, *95*, 51–60.

[6] Gagneur, J., Klamt, S., Computation of elementary modes: A unifying framework and the new binary approach. *BMC Bioinformatics* 2004, *5*, 175.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[7] Jevremovic, D., Trinh, C. T., Srienc, F., Boley, D., On algebraic properties of extreme pathways in metabolic networks. *J. Comput. Biol.* 2010, *17*, 107–119.
[8] Acuña, V., Marchetti-Spaccamela, A., Sagot, M. F., Stougie, L., A note on the complexity of finding and enumerating elementary modes. *Biosystems* 2010, *99*, 210–214.
[9] Klamt, S., Stelling, J., Combinatorial complexity of pathway analysis in metabolic networks. *Mol. Biol. Rep*, 2002, *29*, 233–236.
[10] Orth, J. D., Fleming, R. M. T., Palsson, B. Ø., Chapter 10.2.1 – reconstruction and use of microbial metabolic networks: the core *Escherichia coli* metabolic model as an educational guide, in: Karp, P. D. (Ed.), *EcoSal – Escherichia coli and Salmonella Cellular and Molecular Biology*, ASM Press, Washington DC 2009.
[11] Jungreuthmayer, C., Zanghellini, J., Designing optimal cell factories: Integer programing couples elementary mode analysis with regulation. *BMC Syst. Biol.* 2012, *6*, 103.
[12] Schilling, C. H., Letscher, D., Palsson, B. Ø., Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J Theor. Biol.* 2000, *203*, 229–248.
[13] Stelling, J., Klamt, S., Bettenbrock K., Schuster, S. et al., Metabolic network structure determines key aspects of functionality and regulation. *Nature* 2002, *420*, 190–193.
[14] Behre, J., Wilhelm, T., von Kamp, A., Ruppin E. et al., Structural robustness of metabolic networks with respect to multiple knockouts. *J. Theor. Biol.* 2008, *252*, 433–441.
[15] Sauer, U., Metabolic networks in motion: ^{13}C-based flux analysis. *Mol. Syst. Biol.* 2006, *2*, 62.
[16] Pey, J., Theodoropoulos, C., Rezola, A., Rubio, A. et al., Do elementary flux modes combine linearly at the atomic level? Integrating tracer-based metabolomics data and elementary flux modes. *Biosystems* 2011, *105*, 140–146.
[17] Trinh, C. T., Thompson, R. A., Elementary mode analysis: A useful metabolic path-way analysis tool for reprograming microbial metabolic pathways, in Wang, X., Chen, J., Quinn, P. (Eds.), *Reprogramming Microbial Metabolic Pathways*, volume 64, Springer Netherlands, Dordrecht 2012, pp. 21–42.
[18] Trinh, C. T., Wlaschin, A., Srienc, F., Elementary mode analysis: A useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl. Microbiol. Biotechnol.* 2009, *81*, 813–826.
[19] Carlson, R., Srienc, F., Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: Identification of reactions. *Biotechnol. Bioeng.* 2004, *85*, 1–19.
[20] Carlson, R., Srienc, F., Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: Creation of overall flux states. *Biotechnol. Bioeng.* 2004, *86*, 149–162.
[21] Trinh, C.T., Carlson, R., Wlaschin, A., Srienc, F., Design, construction and performance of the most efficient biomass producing *E. coli* bacterium. *Metab. Eng.* 2006, *8*, 628–638.
[22] Unrean, P., Trinh, C. T., Srienc, F., Rational design and construction of an efficient *E. coli* for production of diapolycopendioic acid. *Metab. Eng.* 2010, *12*, 112–122.
[23] Trinh, C.T., Srienc, F., Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. *Appl. Environ. Microbiol.* 2009, *75*, 6696–6705.
[24] Trinh, C.T., Unrean, P., Srienc, F., Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Appl Environ Microbiol.* 2008, *74*, 3634–3643.
[25] Trinh, C. T., Li, J., Blanch, H. W., Clark, D. S., Redesigning *Escherichia coli* metabolism for anaerobic production of isobutanol. *Appl. Environ. Microbiol.* 2011, *77*, 4894–4904.
[26] Hädicke, O., Klamt, S., Computing complex metabolic intervention strategies using constrained minimal cut sets. *Metab. Eng.* 2011, *13*, 204–213.

