John von Neumann Institute for Computing

From Isotope Labeling Patterns to Metabolic Flux Rates

K. Nöh, M. Weitzel, W. Wiechert

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 345-348, 2008.

http://www.fz-juelich.de/nic-series/volume40

From Isotope Labeling Patterns to Metabolic Flux Rates

Katharina Nöh[1], **Michael Weitzel**[2], **and Wolfgang Wiechert**[2]

[1] Institute of Biotechnology 2, Research Centre Jülich, 52425 Jülich, Germany
E-mail: k.noeh@fz-juelich.de

[2] Institute of Systems Engineering, University of Siegen, 57068 Siegen, Germany
E-mail: {michael.weitzel, wolfgang.wiechert}@uni-siegen.de

Intracellular metabolic flux rates are the manifestation of metabolic activities within organisms. Aiming at a precise quantification of the *in-vivo* fluxes, Metabolic Flux Analysis based on labeling experiments has become an valuable key technology in Systems Biology. Topology-based computational algorithms are derived to facilitate informative high-throughput analyses as well as predictive computational modeling and simulation approaches to generate new knowledge and a robust experimental design of labeling experiments.

1 Introduction

Intracellular metabolic flux rates are the most important indicators of the metabolic activities in organisms. Depending on external conditions, the *fluxome* determines the cells' physiological phenotype and, thus, their metabolic capabilities. The *in-vivo* reaction rates, however, cannot be directly accessed from measurements. Aiming at a precise quantification, Metabolic Flux Analysis (MFA) based on Isotope Labeling Experiments (ILE) became an invaluable tool for Systems Biology.

2 General Procedure: Isotope Labeling Experiments

Over the past decade two types of isotope-based MFA emerged and are being successively refined. Both methods rely on measurements of either labeled biomass components (with GC-MS, NMR) or labeled primary intermediates (LC-MS). The classical, well established *stationary* isotope MFA characterizes a cell's fluxome in a metabolic and isotopic stationary state. Typically, in a continuous culture the feed is switched from naturally labeled to isotopically labeled medium which propagates though the network and progressively replaces the unlabeled intermediates. When the labeling distribution is approximately time-invariant, samples are taken. In this field, recent experimental progress strengthened the development of high-throughput MFA[1], the investigation of extensive metabolic networks as well as the utilization of elaborated nonlinear statistical methods for flux estimation[2].

A current area of research is isotopically *non-stationary* MFA which represents a promising generalization of the classical approach[3]. Here, the cells are likewise kept under metabolic steady state conditions, however, now the time profiles of the labeling patterns are measured upon start of the labeling period in order to monitor the time-dissolved labeling propagation of the isotopic tracer though the network. Compared to classical isotope MFA, non-stationary ILEs are typically more informative and facilitate a cross check between metabolome and fluxome data. In particular for comprehensive networks, however, the required computational effort is prohibitive[3]. Altogether, the new experimental techniques result in an increased demand for more efficient algorithms.

3 Modeling and Computational Machinery

Besides the measurement of labeling patterns in key metabolites, extracellular rates, and pool sizes (non-stationary case only), a biochemical reaction network has to be provided. Additionally, atom transitions of all reactions have to be specified which essentially describe how the flow of labeled material through the cell's metabolic pathways is organized.

The analytic determination of the fluxes as a function of the emerging labeling patterns, however, is impossible for realistic networks. Thus, for unraveling intracellular fluxes, an iterative parameter fitting procedure has to be applied. Summarizing, mathematical modeling and computational algorithms build the foundation of isotope-based MFA. In this context, the computational bottleneck is the forward simulation step, i.e. the calculation of emerging labeling patterns from given parameters. Because all possible labeling combinations of each metabolite have to be considered, at the end, the solution procedure involves the solution of large, cascaded systems of Labeling Balance Equations (LBEs)[4,5].

The requirement for the computational evaluation of the model is in each case high: system dimensions range from 700 with 45 parameters (central metabolism, CM) and 5600 with 65 parameters (CM with biosynthesis pathways) up to 355.000 with 580 parameters for genome-scale models. The classical method involves the solution of an algebraic equation system while for the non-stationary method a system of ordinary differential equations, with possibly stiff characteristic, has to be solved.

