A Software Framework for Modeling and Simulation of Dynamic Metabolic and Isotopic Systems

Jana Tillack[1], Stephan Noack[1*], Katharina Nöh[1], Atya Elsheikh[2], Wolfgang Wiechert[2]

[1] Inst. of Biotechnology, Research Centre Jülich, Leo-Brandt-Strasse, 52425 Jülich, Germany; *s.noack@fz-juelich.de*
[2] Inst. of System Engineering., Dept. of Simulation, University of Siegen, Germany

Abstract. Various experimental approaches including also isotopic tracer information were proposed paving the way for quantitative modeling and detailed *in vivo* studies of biological systems. Accordingly, for the underlying modeling approaches a diversity of *in silico* tools have been developed. The full exploitation of this potential to address metabolic processes is hampered by mainly three principal issues.

First, not all currently realizable experiments are covered by these tools. Secondly, an easy switching between existing tools allowing for a flexible description of different experimental states is not possible. The third item addresses the universality of underlying modeling concepts which usually have a restricted focus.

We propose a general modeling concept which allows modeling and simulation of all combinations of metabolically and isotopically variants in their stationary and dynamic states and which is embedded in an unique software platform.

The basic idea is to build up dynamic metabolic networks relying on mass balances for intermediate labeling pools. A workflow is presented that allows the automatized generation of models of any size and complexity specially tailored for the experiment of choice. Within the software framework, the application of sophisticated methods for statistical analysis and interpretation of simulation results are realized.

Introduction

To reveal underlying kinetic mechanisms of metabolic regulation, pulse experiments have been established to generate dynamic data of metabolic intermediates [11, 7]. For the description of these data sets, dynamic metabolic networks are formulated that are based on kinetic models describing enzyme catalysis and regulatory metabolic interactions [17].

Validation of such models is a challenging task since usually a huge amount of parameters have to be identified with only a limited number of measurements available.

In order to attenuate the solution of these usually ill-posed problems, but nevertheless being able to elucidate metabolic stationary phenotypic behavior, metabolic flux analysis (MFA) has been introduced [14].

This approach has been refined by the addition of isotopically labeled tracers (usually ^{13}C-labeled glucose) [19]. In the last decade ^{13}C-MFA has become one of the major tools of metabolic engineering which is successfully applied to gain biological insight into different organisms of bacteria and plants [10, 13]. However, because it operates under metabolic stationary conditions, ^{13}C-MFA is not capable of describing *in vivo* metabolic regulation and control. Therefore its predictive power is limited.

Ongoing development of experimental and analytical procedures for measuring metabolic intermediates with and without tracer information (e.g. ^{13}C) [12] led to the requirement to formulate different model approaches describ- ing the measurement data. Similar to the previous step from MFA to ^{13}C-MFA recently a supplementation of the dynamic metabolic modeling approach with ^{13}C labeling has been suggested in [16]. Taking all approaches to- gether a classification can be derived between the model's assumption on metabolic and isotopic (non-)stationarity.

In the following we present a general framework for modeling and simulation of all common types of metabolic and isotopic systems.

In order to clarify our concept and introduce necessary definitions we start with a mathematical overview on the established model approaches which are introduced via a consistent example network.

SNE Simulation Notes Europe – Print ISSN 2305-9974 | Online ISSN 2306-0271
SNE 22(3-4), 2012, 147-156 | doi: 10.11128/sne.22.tn.10147

1 Modeling Metabolic Networks

1.1 General Model Assumptions

Our modeling approach relies on the following assumptions:

1. **Continuum:** All chemical species involved in the considered processes have such a high copy number that a continuous concentration value can be used to describe it.
2. **Homogeneity:** Diffusion processes are very fast compared to chemical reactions so that concentrations can be considered to be spatially homogeneous.
3. **No isotope effects:** There are no significant isotopic mass effects, i.e. the reaction rates do not depend on the actual labeling state of the reactants.

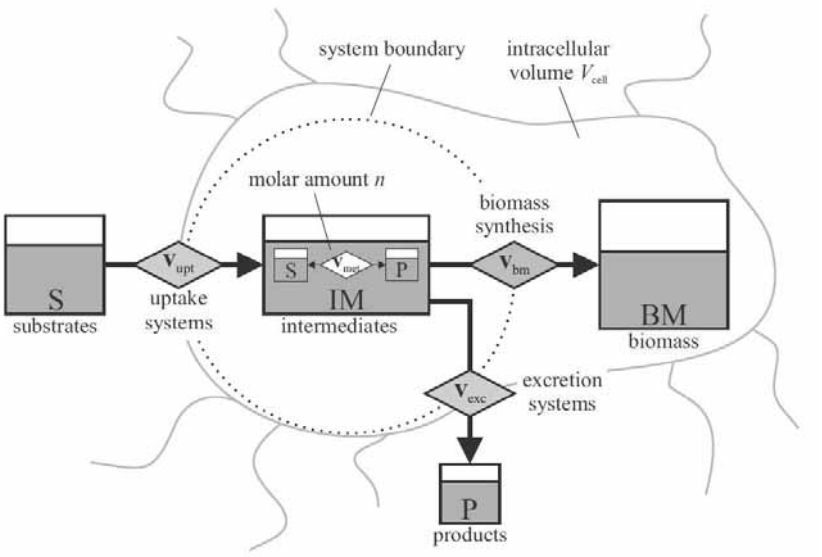


Figure 1. Scheme of metabolic processes within a cell that are covered by metabolic network models. Pools within the cell wall (indicated by the grey line) are mass balanced. The cell is fed with substrates and metabolizes intermediates towards biomass and excretion products.

With these assumptions, which are common for most of the currently available averaging biochemical network concepts [19], it is possible to describe reaction networks with metabolite pools and metabolic fluxes as state variables.

