Systems biology

Functional stoichiometric analysis of metabolic networks

R. Urbanczik* and C. Wagner

Institute of Pharmacology, University of Bern, Friedbühlstrasse 49, CH-3010 Bern, Switzerland

ABSTRACT

Motivation: An important tool in Systems Biology is the stoichiometric modeling of metabolic networks, where the stationary states of the network are described by a high-dimensional polyhedral cone, the so-called flux cone. Exhaustive descriptions of the metabolism can be obtained by computing the elementary vectors of this cone but, owing to a combinatorial explosion of the number of elementary vectors, this approach becomes computationally intractable for genome scale networks.

Result: Hence, we propose to instead focus on the conversion cone, a projection of the flux cone, which describes the interaction of the metabolism with its external chemical environment. We present a direct method for calculating the elementary vectors of this cone and, by studying the metabolism of *Saccharomyces cerevisiae*, we demonstrate that such an analysis is computationally feasible even for genome scale networks.

Contact: robert.urbanczik@pki.unibe.ch

1 INTRODUCTION

Thanks to the abundance of genome data, it has become possible to reconstruct chemical reaction networks that comprehensively model the metabolism of microorganisms such as *Escherichia coli* and *Saccharomyces cerevisiae* (Reed *et al*., 2003; Duarte *et al*., 2004). In the stoichiometric network analysis of such models, the possible flows through the network are constrained by the requirement that in the long term there be no net accumulation or depletion of any chemical compound that cannot cross the cell boundary. Since typically many of the reactions are irreversible, the space of possible flows through the network is further constrained, and in mathematical terms corresponds to a high-dimensional polyhedral cone, which is often called the flux cone.

In Flux Balance Analysis (Ibarra *et al*., 2002) predictions about the response of the microorganism to changing environmental conditions or even to the deletion of certain genes are obtained by using linear programming to study this cone. But not all relevant questions readily fit into the framework of linear or even convex programming. A case in point would be the enumeration of all minimal media able to sustain the microorganism. Hence a second approach to studying chemical reaction networks has been to obtain an exhaustive description of the flux cone by calculating its elementary vectors (Stelling *et al*., 2002; Schuster *et al*., 2002a), also called elementary flux modes. Based on a complete set of these, many conclusions about the metabolism can easily be drawn since each

elementary vector represents a unique, stoichiometrically viable and non-redundant pathway through the entire network. However, the price to pay for this exhaustive description by elementary fluxes is the combinatorial explosion of their number (Klamt and Stelling, 2002) and the approach has therefore been limited to networks that are an order of magnitude smaller than, for example, the full genome-scale network of *S.cerevisiae*.

To achieve genome-scale, we shall adopt a functional perspective and focus on describing the effect of the microorganism on its external chemical environment, for instance, the consumption of energy sources such as glucose and the production of waste such as CO_2. In disregarding the internal mechanism, our concept of conversions between external compounds, which can be effected by the metabolism, is similar to the overall reactions studied by Happel and Sellers (1989). But, crucially, we do not assume all reactions to be reversible. Hence the set of possible conversions in general is not a vector space but a polyhedral cone. While this conversion cone is a projection of the flux cone, we shall show how the elementary vectors of the conversion cone can be found, without first having to compute the flux cone. Using this algorithm, elementary conversion analysis becomes computationally tractable for networks far larger than the ones considered in the elementary flux analysis. This is demonstrated by studying a conversion cone associated with the above mentioned metabolic network of *S.cerevisiae*.

2 BASIC CONCEPTS

In stoichiometric analysis (Clarke, 1980; Heinrich and Schuster, 1996; Price *et al*., 2004; Stucki, 2004) a metabolic network is modeled by an m by n stoichiometry matrix S, which relates the flows ν through the n reactions to changes $\dot{c}$ in the concentrations c of the m metabolites by $\dot{c} = S\nu$. Typically many of the reactions will be irreversible and this means that the flow ν_j through such a reaction cannot be negative. Further, one can often assume that the concentration levels of some (external) metabolites, such as CO_2, are maintained by the environment or large enough that the changes caused by the reaction system become negligible. For many metabolites, however, this will not be the case and hence in a stationary state of the network $\dot{c}_i = 0$ most hold for such an internal metabolite. Thus the possible steady states of the metabolic model are required to satisfy

$$
\begin{aligned}
&\text{I.} \quad && \dot{c} = S\nu, \\
&\text{II.} \quad && \nu_j \geq 0 \quad \text{if } j \in \mathsf{Irrev}, \\
&\text{III.} \quad && \dot{c}_j = 0 \quad \text{if } i \in \mathsf{Int},
\end{aligned}
\tag{1}
$$

where Irrev denotes the subset of irreversible reactions and Int the subset of internal metabolites. Sometimes one further wants to