Jürgen Zanghellini, received his PhD from the Vienna University of Technology in 2004. He is currently heading a group on metabolic modeling at the Austrian Centre of Industrial Biotechnology (ACIB). His current research interests are geared toward numerical methods in biology with a special focus on computational systems biology.

Christian Jungreuthmayer, received his PhD from the Vienna University of Technology in 2005. He is a scientist at the Austrian Centre of Industrial Biotechnology (ACIB). His research background is high performance computing and metabolic modeling. His current research interests include numerical modeling and computational simulations of metabolic networks.

[27] Melzer, G., Esfandabadi, M. E., Franco-Lara, E., Wittmann, C., Flux Design: *In silico* design of cell factories based on correlation of pathway fluxes to desired properties. *BMC Syst. Biol.* 2009, *3*, 120.
[28] Becker, J., Zelder, O., Häfner, H., Schröder, S. et al., From zero to hero – design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. *Metab. Eng.* 2011, *13*, 159–168.
[29] Terzer, M., Stelling, J., Large-scale computation of elementary flux modes with bit pattern trees. *Bioinformatics* 2008, *24*, 2229–2235.
[30] von Kamp, A., Schuster, S., Metatool 5.0: Fast and flexible elementary modes analysis. *Bioinformatics* 2006, *22*, 1930–1931.
[31] Rocha, I., Maia, P., Evangelista, P., Vila¸ca, P. et al., OptFlux: An open-source software platform for *in silico* metabolic engineering. *BMC Syst. Biol.* 2010, *4*, 45.
[32] Hoops, S., Sahle, S., Gauges, R., Lee, C. et al., COPASI – a COmplex PAthway SImulator. *Bioinformatics* 2006, *22*, 3067–3074.
[33] Klamt, S., Saez-Rodriguez, J., Lindquist, J. A., Simeoni, L. et al., A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics* 2006, *7*, 56.
[34] Schwarz, R., Musch, P., von Kamp, A., Engels, B. et al., YANA a software tool for analyzing flux modes, gene-expression and enzyme activities. *BMC Bioinformatics* 2005, *6*, 135.
[35] Fukuda, K., Prodon, A., Double description method revisited, in: Deza, M., Euler, R., Manoussakis, I., (Eds.), *Combinatorics and Computer Science*, *Volume 1120 of Lecture Notes in Computer Science*, Springer, 1996, pp. 91–111.
[36] Klamt, S., Gagneur, J., von Kamp, A., Algorithmic approaches for computing elementary modes in large biochemical reaction networks. *IEE Proceedings – Systems Biology* 2005, *152*, 249–255.
[37] Terzer, M., Stelling, J., Accelerating the computation of elementary modes using pattern trees, in: Bücher, P., and Bernard Moret (Eds.), *Algorithms in Bioinformatics*, *Volume 4175 of Lecture Notes in Computer Science*, Springer, 2006, pp. 333–343.
[38] Jevremović, D., Trinh, C.T., Srienc, F., Sosa, C.P. et al., Parallelization of nullspace algorithm for the computation of metabolic pathways. *Parallel Comput.* 2011, *37*, 261–278.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[39] Jungreuthmayer, C., Ruckerbauer, D. E., Zanghellini, J., regEfmtool: Speeding up elementary flux mode calculation using transcriptional regulatory rules in the form of three-state logic. *Biosystems* 2013, *113*, 37–39.
[40] Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M.J. et al., Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, 2004, *429*, 92–96.
[41] Covert, M. W., and Palsson, B., Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J Biol Chem*. 2002, *277*, 28058–28064.