1.2 Mass Balances of Intermediate Pools

In principle, the cell's fluxome can be devided into two different species (cf. Figure 1):

1. Extracellular fluxes, comprising uptake systems ($\mathbf{v}_{u\,pt}$) for various substrates (*S*) into the cell and excretion systems ($\mathbf{v}_{exc}$) for products (*P*) out of the cell.
2. Intracellular fluxes, comprising reactions ($\mathbf{v}_{met}$) between metabolic intermediates (*IM*) and effluxes $\mathbf{v}_{bm}$ from intermediate precursors into biomass components (*BM*).

Following this classification a mass balance for the molar amount *n* of each metabolic intermediate in the cell can be formulated as:

$$\frac{dn_{IM}}{dt} = \sum v_{upt} - \sum v_{exc} \pm \sum v_{met} - \sum v_{bm} \quad (1)$$

By introducing (a) the relation $c_{IM} = n_{IM}/V_{cell}$ and (b) assuming that the intracellular volume is constant $\frac{dV_{cell}}{dt} = 0$, Equation 1 can be transformed into a mass balance for intermediate concentrations c_{IM}:

$$\frac{dn_{IM}}{dt} \overset{(a)}{=} \frac{d(c_{IM} \cdot V_{cell})}{dt} = \\ = c_{IM}\frac{dV_{cell}}{dt} + V_{cell}\frac{dc_{IM}}{dt} \overset{(b)}{=} V_{cell}\frac{dc_{IM}}{dt} \quad (2,3)$$

For a simple example shown in Figure 2 the time dependent concentration changes of all intracellular metabolites are formulated as:

$$\frac{dc_A}{dt} = v_{upt} - v_1 \qquad \frac{dc_B}{dt} = v_1 - v_2 - v_5$$
$$\frac{dc_C}{dt} = v_2 - v_3 \qquad \frac{dc_D}{dt} = v_3 - v_4 \quad (4)$$
$$\frac{dc_E}{dt} = v_4 + v_5 - v_{bm} \qquad \frac{dc_F}{dt} = v_3 + v_4 - v_{ex}$$

The system can be written in matrix notation:

$$\frac{d\mathbf{c}_{IM}}{dt} = \mathbf{N} \cdot \mathbf{v} \qquad (5)$$

with vectors for all intermediates $\mathbf{c} = (c_A, c_B, \ldots, c_F)^T$ and fluxes $\mathbf{v} = (v_1, v_2, \ldots, v_{ex})^T$. **N** is called stoichiometric matrix, which consists of *m* columns representing all reactions and *n* rows representing all balanced IM's.

Metabolic Stationary States. Metabolic stationary flux analysis (MFA) aims at the quantification of intracellular fluxes in the metabolism of an organism. In order to use this approach, the assumption on metabolic stationary $d\mathbf{c}_{IM}/dt = 0$ must hold, i.e. the concentrations of all IM's do not change over the experimental time. It is generally accepted that this condition is fullfilled in bioreactors under continuous cultivation conditions (chemostat) and in the exponential growth phase of cells cultivated in batch/fed-batch mode.

Assuming a metabolic stationary system state it simplifies as follows according to Equation 5:

$$0 = \mathbf{N} \cdot \mathbf{v} = \mathbf{f}(\mathbf{v}) \qquad (6)$$

Due to the fact that there are usually more reactions than metabolite pools, the algebraic equation system 6 is underdetermined. Including measurements (typically substrate uptake and product excretion rates determined from extracellular metabolome analysis as well as anabolic reactions known from biomass formation) is usually not enough to recover all flux rates. Thus, only a linear combination of all fluxes can be calculated. In the running example a combination of all intracellular fluxes can be determined by constraining the influx v_{feed} and the two effluxes v_{bm} and v_{ex}.

Metabolic Non-stationary States. Strictly speaking a biochemical system under defined cultivation conditions can only approximatively be considered as metabolic stationary, i.e. usually a quasi-stationary state $d\mathbf{c}_{IM}/dt \approx 0$ is attained.

To be more general, dynamic metabolic networks can be formulated (cf. Equation 5):

$$\frac{d\mathbf{c}_{IM}}{dt} = \mathbf{N} \cdot \mathbf{v}(\mathbf{c}_{IM}, \mathbf{c}_S, \alpha) = f(\mathbf{c}_{IM}, \mathbf{c}_S, \alpha), \quad \mathbf{c}_{IM}(0) = \mathbf{c}_{IM,0} \qquad (7)$$

with external (possibly time dependent) concentrations $\mathbf{c}_S$ and the vector of kinetic parameters. Usually, the time dependent reaction rates **v** are modelled mechanistically assuming fast equilibrium of intermediate enzyme complexes (validity of Michaelis-Menten Quasi-Steady State). For example the reaction rate v_1 of the example can be modelled by a reversible reaction of Michaelis-Menten type

$$v_1 = v_{max,1} \cdot \frac{c_A - \frac{c_B}{K_{eq,1}}}{K_{A,1} \cdot \left(1 + \frac{c_B}{K_{B,1}}\right) + c_A} \qquad (8)$$

where kinetic parameters are the maximal reaction rate v_{max}, the equilibrium constant K_{eq}, and the affinity constants K_A and K_B.

