BIOINFORMATICS ORIGINAL PAPER Vol. 25 no. 21 2009, pages 2816–2823
doi:10.1093/bioinformatics/btp451

Systems biology

Accessible methods for the dynamic time-scale decomposition of biochemical systems

Irina Surovtsova[1,*], Natalia Simus[1], Thomas Lorenz[2], Artjom König[1], Sven Sahle[1] and Ursula Kummer[1]

[1]Department of Modeling of Biological Processes and [2]Interdisciplinary Center for Scientific Computing, University of Heidelberg, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany

Received on December 11, 2008; revised on May 22, 2009; accepted on June 29, 2009

Advance Access publication July 24, 2009

Associate Editor: Trey Ideker

ABSTRACT

Motivation: The growing complexity of biochemical models asks for means to rationally dissect the networks into meaningful and rather independent subnetworks. Such foregoing should ensure an understanding of the system without any heuristics employed. Important for the success of such an approach is its accessibility and the clarity of the presentation of the results.

Results: In order to achieve this goal, we developed a method which is a modification of the classical approach of time-scale separation. This modified method as well as the more classical approach have been implemented for time-dependent application within the widely used software COPASI. The implementation includes different possibilities for the representation of the results including 3D-visualization.

Availability: The methods are included in COPASI which is free for academic use and available at www.copasi.org.

Contact: irina.surovtsova@bioquant.uni-heidelberg.de

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

An important computational aspect in systems biology is the fact that the increasing size and complexity of studied biochemical systems lead not only to experimental results which are hard to understand, but also to computational results which are not easy to comprehend. Here, the so-called complexity reduction aims for two different directions: increasing the speed of simulations and dissecting the biochemical network into smaller subsystems that can be studied independently. The first one is usually achieved by reducing the number of equations mathematically. This has often been tackled by analyzing the presence of a wide range of characteristic time scales in the respective systems. Among the most prominent approaches built on this concept are the Computational Singular Perturbation (CSP) method of Lam and Goussis (1994) and the so-called Intrinsic low-dimensional manifolds (ILDM) method of Maas and Pope (1992). Several variants of these two methods have been used successfully, e.g. in atmospheric and combustion chemistry modeling (see e.g. Schmidt *et al*. 1998). The CSP method was recently applied to the model reduction of circardian rhythm in Drosophila by Goussis and Najm (2006). Besides, Lovrics *et al.* (2006) have used time-scale analysis to determine the stability of the budding yeast-cycle trajectories.

A smaller number of equations, though, does not guarantee the reduction of biochemical species in the system, since many different species might contribute to one and the same transformed equation. Therefore, a likewise or even more important aspect is to dissect the biochemical system into several modules that can be studied independently. This is needed to understand the interplay of specific subsystems. Most of the past approaches to complexity reduction of biochemical systems have focused mainly on methods studying the steady-state behavior of the system (Kauffman *et al*., 2002; Price *et al*., 2003), or on dissecting the system based on its network topology using heuristic rules (Holme *et al*., 2003; Schuster *et al*., 2002). The first approach is helpful for biochemical systems that indeed can be expected to display steady-state behavior. Nevertheless, most organisms will not display steady states, but rather transient behavior of different kinds at all times. Moreover, for large and complex biochemical reaction networks an additional requirement is that reduction methods can be applied in an automatic way according to well-defined principles. For this purpose an automatic method based on a time-dependent ILDM was developed by Zobeley *et al.* (2005) for biochemical systems. An additional analysis of species contribution to the slow dynamics has been introduced making the actual decomposition of a system feasible (see Zobeley *et al.*, 2005, as well as Shaik *et al.*, 2005). Another automatic complexity reduction method for biochemistry was presented recently by Lebiedz *et al.* (2008). This combines dynamic sensitivity analysis, singular value decomposition and further computation of metabolite contribution to the active dynamics of the system.

However, in addition to automation, it is equally important to ensure accessibility of the respective methods, e.g. by integrating them in commonly available and frequently used software packages. That goal is not trivial since the representation of the results, e.g. via visualization is quite difficult.