[42] Kornberg, H. L., The role and maintenance of the tricarboxylic acid cycle in *Escherichia coli*. *Biochem. Soc. Symp.* 1979, *30*, 155–171.
[43] Kornberg, H. L., The role and control of the glyoxylate cycle in *Escherichia coli*. *Biochem. J.* 1966, *99*, 1–11.
[44] Jol, S. J., Kümmel, A., Terzer, M., Stelling, J. et al., System-level insights into yeast metabolism by thermodynamic analysis of elementary flux modes. *PLoS Comput. Biol.* 2012, *8*, e1002415.
[45] Kümmel, A., Panke, S., Heinemann, M., Putative regulatory sites unraveled by networkem-bedded thermodynamic analysis of metabolome data. *Mol. Syst. Biol.* 2006, *2*, 2006.0034.
[46] Zhu, Y., Song, J., Xu, Z., Sun, J. et al., Development of thermodynamic optimum searching (TOS) to improve the prediction accuracy of flux balance analysis. *Biotechnol. Bioeng.* 2013, *110*, 914–923.
[47] Hoppe A., Hoffmann, S., Holzhtter, H. G., Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Syst. Biol.* 2007, *1*, 23.
[48] Henry, C. S., Broadbelt, L. J., Hatzimanikatis, V., Thermodynamics-based metabolic flux analysis. *Biophys. J.* 2007, *92*, 1792–1805.
[49] Verwoerd, W. S., A new computational method to split large biochemical networks into coherent subnets. *BMC Syst. Biol.* 2011, *5*, 25.
[50] Schuster, S., Pfeiffer, T., Moldenhauer, F., Koch, I. et al., Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae*. *Bioinformatics* 2002, *18*, 351–361.
[51] Kaleta, C., de Figueiredo, L. F., Schuster, S., Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. *Genome Res.* 2009, *19*, 1872–1883.
[52] De Figueiredo, L. F., Podhorski, A., Rubio, A., Kaleta, C. et al., Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics*, 2009, *25*, 3158–3165.
[53] Unrean, P., Srienc, F., Predicting the adaptive evolution of *thermoanaerobacterium saccharolyticum*. *J. Biotechnol*. 2012, *158*, 259–266.
[54] Unrean, P., Srienc, F., Metabolic networks evolve towards states of maximum entropy production. *Metab. Eng.* 2011, *13*, 666–673.
[55] Srienc, F., Unrean, P., A statistical thermodynamical interpretation of metabolism. *Entropy* 2010, *12*, 1921–1935.
[56] Pérès, S., Vallée, F., Beurton-Aimar, M., Mazat. J. P., ACoM: A classification method for elementary flux modes based on motif finding. *Biosystems* 2011, *103*, 410–419.
[57] Kaleta, C., de Figueiredo, L. F., Behre, J., Schuster, S., EFMEvolver: Computing elementary flux modes in genome-scale metabolic networks. *Proceedings of the 14th German Conference on Bioinformatics (GCB)* 2009, *2*, 180–190.
[58] Machado, D., Soons, Z., Patil, K. R., Ferreira, E. C. et al., Random sampling of elementary flux modes in large-scale metabolic networks. *Bioinformatics* 2012, *28*, i515–i521.
[59] Ip, K., Colijn, C., Lun, D., Analysis of complex metabolic behavior through pathway decomposition. *BMC Syst. Biol.* 2011, *5*, 91.
[60] Chan, S. H. J., Ji, P., Decomposing flux distributions into elementary flux modes in genome- scale metabolic networks. *Bioinformatics* 2011, *27*, 2256–2562.
[61] Schwartz, J. M., Kanehisa, M., Quantitative elementary mode analysis of metabolic pathways: The example of yeast glycolysis. *BMC Bioinformatics* 2006, *7*, 1–20.
[62] Schwartz, J. M., Kanehisa, M., A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes. *Bioinformatics* 2005, *21*, ii204–ii205.
[63] Gebauer, J., Schuster, S., de Figueiredo, L. F., Kaleta, C., Detecting and investigating substrate cycles in a genome-scale human metabolic network. *FEBS J.* 2012, *279*, 3192–3202.