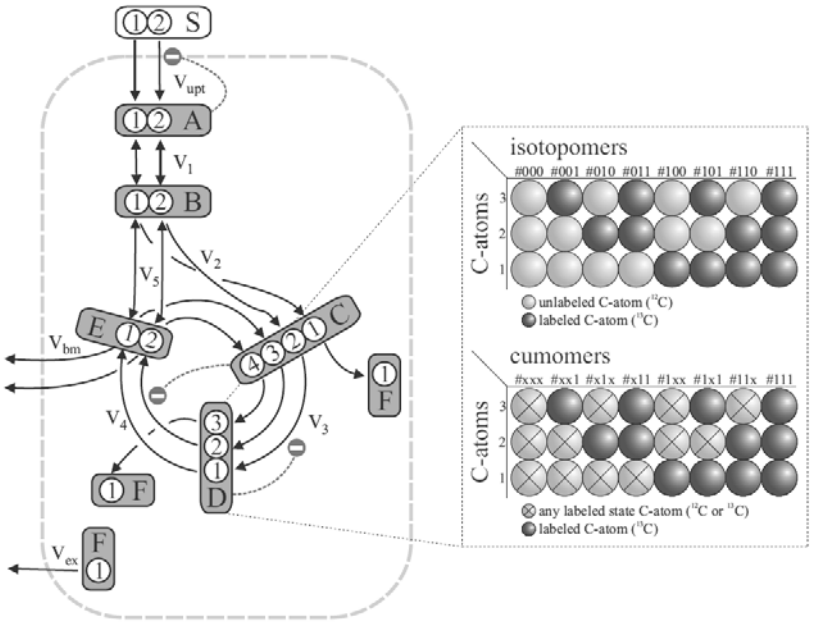


Figure 2. Left: Different levels of the example network regarded throughout the text. Level 1 (orange boxes): metabolic network model describing the reaction of rates *v* between intracellular metabolite pools. Level 2 (white cycles): C- atom transition network decribing the carbon atom traces. Level 3 (red interactions): regulatory/inhibitory constraints. Right: Two different ways to express the isotopic composition of metabolite pools: isotopomers (isotopic isomers) and cumomers (cumulative isotopomers).

As a result of the mechanistic enzyme descriptions, dynamic models contain a high number of model parameters. Although rapidly sampled measurement data of time dependent concentration changes are used for parameter fitting, often the adequate information is missing for identifying all kinetic parameters [2]. To reduce the amount of unknown model parameters, several methods for simplifying enzyme kinetics have been proposed [5, 4]. Inserting mechanistic or approximative kinetic approaches in Equation 4 results in the ordinary differential equation (ODE) system for the running example under metabolic dynamic conditions.

1.3 Mass Balance of Labeled Intermediate Pools

The application of cultivation experiments with labeled substrates is motivated by the better relation between measurement data and model unknowns. Depending on the mathematical formalism for modeling the labeling state of a certain metabolic network this results in a combinatorial blowup of the number of equations [18, 15, 1].

Clearly, the most general representation of isotopic systems is given by the concepts of isotopomers and, equivalently, cumomers (cf. Figure 2). In both cases there is not only one single mass balance for the overall concentration of an IM, but rather 2^n equations for the IM's isotopomer or cumomer concentrations, respectively, with *n* the number of IM's C-atoms.

Following the well-known isotopomer concept the mass balance for intermediate labeling pools $\mathbf{c}_{IM}^{Iso}$ is formulated

$$\frac{d\mathbf{c}_{IM}^{Iso}}{dt} = \frac{d(c_{IM}\cdot \mathbf{x}_{IM}^{Iso})}{dt} = c_{IM}\frac{d\mathbf{x}_{IM}^{Iso}}{dt} + \mathbf{x}_{IM}^{Iso}\frac{dc_{IM}}{dt}$$
$$\frac{dc_{IM}}{dt} = \sum_{j=1}^{n}\frac{dc_{IM}^{Iso}}{dt} \qquad (9)$$

There $x_{IM}^{Iso} = x_{IM}^{Iso}/c_{IM}$ denotes an isotopomer fracztion of the intermediate IM.

When modeling systems where isotopic labeling plays a role, it is necessary to formulate mass balances around single intermediates that are based on a separation of the net rates $v^{net} = v^{\rightarrow} - v^{\leftarrow}$ into forward $v^{\rightarrow}$ and backward rates $v^{\leftarrow}$. The reason is simply given by the fact that a IM's labeling is influenced by the labeling fraction of all pools contributing to its (isotope) mass balance. Consequently, in a reversible reaction step, the labeling that arrives the substrate pool through the backward rate depends on the labeling of the product pool and is not necessarily consistent with the labeling state of the substrate, except high exchange rates can be assumed.

As an example, consider the isotopomer mass balance around the labeled pool B of the example (cf. Figure 2):

$$\frac{d\mathbf{c}_B^{\#ij}}{dt} = v_1^{\rightarrow}\cdot \mathbf{x}_A^{\#ij} + v_5^{\leftarrow}\cdot \mathbf{x}_E^{\#ij} - \\ -(v_1^{\leftarrow} + v_2^{\rightarrow} + v_5^{\rightarrow})\cdot \mathbf{x}_B^{\#ij} \quad i,j \in \{0,1\} \qquad (10)$$

Special cases. In analogy to section 1.2, Equation 9 contains two special cases:

1. Metabolic and Isotopic Stationary States: Referring to Eq. 9 and assuming a metabolic and isotopic stationary $d\mathbf{c}_{IM}^{Iso} = 0$ system state it follows:

$$\mathbf{0} = \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, \mathbf{x}_{IM}^{Iso}, \mathbf{x}_S^{Iso}) \qquad (11)$$

2. Metabolic Stationary and Isotopic Non-stationary States: Assuming only a metabolic stationary system state ($d\mathbf{c}_{IM}/dt \approx 0$) Equation 9 leads to:

$$c_{IM}\frac{d\mathbf{x}_{IM}^{Iso}}{dt} = \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, c_{IM}, \mathbf{x}_{IM}^{Iso}, \mathbf{x}_S^{Iso}), \\ \mathbf{x}_{IM}^{Iso}(0) = \mathbf{x}_{IM,0}^{Iso} \qquad (12)$$

In both cases the function **f** depends in a linear way on the fluxes $\mathbf{v}^{\rightleftharpoons}$ and the substrate isotopomer fractions $\mathbf{x}_S^{Iso}$ while it might be nonlinear in the intermediate isotopomer fractions $\mathbf{x}_{IM}^{Iso}$.

