Combining kinetics and network-based pathway analysis

Jean-Marc Schwartz[1]
jean@kuicr.kyoto-u.ac.jp

Minoru Kanehisa[1]
kanehisa@kuicr.kyoto-u.ac.jp

[1]Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

Keywords: pathway analysis, elementary modes, kinetics, S-system

1 Introduction

As genome-scale metabolic networks can now be reconstructed, the systemic biochemical properties of these large networks are being studied with more and more interest. As with any network of interacting elements, the behaviour of such a system cannot be comprehended by the sole characterization of its individual components or pair-wise relations. Even systems with as little as three independent variables may exhibit very complex behaviour, and new properties emerge from the interaction of different components. Systems-based approaches are needed to gain knowledge about complex cellular processes, and may allow in a further future to provide an integrated and predictive description of a complete organism.

Metabolic pathways are an essential key to the systemic behaviour of a biological cell, as they describe a multitude of enzymatic reactions carrying out various functions. One possible approach to the study of these large-scale networks is dynamical simulation. But precise mathematical modelling of cellular processes is in most cases impossible because the required kinetic data are missing. Therefore alternative network-based approaches have been developed and have found much interest recently, in particular two very similar mathematical concepts called *elementary flux modes* and *extreme pathways* [3]. In contrast to dynamical simulation approaches, these descriptions only require the knowledge of network topology and stoichiometry, which are well known is many cases.

2 Network-based pathway analysis

In network-based pathway analysis, a biological network of interacting components is represented by a stoichiometric matrix which relates reactions and metabolites. This matrix is analyzed by an algorithm that computes a set of routes satisfying specified conditions. For a route to be a valid elementary mode, three conditions are required: a pseudo steady-state condition (*i.e.* none of the metabolites must be consumed or accumulated), a feasibility condition (*i.e.* no negative flux is allowed for irreversible reactions), and a non-decomposability condition (*i.e.* no subset of the route can hold the network in a steady-state with non-zero fluxes). Extreme pathways additionally require a systemic independence condition, *i.e.* no extreme pathway can be represented by a non-negative linear combination of other extreme pathways. For this reason the set of extreme pathways is minimal and is always a subset of the set of elementary modes.

In a mathematical multidimensional representation where each axis corresponds to a reaction flux, all possible steady-state flux distributions lie within a multidimensional *flux cone*, and extreme pathways form the edges of this cone. The additional elementary modes may lie on the surface or the interior of the cone.

These structure-oriented approaches have already lead to a significant number of applications, and they can help predicting key aspects of functionality, robustness and gene regulation. For example, the number of elementary modes can serve as a measure of network robustness and fragility towards disturbances such as mutations or drug inhibition, and allow the prediction of gene expression patterns [5]. Prediction of the function or functional class of orphan genes can be achieved by comparing the pathway structure of different

strains [1], and genome-scale pathway analysis can provide useful system-level characterizations [2].

3 Analyzing the space of steady-states

An important question yet remains open, do extreme pathways or elementary modes represent physiologically relevant functional states of a biological network? The flux cone generated by the extreme pathways is known to contain all possible steady-state flux distributions for the system, but whether all these states can actually be reached by a real biological system is unknown. Therefore, new approaches are required to further characterize the flux cone and to understand how environmental or expression conditions restrict the space of allowable states.

This project aims at studying the influence of kinetic parameters onto the extent of the space of allowable states. Kinetics plays a crucial role as it determines which steady-state will actually be reached by the system. Different kinetic representations may be used for this type of analysis. The Michaelis-Menten description is frequently used for modelling biochemical reactions, but it is rather inflexible and has been shown to be incompatible with certain types of reactions [4]. The S-system representation on the opposite presents a number of advantages. It covers a wide range of behaviours, its equivalence to other formalisms has been proven for complex systems, and straightforward algorithms based on linear algebra exist for the computation of steady-states. For these reasons, the S-system approach has been chosen in this project.

The approach followed here consists in decomposing a given steady-state onto the set of elementary modes or extreme pathways. This decomposition is not unique, therefore a description in terms of range of weighting, as developed by the concept of α-spectrum [6], has to be used. This will allow the effect of parameters onto the range of weightings to be studied. Analyzing how the value of a single kinetic parameter affects the space of allowable states can lead to information about the functionality of that given reaction and of the associated genes. Comparing or correlating the effects of several parameters can give information about correlated functions or regulation. Assuming uncertainty for all the parameters in a network may either allow the system to reach all the states contained in the unconstrained flux cone, or still produce a smaller allowed space, because of the inherent constraints provided by kinetics.

In the current state of this project, an S-system model of biochemical pathways and an algorithm for steady-state computation have been implemented using Java. Algorithms for the computation of elementary modes and extreme pathways have been implemented and are being integrated into the S-system framework. These algorithms have been successfully validated on model pathways by comparing their results to published data. Further developments regarding the decomposition of steady-states onto the set of extreme pathways and the analysis of kinetics influence onto that decomposition are under way.

References

[1] Förster, J., Gombert, A.K., Nielsen, J., A functional genomics approach using metabolomics and *in silico* pathway analysis, *Biotechnology and Bioengineering*, 79(7): 703-712, 2002.

[2] Papin, J.A., Price, N.D., Edwards, J.S., Palsson, B.Ø., The genome-scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy, *Journal of Theoretical Biology*, 215(1): 67-82, 2002.

[3] Papin, J.A., Price, N.D., Wiback, S.J., Fell, D.A., Palsson, B.Ø., Metabolic pathways in the post-genome era, *Trends in Biochemical Sciences*, 28(5): 250-258, 2003.

[4] Shiraishi, F., Savageau, M.A., The tricarboxylic acid cycle in *Dictyostelium discoideum*. I. Formulation of alternative kinetic representations, *Journal of Biological Chemistry*, 267(32): 22912-22918, 1992.

[5] Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S., Gilles, E.D., Metabolic network structure determines key aspects of functionality and regulation, *Nature*, 420(6912): 190-193, 2003.

[6] Wiback, S.J., Mahadevan, R., Palsson, B.Ø., Reconstructing metabolic flux vectors from extreme pathways: defining the α-spectrum, *Journal of Theoretical Biology*, 224(3): 313-324, 2003.