*To whom correspondence should be addressed.

subclassify some of the external metabolites into inputs ($\dot{c}_i \leq 0$) and outputs ($\dot{c}_i \geq 0$). While such a distinction can easily be incorporated into the formal framework given below, for brevity, we shall not explicitly consider this here.

The above notation is a bit unusual since we do not incorporate exchange fluxes into the stoichiometry matrix S. Often the system is augmented by pseudo reactions ($C_i \leftrightarrow 0$ for $i \notin \mathsf{Int}$) for the external metabolites, and the stoichiometry matrix of the augmented system is denoted by N. Then a steady state flux vector J of the augmented system must satisfy $NJ = 0$ and the irreversibility conditions. But, if one orders the columns in N appropriately, any pair $(\dot{c}, \nu)$ satisfying Equation (1) corresponds to the steady-state flux vector $J = (\dot{c}_{|\mathsf{Ext}}, \nu)$, where $\dot{c}_{|\mathsf{Ext}}$ denotes the restriction of $\dot{c}$ to the external metabolites. Since condition III means that $\dot{c}_{|\mathsf{Int}} = 0$, this correspondence is easily seen to be a bijection, and the two notations are entirely equivalent.

The set of all pairs $(\dot{c}, \nu)$ satisfying the three conditions in Equation (1) is a polyhedral cone, and we shall call it the flux cone $\mathcal{F}$. Our notation highlights that the two components of the pair $(\dot{c}, \nu)$ provide quite different information on the stationary state of the network. Whereas ν tells us how the network is operating in terms of the flows through its reactions, $\dot{c}$ describes what effect the networks operation has on the environment, in terms of the conversions taking place between the external metabolites. In order to separate the 'what' from the 'how', we now introduce the conversion cone $\mathcal{C}$ as the set of all vectors $\dot{c}$, which can be part of a stationary flux $(\dot{c}, \nu)$. Formally

$$\mathcal{C} = \{\dot{c} \mid \text{There is a } \nu \text{ such that } (\dot{c}, \nu) \in \mathcal{F}\}. \quad (2)$$

The definition of the conversion cone in Equation (2) is rather oblique and our main goal will be to find a more explicit description of this cone in terms of elementary conversions. We call a non-zero vector e an elementary vector of $\mathcal{C}$ or simply an elementary conversion if e lies, for some $s \in \{-1, 1\}^m$, on an edge of the pointed cone $\mathcal{C} \cap \mathcal{O}_s$, where $\mathcal{O}_s$ is the orthant:

$$\mathcal{O}_s = \{x \in \mathbb{R}^m \mid s_i x_i \geq 0 \quad \text{for } i = 1, \ldots, m\}. \quad (3)$$

The concept is illustrated in Figure 1. We call a set of elementary conversions complete if for any orthant $\mathcal{O}_s$ each edge of the intersection cone $\mathcal{C} \cap \mathcal{O}_s$ is represented by exactly one vector in this set.

The importance of elementary conversions stems from the following observation. Any vector $\dot{c}$ of the conversion cone will lie in at least one orthant $\mathcal{O}_s$ and can hence be decomposed into vectors $e^{(l)}$ lying on the edges of $\mathcal{C} \cap \mathcal{O}_s$ as

$$\dot{c} = e^{(1)} + e^{(2)} + \cdots + e^{(k)}. \quad (4)$$

Since the $e^{(l)}$ are edge vectors of $\mathcal{C} \cap \mathcal{O}_s$, in the matrix $(e^{(1)}, e^{(2)}, \ldots, e^{(k)})$ each row is either semi-positive or semi-negative and no cancellation can occur during the addition in Equation (4). Of course, by definition, the $e^{(l)}$ are elementary. So we have established that any vector in $\mathcal{C}$ can be decomposed into elementary conversions without cancellation and this implies that many properties of the conversion cone can be found by simply inspecting the elementary vectors. For instance, a statement such as 'consumption of glucose must be accompanied by the production of ethanol and CO_2' will hold for any stationary state if and only if it is true for all elementary conversions.