2 A Generalized Modeling Concept

2.1 Metabolic and Isotopic Non-stationary Systems

Here, we present a general modeling concept, which is based on Equation motivated by [16] in terms of isotopomers. In order to benefit from the increase in performance reported for the solution in ^{13}C-MFA, we implemented the cumomer approach (cf. Figure 2):

$$\frac{d\mathbf{c}_{IM}^{Cum}}{dt} = \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, \mathbf{c}_{IM}^{Cum}, \mathbf{c}_S^{Cum}, \alpha), \\ \mathbf{c}_{IM}^{Cum}(0) = \mathbf{c}_{IM,0}^{Cum} \qquad (13)$$

The function **f** in Equation 13 is linearly dependent on the substrate cumomers $\mathbf{c}^{Cum}$ and nonlinear with respect to the reaction rates $\mathbf{v}^{\rightleftharpoons}(\mathbf{c}_{IM}^{Cum}, \alpha)$. The use of cumomers instead of isotopomers is motivated by the fact that the nonlinear ODE system can be partioned into cascaded subsystems of ODE's [18]. In short, all cumomers with equal weight (identical number of labeled C-atoms) form one level of system equations.

These levels are solved consecutively, starting from level 0, which is solely determined by the stoichiometry and corresponding kinetics (cf. Equation 7), up to the highest level, which is determined by the C-atoms of the longest carbon backbone occuring in the network.

To illustrate the concept consider the cumomer balances around the pool C of the running example (cf. Figure 2). For clarity, the elements of the cumomer concentration vector of C are denoted as
$\mathbf{c}_C^{Cum} = (C\#xxxx, C\#xxx1, \ldots, C\#1111)^T, x \in \{0,1\}$
Corresponding cucomer fractions are introduced by
$\mathbf{x}_{IM}^{Cum} = \mathbf{c}_{IM}^{Cum}/c_{IM}$ and given in small letters, e.g. *b#x*1

level 0:

$$\frac{dC\#xxxx}{dt} = v_2^{\rightarrow} - v_3^{\rightarrow}$$

level 1:

$$\frac{dC\#xxx1}{dt} = v_2^{\rightarrow} \cdot e\#x1 - v_3^{\rightarrow} \cdot c\#xxx1$$
$$\frac{dC\#xx1x}{dt} = v_2^{\rightarrow} \cdot e\#1x - v_3^{\rightarrow} \cdot c\#xx1x$$
$$\frac{dC\#x1xx}{dt} = v_2^{\rightarrow} \cdot b\#x1 - v_3^{\rightarrow} \cdot c\#x1xx$$
$$\frac{dC\#1xxx}{dt} = v_2^{\rightarrow} \cdot b\#1x - v_3^{\rightarrow} \cdot c\#1xxx$$

level 2:

$$\frac{dC\#xx11}{dt} = v_2^{\rightarrow} \cdot e\#11 - v_3^{\rightarrow} \cdot c\#xx11$$
$$\frac{dC\#x1x1}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#x1} \cdot \boxed{e\#x1} - v_3^{\rightarrow} \cdot c\#x1x1$$
$$\frac{dC\#x11x}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#x1} \cdot \boxed{e\#1x} - v_3^{\rightarrow} \cdot c\#x11x \quad (14)$$
$$\frac{dC\#1xx1}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#1x} \cdot \boxed{e\#x1} - v_3^{\rightarrow} \cdot c\#1xx1$$
$$\frac{dC\#11xx}{dt} = v_2^{\rightarrow} \cdot b\#11 - v_3^{\rightarrow} \cdot c\#11xx$$

level 3:

$$\frac{dC\#x111}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#x1} \cdot \boxed{e\#11} - v_3^{\rightarrow} \cdot c\#x111$$
$$\frac{dC\#1x11}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#1x} \cdot \boxed{e\#11} - v_3^{\rightarrow} \cdot c\#1x11$$
$$\frac{dC\#11x1}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#11} \cdot \boxed{e\#x1} - v_3^{\rightarrow} \cdot c\#11x1$$
$$\frac{dC\#111x}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#11} \cdot \boxed{e\#1x} - v_3^{\rightarrow} \cdot c\#111x$$

level 4:

$$\frac{dC\#1111}{dt} = v_2^{\rightarrow} \cdot \boxed{b\#11} \cdot \boxed{e\#11} - v_3^{\rightarrow} \cdot c\#1111$$

The cumomer fractions of level 0 (*b#xx* = *c*#xxxx = *e#xx* = 1) are not given explicitly. Considering the levels above 1 the framed cumomer fractions are already determined in lower levels. It should be noticed, that the information on the C-atom transitions, i.e. which C-atom of the substrate is transfered to which C-atom of the product, is essential for the correct formulation of the model equations.

Already, from this simple example it becomes clear, that the whole system equations can hardly be formulated manually. To avoid tedious and error prone typing it is desirable to generate systems like Equation 14 in a fully automized way. We use the existing software toolbox 13CFLUX to take this task [19].

2.2 Kinetics for Labeling Dynamics

Basically for every type of mechanistic enzyme kinetics the separate formulation of steady-state forward and backward rates is possible applying e.g. the King-Altman method [6]. Although automatable, it is questionary if a description of one single reaction step by a complex mechanistic model like e.g. Bi-Bi Random Order comprising 18 rate constants to be fitted, results in a overall valid dynamic model.

