Xeledon: Efficient enumeration of all steady state cycles in metabolic networks

Rene Rex*† Dietmar Schomburg*

Genome-scale metabolic models summarize the knowledge about an organism and thus allow predictions about its features and capabilities. Most widely used are constraint-based models which rely on a set of chemical reactions and optionally include bounds for the flux trough them. Validation and quality assurance are of great importance if the model of an organism is intended to be used for making predictions about the organisms' features or to guide expensive experiments. While the best way is validation against biological data, it is not available in all cases. Here, we address a common error occurring during a metabolic reconstruction: the inclusion of thermodynamically implausible reactions. Such reactions often lead to internal cycles which violate the second law of thermodynamics.[4,6] We exploit this feature by enumerating all cycles in a metabolic network and so ease the search for thermodynamically implausible reactions. Furthermore, a cycle may be a feature of the organism's metabolism. Specifically, it has been shown that a futile cycle can regulate the equilibrium of a metabolite pool.[3] From a regulatory point of view, cycles can be used to produce a feedback signal.[1] In the context of extreme pathways the cycles studied here have been classified as *Type III pathways*.[5]

In this work, we propose a fast algorithm which finds all steady state cycles in a metabolic network (algorithm 1). The key idea is to remove as many reactions as possible before carrying out the computationally expensive convex analysis. Since we are interested in internal fluxes only, all boundary reactions are removed first (lines 1 - 4). This creates blocked reactions which cannot carry any flux without violating the steady-state condition. Obviously, these reactions cannot participate in any cycle and thus are removed in line 5. Some reactions may have a bound which forces them to carry a non-zero flux. This may prevent the detection of unlimited reactions if they either cannot be active at all or their corresponding cycle cannot operate at the same time as some other cycle in the network. Thus, all flux bounds are relaxed in lines 7 and 8 so that they can be zero. All reactions which can still carry a flux at this point must be part of a cycle. Hence, all other reactions are identified and removed in line 9 using fast FVA.[2] Finally, the flux cone described by the remaining reactions is constructed in line 14 and all internal extreme pathways are enumerated. An example of an identified cycle caused by a PEP synthase running backwards is shown in figure 1.

*Department for Bioinformatics and Biochemistry, Technische Universität Braunschweig, Germany
†r.rex@tu-bs.de

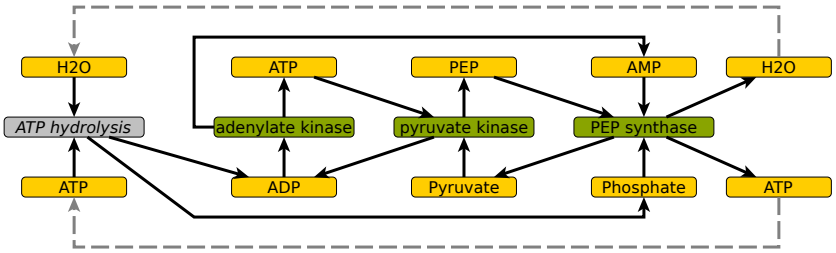

Figure 1: Example cycle which could regenerate ATP caused by a missing lower bound for the PEP synthase reaction

Input : A set R of reactions
Two vectors $\vec{u}$ and $\vec{l}$ with upper and lower flux bounds for each reaction in R
A set O of objective reactions
Output: A set C of sets of reactions which form a cycle

`// remove all objective reactions`
remove each reaction from R if reaction $\in O$
foreach reaction *in* R **do**
  if reaction *has no products* **or** reaction *has no educts* **then**
    remove reaction from R
foreach reaction *in* R **do**
  if reaction *is blocked* **then** remove reaction from R

`// allow each reaction to have zero flux`
Set each element e of $\vec{u}$ to zero if $e < 0$
Set each element e of $\vec{l}$ to zero if $e > 0$

`// use FVA to remove reactions which cannot carry any flux`
$\overrightarrow{minflux}, \overrightarrow{maxflux} = \texttt{fastFVA}(R, \vec{u}, \vec{l})$
foreach reaction *in* R **do**
  if $\overrightarrow{minflux}_{reaction} = 0$ **and** $\overrightarrow{maxflux}_{reaction} = 0$ **then**
    remove reaction from R

`// identify cycles using convex analysis`
$C = \emptyset$
foreach ray *in* `convexAnalysis` (R) **do**
  create a set C^r of reactions corresponding to the ray
  add C^r to C

Algorithm 1: xeledon algorithm

References

[1] J. Ferrell et al. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Current Opinion in Cell Biology*, 14(2):140–148, 2002.

[2] S. Gudmundsson and I. Thiele. Computationally efficient flux variability analysis. *BMC bioinformatics*, 11(1):489, 2010.

[3] H. Qian and D. Beard. Metabolic futile cycles and their functions: a systems analysis of energy and control. In *Systems Biology, IEE Proceedings*, volume 153, pages 192–200. IET, 2006.

[4] J. Schellenberger, N. E. Lewis, and B. O. Palsson. Elimination of thermodynamically infeasible loops in Steady-State metabolic models. *Biophysical Journal*, 100(3):544–553, Feb. 2011.

[5] C. Schilling, D. Letscher, and B. Palsson. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *Journal of Theoretical Biology*, 203(3):229–248, 2000.

[6] J. Wright and A. Wagner. Exhaustive identification of steady state cycles in large stoichiometric networks. *BMC systems biology*, 2(1):61, 2008.