In passing we note that one can define elementary vectors of the flux cone $\mathcal{F}$ analogously by requiring that they lie on an
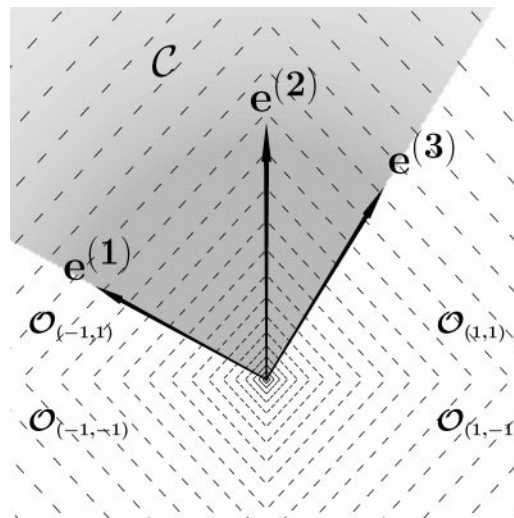

Fig. 1. Elementary vectors of the cone $\mathcal{C}$ (grey). The vectors $e^{(2)}$ and $e^{(3)}$ are elementary because they lie on the edges of the intersection of $\mathcal{C}$ with the semi-positive orthant $\mathcal{O}_{(1,1)}$. Intersecting $\mathcal{C}$ with $\mathcal{O}_{(-1,1)}$ shows that $e^{(1)}$ and $e^{(2)}$ are elementary. Since the intersection of $\mathcal{C}$ with $\mathcal{O}_{(-1,-1)}$, as well as with $\mathcal{O}_{(1,-1)}$, is empty, the only elementary vectors of $\mathcal{C}$ are, upto positive multiples, $e^{(1)}, e^{(2)}$ and $e^{(3)}$.

edge of $\mathcal{F} \cap \mathcal{O}$, where $\mathcal{O}$ is an orthant. While at first glance this definition differs from the one used in Schuster *et al.* (2002a) and Rockafellar (1970), where elementary vectors are defined by having a maximal set of zeroes, in Supplementary Material A.4 (Lemma 2) we show that the two definitions are in fact completely equivalent for flux cones. Hence, any elementary conversion is the projection of at least one elementary flux.

3 CALCULATING ELEMENTARY CONVERSIONS

We have recently presented the Nullspace algorithm (Urbanczik and Wagner, 2005), a double description method (Fukuda and Prodon, 1995) tailored to the calculation of elementary fluxes. The procedure first obtains a representation, in terms of linear inequalities, of a so-called coordinate cone isomorphic to the flux cone $\mathcal{F}$. Then, based on these inequalities, a generating set of the coordinate cone is calculated, i.e. a set such that the entire coordinate cone can be obtained by taking all linear combinations of the vectors in this set, involving only non-negative scalar factors. In fact, the generating set found is minimal, so no proper subset generates the cone. Then, based on these generating vectors, the algorithm finds a complete set of elementary fluxes. Here, we shall use similar ideas to obtain all elementary conversions, without first having to calculate the elementary fluxes.

Our starting point is the cone $\mathcal{C}_0$ of all vectors satisfying the first two conditions in Equation (1),

$$\mathcal{C}_O = \{S\nu | \nu_j \geq 0 \quad \text{if } j \in \mathsf{Irrev}\} \quad (5)$$

Obtaining the conversion cone $\mathcal{C}$, i.e. the vectors in $\mathcal{C}_0$ that also satisfy the condition III in Equation (1), would be quite easy if we had a different representation of $\mathcal{C}_0$. In particular, assume that we have an inequalities representation of $\mathcal{C}_0$, that is, a set of vectors

$H \subset \mathbb{R}^m$ such that

$$\mathcal{C}_0 = \{\dot{c} \in \mathbb{R}^m \mid h \cdot \dot{c} \geq 0 \quad \text{for all } h \in H\}. \tag{6}$$