For that reason we choose the convenience rate law [8] as a simplified mechanistic approach, which can be specified by a small number of parameters and is easy to handle for automatically assign a kinetic model to each reactions step of a dynamic network model. The general rate law of a convenience kinetic for *n* substrates and *m* products is given by [8]:

$$v = v_{max,1} \cdot \frac{\mathbf{v}_{max}^{\rightarrow} \cdot \prod_i^n \left(\frac{c_{S_i}}{K_{S_i}}\right) - \mathbf{v}_{max}^{\leftarrow} \cdot \prod_j^m \left(\frac{c_{P_j}}{K_{P_j}}\right)}{\prod_i^n \left(1 + \frac{c_{S_i}}{K_{S_i}}\right) + \prod_j^m \left(\frac{c_{P_j}}{K_{P_j}}\right) - 1} \quad (15)$$

In case of a reaction where only one substrate is converted to one product the known Michaelis-Menten kinetic is derived (cf. Equation 8).

As an example, consider the cumomer mass balance around the pool B of the example (cf. Figure 2):

$$\frac{d\mathbf{c}_{IM}^{Cum}}{dt} = \mathbf{f}\left(\mathbf{v}^{\rightleftharpoons},\ \mathbf{c}_{IM}^{Cum}, \mathbf{c}_{S}^{Cum}, \alpha\right),$$
$$\mathbf{c}_{IM}^{Cum}(0) = \ \mathbf{c}_{IM,0}^{Cum}$$

$$\frac{d\mathbf{c}_{B}^{\#ij}}{dt} = \frac{1}{1+\frac{c_A}{K_{A,1}}+\frac{c_B}{K_{B,1}}} \cdot$$
$$\cdot\left(\mathbf{v}_{max,1}^{\rightarrow}\cdot\frac{c_A}{K_{A,1}}\cdot\mathbf{x}_{A}^{\#ij} - \mathbf{v}_{max,1}^{\leftarrow}\cdot\frac{c_B}{K_{B,1}}\cdot\mathbf{x}_{B}^{\#ij}\right) +$$
$$+\frac{1}{1+\frac{c_B}{K_{B,5}}+\frac{c_E}{K_{E,5}}}\cdot \qquad (16)$$
$$\cdot\left(\mathbf{v}_{max,5}^{\leftarrow}\cdot\frac{c_E}{K_{E,5}}\cdot\mathbf{x}_{E}^{\#ij} - \mathbf{v}_{max,5}^{\rightarrow}\cdot\frac{c_B}{K_{B,5}}\cdot\mathbf{x}_{B}^{\#ij}\right) -$$
$$-\mathbf{v}_{max,2}^{\rightarrow}\cdot\frac{c_B}{K_{B,2}+c_B}\cdot\mathbf{x}_{B}^{\#ij},\ i,j\in\{0,1\}$$

For modeling metabolic regulation, generic terms for activation and inhibition can be multiplicatively combined with Equation 15:

$$\prod_a\left(\frac{k_{A_a}+c_{A_a}}{k_{A_a}}\right),\ \prod_b\left(\frac{k_{I_b}}{k_{I_b}+c_{I_b}}\right) \qquad (17)$$

When parameterizing a dynamic model only the right combinations of kinetic parameter values and initial pool sizes will lead to a certain stationary state of the undisturbed system.

In order to use our modeling concept for the simulation of different experimental states (including metabolic stationarity) the dynamic network is coupled to a stationary network representing the initial system state.

Thereby the maximal reaction rates $\mathbf{v}_{max}^{\rightleftharpoons} = \mathbf{f}\left(\mathbf{v}_{\mathbf{0}}^{\rightleftharpoons}, \mathbf{c}_{IM,0}, \alpha\right)$are formulated as functions for initial values of fluxes $\mathbf{v}_{\mathbf{0}}^{\rightleftharpoons}$ and intermediate concentrations $\mathbf{c}_{IM,0}$ as well as kinetic parameters (now only comprising affinity constants).

Hence, the model includes one additional term per kinetic rate equation, e.g. for the forward step of reaction v_1 the two equations are considered:

$$v_1^{\rightarrow} = v_{max,1}^{\rightarrow}\cdot\frac{\frac{c_A}{K_{A,1}}}{1+\frac{c_A}{K_{A,1}}+\frac{c_B}{K_{B,1}}}, \qquad (18)$$

$$v_{max,1}^{\rightarrow} = v_{1,0}^{\rightarrow}\cdot\frac{1+\frac{c_{A,0}}{K_{A,1}}+\frac{c_{B,0}}{K_{B,1}}}{\frac{c_{A,0}}{K_{A,1}}},$$

3 An Universal Framework for Simulation and Evaluation of Metabolic Networks

So far we utilized a quite simple example network for explaining our modeling concept. Even for that system the number of cumomer mass balances and kinetic equations to be generated is quite high. A more realistic example network modeling reactions of glycolysis and pentose phosphate pathways already contains 682 ODE's for cumomer concentrations and a total of 112 kinetic parameters.

Clearly, manual generation of such models becomes infeasible and automated model code generation is strongly recommended. Additionally, the resulting dynamic model is a system of highly nonlinear ODE's and therefore demands for a simulation environment that can handle this complexity.