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Special issue: *Metabolic Modeling and Simulation*. Modeling of cellular metabolism has been a major area of research for bioengineers and biomedical researchers alike. This Special Issue collects a series of articles on methods of metabolic modeling, modeling of human metabolism, modeling of microbial metabolism and modeling of bioprocesses. This cover is a visual representation of the essence of metabolic engineering. Image: © rolffimages – Fotolia.com.

Biotechnology Journal – list of articles published in the September 2013 issue.

Editorial: Metabolic modeling in biotechnology and medical research
Diethard Mattanovich and Vassily Hatzimanikatis
http://dx.doi.org/10.1002/biot.201300378

Review
Multi-scale modeling for sustainable chemical production
Kai Zhuang, Bhavik R. Bakshi and Markus J. Herrgård
http://dx.doi.org/10.1002/biot.201200272

Review
Genome-scale modeling of human metabolism – a systems biology approach
Adil Mardinoglu, Francesco Gatto and Jens Nielsen
http://dx.doi.org/10.1002/biot.201200275

Review
Basic concepts and principles of stoichiometric modeling of metabolic networks
Timo R. Maarleveld, Ruchir A. Khandelwal, Brett G. Olivier, Bas Teusink and Frank J. Bruggeman
http://dx.doi.org/10.1002/biot.201200291

Mini-Review
Elementary flux modes in a nutshell: Properties, calculation and applications
Jürgen Zanghellini, David E. Ruckerbauer, Michael Hanscho and Christian Jungreuthmayer
http://dx.doi.org/10.1002/biot.201200269

Review
Predicting complex phenotype–genotype interactions to enable yeast engineering: *Saccharomyces cerevisiae* as a model organism and a cell factory
Duygu Dikicioglu, Pınar Pir and Stephen G. Oliver
http://dx.doi.org/10.1002/biot.201300138

Technical Report
Flux-coupled genes and their use in metabolic flux analysis
Hyun Uk Kim, Won Jun Kim and Sang Yup Lee
http://dx.doi.org/10.1002/biot.201200279

Research Article
Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints
Anirikh Chakrabarti, Ljubisa Miskovic, Keng Cher Soh and Vassily Hatzimanikatis
http://dx.doi.org/10.1002/biot.201300091

Research Article
Metabolic gradients as key regulators in zonation of tumor energy metabolism: A tissue-scale model-based study
Matthias König, Hermann-Georg Holzhütter and Nikolaus Berndt
http://dx.doi.org/10.1002/biot.201200393

Research Article
Genomically and biochemically accurate metabolic reconstruction of *Methanosarcina barkeri* Fusaro, iMG746
Matthew C. Gonnerman, Matthew N. Benedict, Adam M. Feist, William W. Metcalf and Nathan D. Price
http://dx.doi.org/10.1002/biot.201200266

Research Article
Kinetic isotope effects significantly influence intracellular metabolite ^{13}C labeling patterns and flux determination
Thomas M. Wasylenko and Gregory Stephanopoulos
http://dx.doi.org/10.1002/biot.201200276

Research Article
Optimization-driven identification of genetic perturbations accelerates the convergence of model parameters in ensemble modeling of metabolic networks
Ali R. Zomorrodi, Jimmy G. Lafontaine Rivera, James C. Liao and Costas D. Maranas
http://dx.doi.org/10.1002/biot.201200270

Research Article
Metabolic costs of amino acid and protein production in *Escherichia coli*
Christoph Kaleta, Sascha Schäuble, Ursula Rinas and Stefan Schuster
http://dx.doi.org/10.1002/biot.201200267

© 2013 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

18607314, 2013, 9, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/biot.201200269, Wiley Online Library on [23/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License