Then, by writing the equality constraints in condition III of Equation (1) as inequalities, we immediately also have the following inequalities representation of the conversion cone,

$$\mathcal{C} = \{\dot{c} \in \mathbb{R}^m \mid h \cdot \dot{c} \geq 0 \quad \text{for all } h \in H,$$
$$\dot{c}_i \geq 0 \quad \text{and} \quad -\dot{c}_i \geq 0 \quad \text{for } i \in \mathsf{Int}\}. \tag{7}$$

From this inequalities representation of $\mathcal{C}$ first a minimal generating set and then all elementary conversions can be calculated using minor modifications of the Nullspace algorithm described in Supplementary Material A.

We still need to specify how to find the set H representing $\mathcal{C}_0$ by inequalities. For this, we first note that using the definition in Equation (5) a, possibly redundant, generating set of $\mathcal{C}_0$ is easily obtained from the stoichiometry matrix. If there are no reversible reactions, the columns $S^{(j)}$ of S are already a generating set and for reversible reactions we just need to take the reverse direction into account as well. So, a generating set of $\mathcal{C}_0$ is

$$G = \{S^{(j)}\}_{j=1}^n \cup \{-S^{(j)}\}_{j \notin \mathsf{Irrev}}.$$

Now, since they generate $\mathcal{C}_0$, the vectors in G are an inequalities representation of a cone $\mathcal{C}_0^*$, which is called the dual of $\mathcal{C}_0$, $\mathcal{C}_0^* = \{x \in \mathbb{R}^m \mid g \cdot x \geq 0 \text{ for all } g \in G\}$. Clearly, given G, we can use the Nullspace algorithm to obtain a minimal generating set of $\mathcal{C}_0^*$ and, surprisingly, this set is just the set H we are looking for. The reason is that, by standard results in Convex Analysis (Rockafellar, 1970; Fukuda, 2004, http://www.ifor.math.ethz.ch/~fukuda/polyfaq/), the vectors of a generating set of the dual cone $\mathcal{C}_0^*$ are an inequalities representation of the primal cone $\mathcal{C}_0$. In fact any algorithm computing a minimal generating set from a linear inequalities representation also solves the converse problem of obtaining a minimal linear inequalities representation, given a generating set.

In summary, the basic procedure is to obtain H from G by running Nullspace (as detailed in Supplementary Material A), then to augment the set H by the vectors reflecting the additional constraints for the conversion cone (Equation 7) and, finally, to run Nullspace again on the augmented set to get the complete set of elementary conversions. This is illustrated in Figure 2 using a very simple example.

4 APPLICATIONS

To compare the conversion cone with the flux cone we first study the stylized model of the central carbon metabolism of *E.coli* presented in Stelling *et al.* (2002). A version of this metabolic model consisting of 99 reactions, with 11 external compounds (5 input, 5 output and 1 input/output), was analyzed in Klamt and Stelling (2002) and it was found that an excess of 5×10^5 elementary fluxes are needed for a complete set. In contrast to this, we find that a complete set of elementary conversions has only 344 elements, highlighting the drastic reduction in descriptive complexity achieved by focusing on the conversion cone.

In fact an even simpler description is possible if one considers all external compounds as being inputs as well as outputs and only calculates a minimal generating set of the conversion cone. One

Fig. 2. Example calculation of the conversions for the two reaction system $A_{\text{ext}} \to B \to C_{\text{ext}}$. Since the two reactions are irreversible, the columns of the above stoichiometry matrix S already are the generating set G of $\mathcal{C}_0$ from which we obtain H, the inequalities representation of $\mathcal{C}_0$. Note that for the first two columns of H we even have $h^{(1)} \cdot S^{(j)} = h^{(2)} \cdot S^{(j)} = 0$. Next, H^+ is the augmented system of inequalities, Equation (7), reflecting that the metabolite B is internal. Finally, the inequalities representation H^+ of $\mathcal{C}$ yields that the conversion cone is generated by the single conversion $\dot{c}$, i.e. by $A_{\text{ext}} \to C_{\text{ext}}$. A more elegant way of enforcing the equality $\dot{B} = 0$ is described in Supplementary Material A.3.
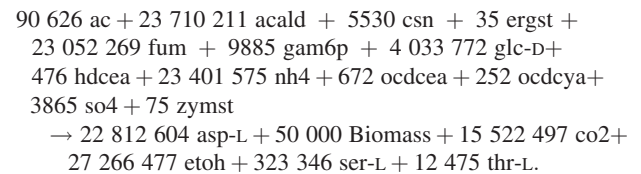
finds that this set consists of only 27 vectors, summarizing the basic metabolic capabilities of the model.