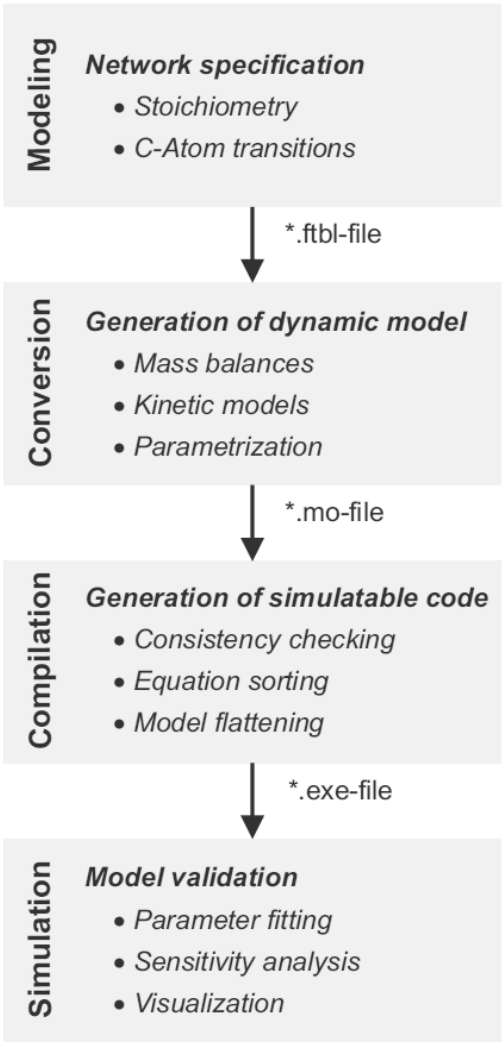
For that reason we developed a software framework, which allows the automatic generation of dynamic metabolic and isotopic network models of any size and complexity and, moreover, offers the perspective to use sophisticated methods for statistical analyis and interpretation of simulation results (cf. Figure 3).

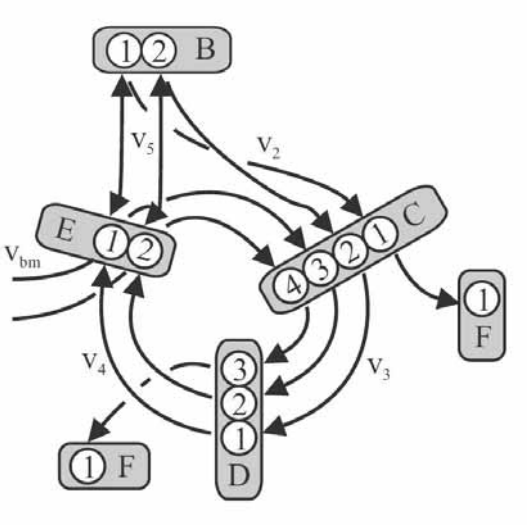
As a core for setting up the model and building executable simulation code we use the Modelica language in combination with the Dymola environment (Dynasim AB, 5.0, www.modelica. org).

3.1 Setup of Dynamic Model Equations

Starting from biochemical network descriptions (including stoichiometry, C-atom transitions, initial conditions) using the software 13CFLUX, consistent algebraic equation systems (AE) of mass balances for all cumomer pools are generated.

These AE's are then transformed into Modelica specific code consisting of ODE-systems for all cumomers, i.e. addition of a left hand side and kinetic equations (cf. Equation 13).






```
model dynamicModel
//Declaration part
  //Model state
parameter Integer flag=5;
 //Cumomer concentrations
Real A_1x(start=0.005267);
...
 //Reaction rates
  Real v1_fwd;
...
 //Kinetic parameters
parameter Real v1_km_A=1;
...
  equation
  //Balances
der(A_x1)=vupt_fwd*S_x1/S_xx+...;
...
  //Kinetics
v1_fwd=v1_fwd_vmax*(A_xx/v1_km_A)/...;
...
end dynamicModel;
```

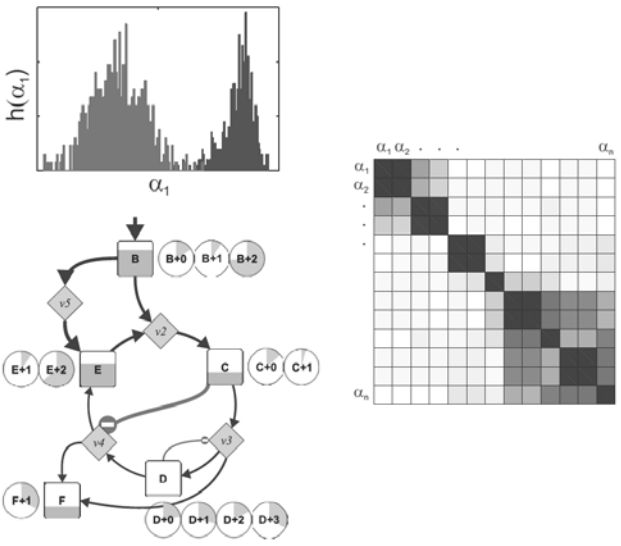


Figure 3: Workflow of model generation. Steps involved can be divided in four categories: (1) spezification of model equations, (2) conversion to a Modelica file, (3) its compilation into an executable code which is then used for versatile simulation tasks (4).

Definition of Model States. Due to the general model structure (cf. Equation 13 and Section 2.3) in total five types of experimental states can be simulated using one model with different scenarios settings:

