A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes

Jean-Marc Schwartz* and Minoru Kanehisa
Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

1 INTRODUCTION

Metabolic pathways are an essential key to the systemic behaviour of biological cells. Accurate dynamical modelling of cellular processes is in most cases impossible because the required kinetic data are missing. Therefore alternative network-based approaches have been developed, which require only the knowledge of network topology and stoichiometry. In particular, two very similar concepts called elementary modes and extreme pathways have been introduced (Papin et al., 2003). An elementary mode is a minimal set of reactions that can operate in steady state through a metabolic network, and the set of extreme pathways is the systemically independent subset of the elementary modes. These network-based approaches have already been used for a variety of applications, including the evaluation of pathway robustness and flexibility, prediction of network functionality, assessment of the effects of gene deletions or environmental changes, etc.

Every steady-state flux distribution can be expressed as a non-negative linear combination of elementary modes or extreme pathways. But all these states may not necessarily be attainable by a real biological system. To understand to which extent individual elementary modes contribute to the reconstruction of actual physiological flux distributions, there is usually an infinite number of possible decompositions. Only certain combinations within the α-spectrum are allowed. Furthermore, correlations between the weightings of different modes could not be studied. We therefore sought a different approach for flux distribution decomposition, with the aim of finding a particular set of weightings that is suitable for biological interpretation.

The following sections refer to elementary modes, but an identical process could be applied to extreme pathways. Discussions about the relative qualities of both concepts have been published (Papin et al., 2004). Here we chose to concentrate on elementary modes because the set of extreme pathways may in some cases not contain all biologically relevant direct routes.

2 METHODS

Since elementary modes are not a linear basis of the flux space, there is usually an infinite number of possible weighting combinations to reconstruct a given flux distribution. We chose to select the particular set that minimizes the length of the weighting vector. The idea behind this choice is that a biologically meaningful description should favour the modes that are closest to the actual state of the system. For a given steady-state flux vector v and a set of elementary modes v₁, v₂, ..., vₘ, we are thus looking for a set of scalars a₁, a₂, ..., aₘ sharing the following properties:

\[ v = \sum_{i=1}^{m} a_i v_i \]  (1)
\[ a_i \geq 0, \quad \forall i \in [1, m] \]  (2)
\[ \text{Minimize} \sum_{i=1}^{m} a_i^2 \]  (3)

These conditions define a quadratic programming problem. Different algorithms are known for solving this type of optimization problem (Fletcher, 1987). We implemented an algorithm based on the active set method. A few adaptations were added to the standard algorithm in order to meet specific requirements to this problem. The whole decomposition process is summarized in the following steps:

- Reactions that always operate in fixed flux proportions in steady state are lumped together in order to reduce the dimension of the system.
- The stoichiometric matrix of the reduced network is constructed, and elementary modes are computed according to the algorithm presented by Schuster et al. (2000).
- The matrix of elementary modes is used as input to the active set algorithm. That algorithm requires an initial feasible point which...
Quadratic programming approach

Fig. 1. (a) Lumped model of yeast glycolysis and steady state fluxes predicted by Teusink et al. (2000), with concentrations of glucose and ethanol fixed at 50 mM. Grey numbers indicate fluxes in mmol min\(^{-1}\) l-cytosol\(^{-1}\), while bold black numbers indicate stoichiometries when other than 1. Abbreviations: AcAld, acetaldehyde; DHAP, glycerone phosphate; GAP, D-glyceraldehyde 3-phosphate; G6P, glucose 6-phosphate. (b) Weight values of each elementary mode in the given flux distribution obtained after quadratic programming decomposition.

is found by linear programming. In some cases, the initial point may already be the solution of the quadratic programming problem; otherwise, the active set algorithm is run.

• Depending on the method or software used to compute steady-state flux distributions, computational approximations may lead to insignificant violations of flux conservation. The active set algorithm showed serious convergence problems when flux conservation was not strictly held. Therefore flux values are checked for conservation before the decomposition is performed, and adjusted if small violations are found. This operation moreover allows proceeding to a further reduction of the dimension of the system, as only a basis set of independent fluxes is retained.

3 RESULTS AND DISCUSSION

An application of this approach to a particular example is presented in this section. We used a branched model of yeast glycolysis presented by Teusink et al. (2000), available from the JWS Online repository (Olivier and Snoep, 2004). This model has been constructed after experimental determination of all kinetic parameters, and can therefore be assumed to produce physiologically meaningful metabolic states. The already lumped network is shown in Figure 1a, together with a steady-state flux distribution predicted by the model.

The elementary modes of this network are shown in Figure 2. The raw weight values obtained after decomposition of the above flux distribution by our quadratic programming algorithm are displayed in Figure 1b. The highest weight was attributed to EM8, which represents the usual glycolysis pathway. This observation strengthens the choice made in our decomposition approach, as the present flux distribution indeed corresponds to standard glycolytic conditions. The highest score has been given to the mode which is most relevant for biological interpretation in this state. We can also observe that three modes received zero weights, meaning that these modes are not required to describe the present state. However, these modes are not identical to the elementary modes that can be recovered by combinations of extreme pathways (i.e. EM4, EM5 and EM7, the extreme pathways of this network being EM1, EM2, EM3, EM6 and EM8).

This decomposition approach may provide a valuable tool for the analysis of metabolic network flux distributions and for further characterization of the range of possible behaviours of metabolic networks. In particular, answers to the following questions may be investigated in the future through this approach: Are all elementary modes necessary to describe feasible states in metabolic networks? Does the space of physiologically feasible states span the whole space of stoichiometrically feasible states? Can the dominant behaviour of a metabolic network be predicted when only partial knowledge of the kinetic parameters is available?

ACKNOWLEDGEMENTS

This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan, the Japan Society for the Promotion of Science, and the Japan Science and Technology Corporation.

Conflict of Interest: none declared.

REFERENCES