We next consider the genome-scale metabolic network of *S.cerevisiae* published in Duarte *et al.* (2004). This fully compartmentalized model has 1149 reactions as well as an additional reaction reflecting the metabolic needs of the organism for sustaining growth. To track the flow through this growth reaction we added the formal external metabolite *Biomass* to its output.

A coupling to the environment was used, where 21 of the metabolites marked as extracellular in the model are external. Interestingly, the conversion cone obtained for this coupling is only 10 dimensional, although it is embedded in the $21 + 1$ dimensional space given by the external metabolites. Further, the complete set for this conversion cone consists of 40 969 extremal conversions, 98% of which sustain growth. However, there are only 1313 extremal conversions in this set that do not consume any amino-acid but nevertheless show growth. These are analyzed in more detail in Figure 3.

Having obtained an overview of the metabolic capabilities of an organism, it will often be of interest to determine how a specific function is achieved. Luckily, it is conceptually simple to determine just the elementary fluxes that give rise to a single elementary conversion. To analyze an elementary conversion e in this manner, one will define the matrix S' obtained by augmenting the stoichiometry matrix S, with $-e$ as an irreversible column vector, as $S' = (S, -e)$. One can then compute the elementary fluxes of S', treating all compounds as internal. Among the elementary vectors $\binom{\nu}{\lambda}$ of the restricted flux cone thus obtained, a few may be futile cycles ($\lambda = 0$) but all others give rise to the conversion e, since $S\nu = \lambda e$.

To illustrate this, we analyzed the following conversion:

90 626 ac + 23 710 211 acald + 5530 csn + 35 ergst + 23 052 269 fum + 9885 gam6p + 4 033 772 glc-D+ 476 hdcea + 23 401 575 nh4 + 672 ocdcea + 252 ocdcya+ 3865 so4 + 75 zymst
→ 22 812 604 asp-L + 50 000 Biomass + 15 522 497 co2+ 27 266 477 etoh + 323 346 ser-L + 12 475 thr-L.

This was chosen because it maximizes ethanol production relative to biomass among the 1313 elementary conversions described
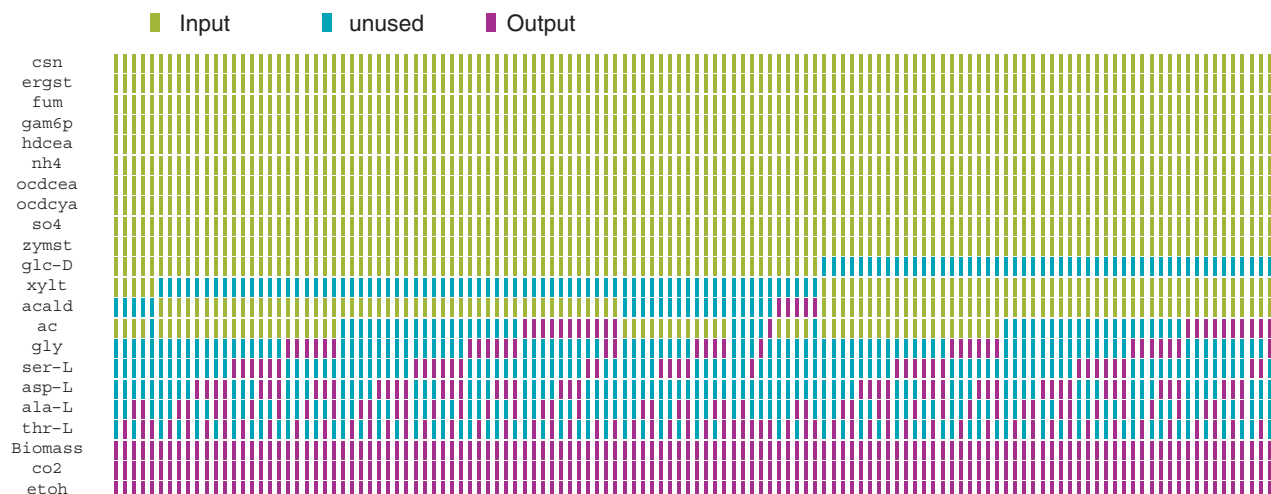
Fig. 3. Input/output sets of the 1313 elementary conversions that sustain growth but do not consume any of the amino acids: *ala-L, asp-L, gly, ser-L* and *thr-L*. The input/output sets are found by ignoring the stoichiometric factors in the conversion and only noting if a metabolite is consumed or produced. Each column in the figure shows one of the 129 distinct input/output sets thus obtained. Ten metabolites are essential for growth but have to be supplemented by either *glc-D* or both *acald* and *xylt* to form a minimal set of inputs enabling growth. Further, *co2* and *etoh* are necessary by-products of growth. However, if one considers the full set of conversion, where amino acids are allowed as inputs (data not shown), one finds 14 different minimal sets of inputs sustaining growth, and *etoh* is no longer a necessary by-product of growth. The full metabolite names are as follows: *ac*, acetate; *acald*, acetaldehyde; *ala-L*, L-alanine; *asp-L*, L-aspartate; *co2*, CO$_2$; *csn*, cytosine; *ergst*, ergosterol; *etoh*, ethanol; *fum*, fumarate; *gam6p*, D-glucosamine 6-phosphate; *glc-D*, D-glucose; *gly*, glycine; *hdcea*, hexadecenoate (n-C16:1); *nh4*, ammonium; *ocdcea*, octadecenoate (n-C18:1); *ocdcya*, octadecynoate (n-C18:2); *ser-L*, L-serine; *so4*, sulfate; *thr-L*, L-threonine; *xylt*, xylitol and *zymst*, zymosterol.