1. The external substrates $\mathbf{c}_S^{Cum}$ are unlabeled and
   - do not change over time resulting in an AE-system ($d\mathbf{c}_{IM}^{Cum}/\,dt = 0 = \mathbf{f}(\mathbf{v})$) depending only on steady state fluxes ($\mathbf{v}$ = *const*). This also equals the starting point, where the system is in a unlabeled metabolic stationary state ($0 = \mathbf{f}(\mathbf{v}_0)$),
   - change over time resulting in an ODE-system ($d\mathbf{x}_{IM}^{Cum}/\,dt = 0 \implies d\mathbf{c}_{IM}/dt = \mathbf{f}(\mathbf{v}, \mathbf{c}_{IM}, \mathbf{c}_S, \boldsymbol{\alpha})$) depending on dynamic intermediate concentrations ($\mathbf{c}_{IM}$).
2. The external substrates substrates $\mathbf{c}_S^{Cum}$ are based on specifically labeled mixtures $\mathbf{x}_S^{Cum}$ and
   - do not change over time leading to an AE-system ($d\mathbf{c}_{IM}^{Cum}/\,dt = 0 := \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, \mathbf{c}_{IM}^{Cum}, \mathbf{c}_S^{Cum}, \alpha)$) depending on steady state fluxes ($\mathbf{v}$ = *const*) and cumomer labeling fractions ($\mathbf{x}_{IM}^{Cum} = const$),
   - the substrate cumomer fractions ($\mathbf{x}_S^{Cum}$) change over time leading to an DAE-system
     $$d\mathbf{c}_{IM}/dt = 0 \implies \mathbf{c}_{IM} \cdot d\mathbf{x}_{IM}^{Cum}/\,dt = \mathbf{f}(\mathbf{v}^{\rightleftharpoons}, \mathbf{c}_{IM}, \mathbf{x}_{IM}^{Cum}, \mathbf{c}_S^{Cum})$$
     depending on steady state fluxes ($\mathbf{v}$ = *const*), dynamic cumomer labeling fractions ($\mathbf{x}_{IM}^{Cum}$) and metabolite pool sizes ($\mathbf{c}_{IM}$= *const.*),
   - change over time resulting in the general ODE-system of Equation 13.

Manual Adaption of Model Code. The initial parametrization of the automatically generated model code allows directly starting forward simulations of the different model states. In order to describe a real biochemical system under certain experimental conditions manual adaption of the following items are possible:

- **Kinetic types:** As the standard kinetic model for all reactions, the convenience rate law is chosen (cf. Section 1.2). Nevertheless, if other kinetic types for single reaction steps are needed, they can be easily substituted.
- **Substrate labeling mixture:** For each external substrate the standard mixture is formed by an amount of unlabeled, single labeled (1-^{13} C) and fully labeled (U-^{13} C) substrate. Different compositions can be fixed manually.

3.2 Model Validation

After generation, Modelica models can be directly compiled into highly efficient executable simulation code (cf. Figure 3). Model simulations, parameter fittings as well as comprehensive statistical evaluations are performed under MATLAB (Mathworks, R2008b) on a high-performance workstation under Linux.

Sensitivity Analysis. As an essential ingredient of model based inference, sensitivity analysis for model variables and parameters is performed using an automatic differentiation (AD) method developed for Modelica source code [3]. In short, ADModelica strives to semantically augment Modelica models with Modelica code for computing certain sensitivities, with minimal user efforts.

Aiming at the full-support of Modelica language constructs, the current version supports most basic constructs of Modelica. Clearly, the number of equations in a differentiated model (> 70000 for a realistic model) increases proportionally with the number of model parameters.

Visualization of Simulation Results. For an intuitive interpretation of simulation data under dynamic as well as steady state conditions a network visualization approach can be applied.

Huge amounts of different kinds of simulation data, e.g. dynamically changing intermediate concentrations and labeling fractions can be analyzed in a quick and comprehensive way (Figure 4).

The visualization of regulatory interactions in a given metabolic network is based on a concept defining the Regulatory Strength of effectors regulating certain reaction steps [9].

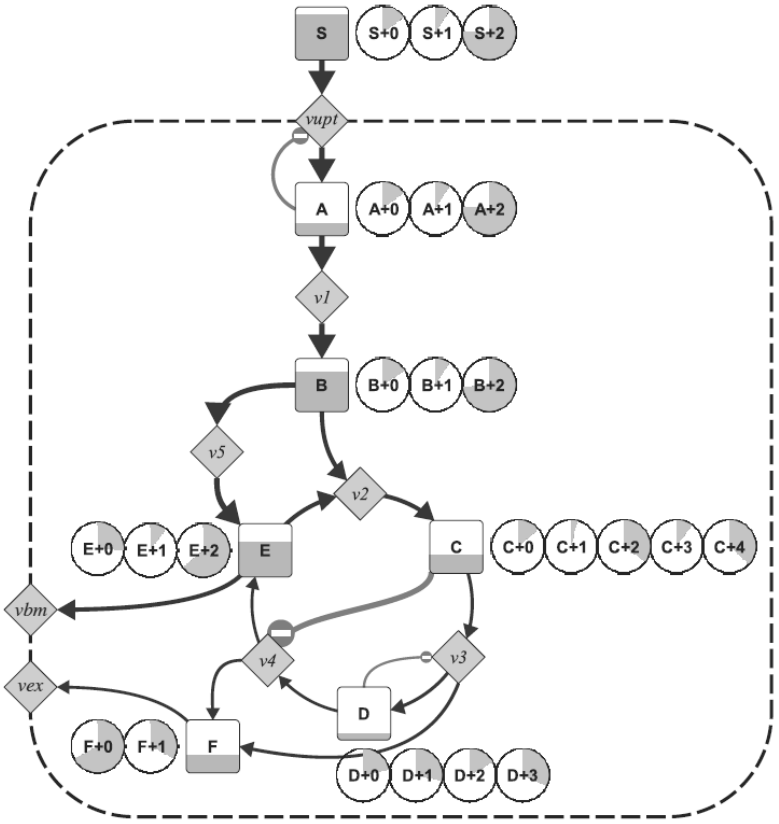


Figure 4. Visualization of the example network showing a metabolic and isotopic snapshot during dynamic simulation. Scaled simulation data of intermediate concentrations and reaction rates are represented by correspondingly filled boxes and varying arrow widths, respectively. The labeling state is visualized in form of mass isotopomer fractions using pie charts. Regulatory interactions are visualized by red (inhibition) or green (activation, not included) circles and the corresponding strength of each effector is represented by the circle size.

4 Conclusions

The proposed software framework allows automatized generation of dynamic network models for simulating all currently realizable experiments in the fields of Metabolomics and Fluxomics, i.e. metabolic and isotopic (non-)stationary systems.

