

# Minimal Sets of External Compounds in Metabolic Networks

Thomas Handorf

handorf@physik.hu-berlin.de

Oliver Ebenhöh

oliver.ebenhoe@rz.hu-berlin.de

Theoretical Biophysics, Institute of Biology, Humboldt University Berlin, Germany

**Keywords:** metabolic networks, external compounds, fluxes, structural analysis

## 1 Introduction

Cellular metabolism has been subject to intense research in the past decades. The vast amount of knowledge on metabolic systems resulted in the development of metabolic databases such as KEGG. Whereas the stoichiometry, that is the structure, of metabolic pathways is largely known, exact information on the kinetic parameters of the metabolic reactions is still limited. With the help of structural analysis, valuable information about metabolic networks can be gained even without the knowledge of these parameters. Possible metabolite interconversions and alternative routes through the system can be determined by calculating the elementary flux modes or extreme pathways [1, 2]. This, however, requires the information on which metabolites undergo a net consumption or production (external compounds) and which metabolites are balanced (internal compounds). If biological information is available on which metabolites are produced or consumed by other processes, these sets can in principle be manually determined. However, often such biological knowledge is incomplete.

In this work we present an algorithm which automatically determines the external compounds for arbitrarily sized networks while additional biological information can optionally be provided.

## 2 Method and Results

Metabolic networks comprising  $r$  reactions and  $q$  compounds define a  $r \times q$ -matrix  $\mathbf{N}$ , the so-called stoichiometric matrix. Possible fluxes  $\mathbf{v}$  through the system are calculated by solving the equation  $\mathbf{N}_{\mathbf{S}} \cdot \mathbf{v} = 0$  using the part of the stoichiometric matrix  $\mathbf{N}_{\mathbf{S}}$  describing only the internal compounds. The general solution is  $\mathbf{v} = \sum_i \alpha_i \mathbf{b}_i$  where the  $\mathbf{b}_i$  are a set of linearly independent vectors spanning the solution space. In this work all reactions are considered to be reversible, hence no sign restrictions apply to the components of the solution vectors.

In the following we describe an algorithm determining a minimal set of external compounds. Such sets must fulfil a utilisation condition which demands that for each reaction there exists at least one base vector  $\mathbf{b}_i$  in which the corresponding component is non-zero. This reflects the assumption that networks containing unused reactions are not reasonable in an evolutionary context.

The algorithm is started with considering all compounds as external resulting in an empty matrix  $\mathbf{N}_{\mathbf{S}}$  and consequently fulfilling the utilisation condition. In the next step, a first compound is tested whether it can be internalised or not. If this compound participates only in one reaction, this is obviously not possible. Otherwise, the test is performed by adding the corresponding row to  $\mathbf{N}_{\mathbf{S}}$ , calculating the  $\mathbf{b}_i$ , and checking whether the utilisation condition is fulfilled. If so, the compound is considered internal and the corresponding row remains in  $\mathbf{N}_{\mathbf{S}}$ , otherwise the row is removed and the compound remains external. This step is repeated for each compound.

Clearly, the result of this algorithm depends on the order in which the compounds are tested. Therefore, the calculated set is locally minimal in the sense that none of the external compounds can be internalised without violating the utilisation condition.

To determine the globally minimal set, in principle this calculation has to be carried out for every possible permutation of the list of compounds. However, performing  $q!$  calculations is not feasible even for small networks. Fortunately, many permutations eventually result in identical sets of external compounds. We reduce the computational effort dramatically by detecting such identical sets at an early stage in the algorithm.

We have applied the algorithm to the glycolytic pathway which contains 20 compounds. We found 108 different minimal sets ranging from five to nine external compounds. Two of these sets are depicted in Figure 1. The upper graph shows an intuitive solution of a locally minimal set of external compounds including all cofactors. The lower graph represents a globally minimal set which contains no cofactors.

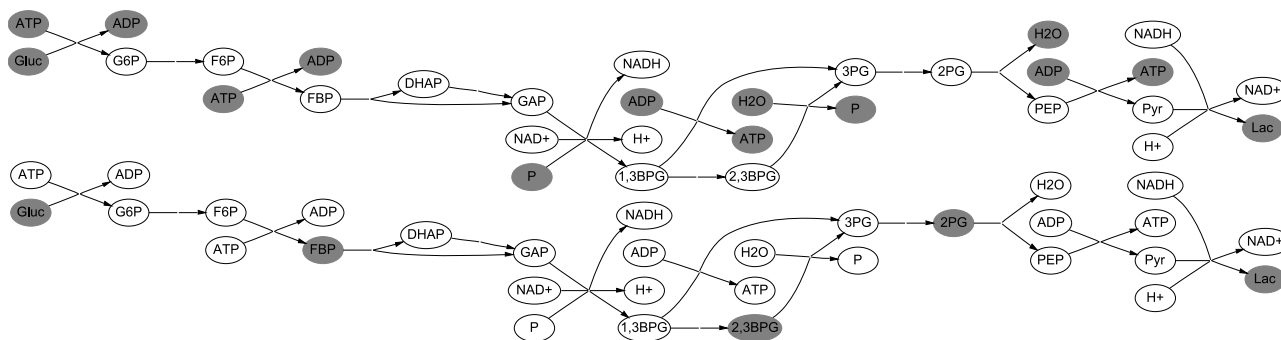


Figure 1: Two representations of the glycolytic pathway with different minimal sets of external compounds (marked grey).

### 3 Discussion

The presented algorithm provides a tool to determine globally and locally minimal sets of external compounds for given metabolic pathways. The sets are minimal in the sense that internalisation of one of the external compounds would result in unused reactions. The results obtained by the algorithm are of great importance for understanding complex biological systems. The calculated sets define minimal requirements of the metabolic pathway on the environment. This information allows to draw conclusions on which resources have to be provided by other processes.

The derived methods are also valuable from a mathematical perspective. The knowledge of sets of external metabolites is a prerequisite for structural pathway analysis. The presented algorithm allows to determine such sets even if the biological information on external compounds is not available or insufficient.

### References

- [1] Schuster, S. and Dandekar, T. and Fell, D. A., Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering, *Trends Biotechnol.* , 17, 53–60, 1999
- [2] Schilling, C. H. and Letscher, D. and Palsson, B. O., Theory for the Systemic Definition of Metabolic Pathways and their use in Interpreting Metabolic Function from a Pathway-Oriented Perspective, *J. Theor. Biol.* 203, 229–248, 2000.