above, which sustain growth without consuming any amino-acid. Using the Nullspace procedure, we first find that already a minimal generating set of the restricted flux cone for this one conversion has 1039 elements. But the true scale of redundancy of the metabolic network only becomes apparent when the procedure goes on to enumerate the complete set for the restricted cone. This yields 3.71×10^7 elementary flux vectors, with only 28 corresponding to futile cycles and the remaining ones all providing different ways of implementing the single conversion displayed above.

5 DISCUSSION

We have shown that the analysis of elementary conversion is possible for genome-scale metabolic systems in situations where it seems highly probable that the enumeration of all elementary fluxes will remain computationally intractable for the foreseeable future. Although it is difficult to make rigorous statements about the complexity of our procedure, in the examples we have considered, the computational requirements are quite benign. For instance, the calculation of the 40 969 elementary conversion for *S.cerevisiae* takes <3 h of computing time on a standard workstation despite of the fact that we do not rely on machine precision but use infinite precision arithmetic in a mixed Mathematica/Matlab/C implementation of the Nullspace procedure. In part, this relatively fast computation time results from employing transformations (detailed in Supplementary Material B) that simplify the metabolic network but leave the conversion cone invariant.

While the conversion cone is an interesting object in its own right, enabling the determination of minimal media or the study of carbon fates, it is also important that conversion analysis can be linked up with flux analysis. We have already mentioned the possibility of directly determining the elementary fluxes that give rise to a given

elementary conversion. Incidentally, this technique can also be used to study the degeneracy of the optimization problems considered in the Flux Balance Analysis, and this seems computationally cheaper than the approach considered in Lee *et al.* (2000) and Reed and Palsson (2004), where the enumeration of the alternative optimal flux vectors is carried out by solving a sequence of NP-complete problems (mixed-integer linear programming).

A second way of linking conversion analysis with flux analysis is to use our technique to study more general projections of the flux cone. If one is interested only in the flows through a few reactions, one can modify the network by adding a new formal external metabolite as a tag for each reaction one wishes to track. For instance, to track the reaction A → B one could replace it by A → B + AtoB$_{\text{ext}}$. Then the flows through the reactions, which have been tagged in this way, can be read of from the stoichiometric factors of the formal metabolites appearing in the elementary conversions of the modified system. In a rudimentary form, we have already used this in our analysis of *S.cerevisiae* when employing the formal compound *Biomass* to track the growth reaction.