Due to the universality of the underlying modeling concept the different experimental states an be simulated using one software platform. Since, model setup and compilation is performed under the Modelica/Dymola environment realistic network models comprising more than 1000 ODE's can be simulated in accept- able time and therefore used for experimental validation.

The developed automatic differentiation method for Modelica code allows exact calculation of Jacobians that can be used for sensitivity analyis and gradient based optimization routines. Finally, the whole framework covers a visualization approach where the resulting simulation data can be easily interpretated in the network context.

References

[1] Antoniewicz, M. R., Kelleher, J. K., and Stephanopoulos, G. (2007). Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab Eng 9*, 1 (2007), 68–86. 1096-7176.

[2] Degenring , D., Froemel , C., Dikta , G., and Takors, R. (2004). Sensitivity analysis for the reduction of complex metabolism models. *J Process Contr 14*, 7 (2004), 729–745.

[3] Elsheik, A., Noack, S., and Wiechert, W. (2008). Sensitivity analysis of modelica applications via auto- matic differentiation. In *6th International Modelica Conference* (Bielefeld, Germany, 2008), p. 669.

[4] Hadlich, F., Noack, S., and Wiechert, W.(2008). Translating biochemical network models between different kinetic formats. *Metab Eng* (2008). 1096-7184 (Electronic) Journal article.

[5] Heijnen, J. J. (2005). Approximative kinetic formats used in metabolic network modeling. *Biotechnol Bioeng 91*, 534–45. 0006-3592.

[6] King, E. L., and Altman, C. (1956). A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *J Phys Chem 60* (1956), 1375–1378.

[7] Kresnowati, M. T., Suarez-Mendez, C. M., van Winden, W. A., Vangulik, W. M., and Heijnen, J. J. (2008). Quantitative physiological study of the fast dynamics in the intracellular pH of Saccharomyces cerevisiae in response to glucose and ethanol pulses. *Metab Eng 10*, 1 (2008), 39–54. 1096-7176 (Print) Journal Article Research Support, Non-U.S. Gov't.

[8] Liebermeister, W., and Klippe, E. (2006). Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model 3* (2006), 41. 1742-4682 (Electronic) Journal Article, Research Support, Non-U.S. Gov't.

[9] Noack, S., Wahl, A., Qeli, E., and Wiechert, W. (2007). Visualizing regulatory interactions in metabolic networks. *BMC Biol 5* (2007), 46. 1741-7007 (Electronic) Journal Article Research Support, Non-U.S. Gov't.

[10] Nöh, K., Grönke , K., Luo, B., Takors , R., Oldiges, M., and Wiechert, W. (2007). Metabolic flux analysis at ultra short time scale: isotopically non-stationary 13C labeling experiments. *J Biotechnol 129*, 2 (2007), 249–67. 0168-1656 (Print) Journal Article Research Support, Non-U.S. Gov't.

[11] Oldiges, M., Kunze, M., Degenring, D., Sprenger, G. A., and Takors, R. (2004). Stimulation, monitoring, and analysis of pathway dynamics by metabolic profiling in the aromatic amino acid pathway. *Biotechnol Prog 20*, 6 (2004), 1623–33. 8756-7938 (Print) Evaluation Studies Journal Article Research Support, Non-U.S. Gov't.

[12] Oldiges, M., Lütz, S., Pflug, S., Schroer, K., Stein, N., and Wiendahl, C. (2007). Metabolomics: current state and evolving methodologies and tools. *Appl Microbiol Biotechnol 76*, 3 (2007), 495–511, 0175-7598 (Print) Journal Article Review.

[13] Schwender, J. (2008). Metabolic flux analysis as a tool in metabolic engineering of plants. *Curr Opin Biotechnol 19*, 2 (2008), 131–7. 0958-1669 (Print) Journal Article Research Support, U.S. Gov't, Non-P.H.S. Review.

[14] Stephanopoulos, G., and Vallino, J. J. (1991). Network rigidity and metabolic engineering in metabolite overproduction. *Science 252*, 5013 (1991), 1675–81. 0036-8075 (Print) Journal Article Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Review.

[15] Van Winden, W. A., Heijnen, J. J., and Verheijen, P. J. (2002). Cumulative bondomers: a new concept in flux analysis from 2D [13C,1H] COSY NMR data. *Biotechnol Bioeng 80*, 7 (2002), 731–45. 0006-3592 (Print) Journal Article.

[16] Wahl, S., Nöh, K., and Wiechert, W. (2008). 13C labeling of experiments at metabolic nonstationary conditions: An explaratory study. *BMC Bioinf 9*, 152 (2008).

[17] Wahl, S. A., Haunschild, M. D., Oldiges, M., ans Wiechert, W. (2006). Unravelling the regulatory structure of biochemical networks using stimulus response experiments and large-scale model selection. *Syst Biol (Stevenage) 153*, 4 (2006), 275–85. 1741-2471 (Print) Journal Article.

[18] Wiechert, W., Möllney, M., Isermann, N., Wurzel, M., and Degraaf, A. A. (1999). Bidirectional reaction steps in metabolic networks: III. Explicit solution and analysis of isotopomer labeling systems. *Biotechnology and Bioengineering 66*, 2 (1999), 69–85.

[19] Wiechert, W., Möllney, M., Petersen, S., and Degraaf, A. A. (2001). A universal framework for 13C metabolic flux analysis. *Metab Eng 3*, 3 (2001), 265–83. 1096-7176 (Print) Journal Article Research Support, Non-U.S. Gov't.

Submitted: March, 2010 – *SNE Accepted*

Submitted MATHMOD 2009: November 2008
Accepted MATHMOD 2009: January 2009
Submitted SNE: January 2010
Accepted: February 5, 2010