4 Topological Techniques for Dimension Reduction

Efficient non-standard solution algorithms basically all rely on the structure of the Isotope Labeling Network (ILN) graph associated with the metabolic network. Although the solution algorithms for the stationary and non-stationary approaches are quite different, both approaches certainly benefit from a reduction of the problem size. Dimension reduction is performed by a careful analysis of the ILN graph followed by a removal of specified nodes and edges or a decomposition into smaller subsystems. The two basic approaches for dimension reduction are:

Path Tracing. Typically, measurement data describe only a small subset of metabolites. The forward simulation step can be restricted to a relevant sub-network which sufficiently describes the transport of labeling from the substrates to the measured metabolites[5]. The necessary path tracing procedure relies on the computation of the transitive closure of the network graphs and is performed in two directions – a *forward tracing*, which determines the fate of the isotopic labeling found in the substrates, and a *backward tracing* which determines the topological predecessors of a labeling pattern or fragment.

Network Graph Decomposition. In a divide-and-conquer approach the network graphs are decomposed into *disconnected*, unilaterally connected (*acyclic*) and *cyclic* subnetwork components (so called CCs, DAGs, and SCCs, respectively). This decomposition heavily uses the unidirectionality of reactions. Once labeling leaves a cyclically connected subnetwork (by taking an unidirectional reaction route), it is impossible for it to return. This essential information can be used for decomposition of a network into smaller subnetworks. The decomposition results in subproblems with lower dimension and, thus, dramatically reduces the running time of the solution algorithms[5]. Depending on the network connectivity the speed-up is at least in the order of two to three magnitudes.

Although the potential of these two methods is high, further reduction is possible using an even more fine-grained approach: due to their combinatorial origin, the ILNs contain isomorphic subgraphs in form of parallel paths. Since the generated LBEs are likewise isomorphic, this property results in redundant computations which are usually hard to eliminate. Fortunately, this problem is solved by the generation of analytical solutions which facilitate the elimination of common subexpressions. As a by-product analytical solutions enable exact evaluation of sensitivities and the generation of highly efficient machine code.

In case of non-stationary ILEs, typically with growing label exchange between neighboring metabolites the differential equation systems tend to be stiff. Application of e.g. a s-stage implicit Runge-Kutta scheme involves a $sN \times sN$-dimensional linear equation system, where N denotes the number of differential equations to be solved. By imposing specialized structures on the Runge-Kutta matrix, e.g. by choosing a SDIRK scheme, the computational cost for its solution can be reduced. However, the complexity of performing at least one Newton step remains. Clearly, both topological approaches presented above can be directly applied. Moreover, because the sparsitivity pattern of the Runge-Kutta matrix remains the same for all time steps, the network decomposition has to be performed only once.

5 Conclusion and Outlook

The computational routines used in stationary, and in particular non-stationary isotopic MFA suffer from the inherent computational complexity of the approach. However, exploiting the nature of the underlying (algebraic and differential) equation systems improves the efficiency of the solution methods. A careful study of the labeling network topology leads to a significant increase in performance. New algorithms emerge, having their roots in Graph Theory, Linear Algebra, and Compiler Theory. However, particularly for the classical approach the new techniques providing analytical and fast numerical solutions open the perspective to simulate even genome-scale metabolic models.

References

1. U. Sauer, *Metabolic networks in motion:* ^{13}*C-based flux analysis.* Mol. Syst. Biol. **2**, 62, 2005.
2. V. Kadirkamanathan, J. Yang, S.A. Billings, P.C. Wright PC, *Markov Chain Monte Carlo Algorithm based metabolic flux distribution analysis on Corynebacterium glutamicum.* Bioinf. **22(21)**, 2681–2687, 2006.
3. K. Nöh, K. Grönke, B. Luo, R. Takors, M. Oldiges, W. Wiechert, *Metabolic flux analysis at ultra short time scale: isotopically non-stationary* ^{13}*C labeling experiments.* J. Biotechnol. **129(2)**, 249–267, 2007.
4. K. Nöh, A. Wahl, W. Wiechert, *Computational tools for isotopically instationary* ^{13}*C labeling experiments under metabolic steady state conditions.* Metab. Eng. **8(6)**, 554–577, 2006.
5. M. Weitzel, W. Wiechert, K. Nöh, *The topology of metabolic isotope labeling networks.* BMC Bioinformatics **8**, 315, 2007.