A third option for linking the two approaches, which we have yet to explore, arises in the analysis of subnetworks. Elementary flux analysis has variously been applied to larger networks by decomposing these into subnetworks and treating any metabolite linking two or more of the subnetworks as external (Schilling and Palsson, 2000; Schuster *et al.*, 2002b; Dandekar *et al.*, 2003). While regarding such linkage metabolites as external, formally makes it possible to consider each subnetwork in isolation, it has remained difficult to relate the elementary fluxes of the subnetworks to those of the entire network. An alternative is to focus on a single subnetwork and also treat the linkage metabolites as external but to use this to calculate the conversion cone of all of the reactions not belonging to the subnetwork. The minimal generating set of this

cone enumerates the potential contributions of the rest of the network to the operation of the subnetwork. One can now augment the subnetwork by the conversions in this generating set. Then any elementary flux of the augmented subnetwork represent at least one elementary flux of the entire system. In this manner one arrives at valid descriptions of the operation of the entire network, which are detailed with respect to the reactions in the subnetwork but only sketch the operations of the other reactions.

Large metabolic networks show a wealth of possible modes of operation. While this is probably essential for the survival of the microorganism in changing environmental conditions, it makes it practically impossible to enumerate all the possibilities. Hence, in describing the metabolism it is useful to focus on specific aspects of its operation and abstract away details that are not of interest in a given context. In stoichiometric network analysis a formal equivalent of this is to focus on suitable projections of the flux cone. The results presented above show that in this manner it does indeed become possible to tackle genome-scale networks.

ACKNOWLEDGEMENTS

REFERENCES

Clarke,B. (1980) Stability of complex reaction networks. In Prigogine,I. and Rice,S., (eds.), *Advances in Chemical Physics*, Wiley, New York, Vol. 43, pp. 1–216.

Dandekar,T. *et al.* (2003) A method for classifying metabolites in topological pathway analyses based on minimization of pathway number. *Biosystems*, **70**, 255–270.

Duarte,N., Herrgard,M. and Palsson,B. (2004) Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.*, **14**, 1298–1309.

Fukuda,K. (2004) *Frequently Asked Questions in Polyhedral Computation*.

Fukuda,K. and Prodon,A. (1995) Double description method revisited. M.Deza, R.Euler and I.Manoussakis (eds), *Combinatorics and Computer Science.* **1120**, Lecture Notes in Computer Science, Springer-Verlag, London, UK., pp. 91–111.

Happel,J. and Sellers,P. (1989) The characterization of complex systems of chemical reactions. *Chem. Eng. Commun.*, **83**, 221–240.

Heinrich,R. and Schuster,S. (1996) *The Regulation of Cellular Systems*. Chapman & Hall, New York.

Ibarra,R. *et al.* (2002) *Escherichia coli* k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*, **420**, 186–189.

Klamt,S. and Stelling,J. (2002) Combinatorial complexity of pathway analysis in metabolic networks. *Mol. Biol. Rep.*, **29**, 233–236.

Lee,S. *et al.* (2000) Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. *Comput. Chem. Eng.*, **24**, 711–716.

Price,N. *et al.* (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.*, **2**, 886–897.

Reed,J. and Palsson,B. (2004) Genome-scale *in silico* models of *E.coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome Res.*, **14**, 1797–1805.

Reed,J. *et al.* (2003) An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.*, **4**, R54.1–R54.12.

Rockafellar,R. (1970) *Convex Analysis*. Princeton University Press, Princeton.

Schilling,C. and Palsson,B. (2000) Assessment of the metabolic capabilities of *Haemophilus influenzae* rd through a genome-scale pathway analysis. *J. Theor. Biol.*, **203**, 249–283.

Schuster,S. *et al.* (2002a) Reaction routes in biochemichal reactions systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J. Math. Biol.*, **45**, 153–181.

Schuster,S. *et al.* (2002b) Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae*. *Bioinformatics*, **18**, 351–261.

Stelling,J. *et al.* (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature*, **420**, 190–193.

Stucki,J. (2004) Chromokinetics of metabolic pathways. *Eur. J. Biochem.*, **271**, 2745–2754.

Urbanczik,R. and Wagner,C. (2005) An improved algorithm for stoichiometric network analysis: theory and applications. *Bioinformatics*, **21**, 1203–1210.