This article presents the development of a new method, as well as the implementation of this and an established method of Zobeley *et al.* (2005) for the systematic and automated reduction procedures that may be applied to any explicit biochemical mechanism.

*To whom correspondence should be addressed.

Both methods rely on the principles of the classical ILDM method which decomposes biochemical systems computationally with respect to time scales. However, in comparison to the classical ILDM method applied, e.g. to chemical systems, the methods have an additional focus on the reduction of the underlying biochemical network in a time-dependent manner. The first approach has already been published in Zobeley *et al*. (2005) and also used in Shaik *et al*. (2005) and Surovtsova *et al*. (2006). It is now implemented in a user-friendly way in the software framework COPASI (Hoops *et al*., 2006). The implementation includes a visualization of the results in the form of tables and 3D bar graphs.

The second algorithm—also implemented in COPASI—uses essentially the same numerical tools as the classical ILDM. Its main goal, however, is to distinguish between 'slow' and 'fast' metabolites—instead of 'modes'. Hence, the final result of the analysis can be formulated without linear transformation of the reaction system. Both methods can be applied to arbitrary biochemical reaction systems and work independently of assumptions about the specific dynamic regime.

The article is organized in the following way: first, we describe the algorithms for time-scale separation as implemented in COPASI (Hoops *et al*., 2006). Then, we proceed to describe details of the respective implementation and finally, both methods are applied to several biochemical examples like calcium oscillations and yeast glycolysis. We also discuss the tolerance criterion used for detecting the dimension of the fast space.

2 TIME-SCALE DECOMPOSITION: ALGORITHM

Both algorithms presented here rely on the presence of a wide range of characteristic time scales in biological systems and are based on the local analysis of the Jacobian, which is partitioned into fast and slow components at the initial point of a user-chosen interval. We assume that the dynamics is determined by a system of n ordinary differential equations (ODE) together with a given state $\mathbf{c}_0$ at time t_0:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{c}(t) = \mathbf{f}(\mathbf{c}(t)), \quad \mathbf{c}(t_0) = \mathbf{c}_0 \in \mathbb{R}^n \tag{1}$$

Moreover, the metabolite concentration $\mathbf{c}$, reaction rates $\mathbf{v}$ and stoichiometry $\mathbf{N}$ are related by: $\mathrm{d}\mathbf{c}/\mathrm{d}t = \mathbf{N}\cdot\mathbf{v}$.

We assume that all linear dependencies due to conservation relationships have already been removed from the system.

2.1 Distinction between 'fast' and 'slow': QSSA

In this article, the terms 'fast' and 'slow' are used for both metabolites [described by their time-dependent concentrations $(c_1(t)\ldots c_n(t)) \in \mathbb{R}^n$] and 'modes' [resulting from appropriately constructed transformations of the metabolite concentrations as in Deuflhard and Heroth (1996) and Zobeley *et al*. (2005)]. Their basic notion, however, is always closely related to the question how rapidly they change in time. The rate of change is usually described by the time derivative.

The starting point is partitioning the reaction system into slow and fast contributions. Such a splitting is related to a singular perturbation description of the transformed ODE system

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_{\mathrm{slow}} &= \mathbf{g}_{\mathrm{slow}}\left(\mathbf{x}_{\mathrm{slow}}, \mathbf{x}_{\mathrm{fast}}\right) \\ \varepsilon\cdot\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_{\mathrm{fast}} &= \mathbf{g}_{\mathrm{fast}}\left(\mathbf{x}_{\mathrm{slow}}, \mathbf{x}_{\mathrm{fast}}\right) \end{cases}$$

Here, ε is a singular perturbation parameter. The concentrations of the fast species change with time, but these species can be described by algebraic relations instead of differential equations. The quasi steady-state assumption (QSSA) yields the associated differential–algebraic system (DAE) for the reduced problem by replacing

$$\varepsilon\cdot\tfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_{\mathrm{fast}} \approx 0.$$

The step to DAEs, however, provides difficulties with regard to the initial conditions. In general, the given initial state $(\mathbf{x}_{0,\mathrm{slow}}, \mathbf{x}_{0,\mathrm{fast}})$ does not satisfy the algebraic equation $\mathbf{g}_{\mathrm{fast}}(\mathbf{x}_{0,\mathrm{slow}}, \mathbf{x}_{0,\mathrm{fast}}) = 0$. For this reason, we replace the known initial value $\mathbf{x}_{0,\mathrm{fast}}$ by an approximation $\mathbf{x}^{\mathrm{alg}}_{0,\mathrm{fast}}$ consistent with the algebraic equation

$$\mathbf{0} = \mathbf{g}_{\mathrm{fast}}(\mathbf{x}_{0,\mathrm{slow}}, \mathbf{x}^{\mathrm{alg}}_{0,\mathrm{fast}}). \tag{2}$$

These consequences of QSSA provide our mathematical characterization of 'slow' and 'fast' components: we regard those components as 'fast' for which the QSSA can be justified (on the basis of a fixed error tolerance). The other components are considered as 'slow' and, they are still described by the full ODE.

2.2 Time-dependent classical ILDM

Step 1: As a starting point of analysis, a linearization of Equation (1) with respect to the state vector $\mathbf{c}_0$ is performed:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{c}(t) &= \mathbf{f}\left(\mathbf{c}_0\right) + \mathbf{J}_{\mathbf{c}_0}\cdot\left(\mathbf{c}(t) - \mathbf{c}_0\right) \\ \mathbf{c}(t_0) &= \mathbf{c}_0 \ \left(\mathbf{J}_{\mathbf{c}_0}\text{ Jacobian of }\mathbf{f}\text{ in }\mathbf{c}_0\right) \end{cases}$$

Step 2: An orthogonal similarity transformation is applied to the Jacobian $\mathbf{J}_{\mathbf{c}_0}$. The resulting matrix $\mathbf{S}$ has real Schur form, i.e. it is a blocks upper triangular matrix (Golub and van Loan, 1996):

$$\mathbf{Q}^{-1}\cdot\mathbf{J}_{\mathbf{c}_0}\cdot\mathbf{Q} = \mathbf{S} = \begin{pmatrix} \mathbf{S}_{\mathrm{slow}} & \mathbf{S}_{\mathrm{coup}} \\ 0 & \mathbf{S}_{\mathrm{fast}} \end{pmatrix}$$

The matrix $\mathbf{S}$ is reordered according to characteristic time-scales (TS): $\tau_i = -(1/\mathrm{Re}\lambda_i)$ by means of Givens rotations (λ_i denote the eigenvalues of the Jacobian).

Step 3: Using the solution $\mathbf{Z}$ of the Sylvester equation

$$\mathbf{S}_{\mathrm{slow}}\,\mathbf{Z} - \mathbf{Z}\,\mathbf{S}_{\mathrm{fast}} = -\mathbf{S}_{\mathrm{coup}}$$

we realize that the transformed Jacobian is decoupled additionally:

$$\mathbf{T}^{-1}\cdot\mathbf{J}_{\mathbf{c}_0}\cdot\mathbf{T} = \begin{pmatrix} \mathbf{S}_{\mathrm{slow}} & 0 \\ 0 & \mathbf{S}_{\mathrm{fast}} \end{pmatrix}$$

where

$$\mathbf{T} = \mathbf{Q}\left(\mathbf{Id_n} + \begin{pmatrix} 0 & \mathbf{Z} \\ 0 & 0 \end{pmatrix}\right), \quad \mathbf{T}^{-1} = \left(\mathbf{Id_n} - \begin{pmatrix} 0 & \mathbf{Z} \\ 0 & 0 \end{pmatrix}\right)\mathbf{Q}^{-1}$$

Step 4: Applying $\mathbf{T}^{-1}$ to the state $\mathbf{c}$ and reaction rate $\mathbf{f}$ results in a decoupled representation of the system dynamics:

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{\mathrm{slow}} \\ \mathbf{x}_{\mathrm{fast}} \end{pmatrix} = \mathbf{T}^{-1}\cdot\mathbf{c}, \quad \mathbf{g} = \begin{pmatrix} \mathbf{g}_{\mathrm{slow}} \\ \mathbf{g}_{\mathrm{fast}} \end{pmatrix} = \mathbf{T}^{-1}\cdot\mathbf{f}(\mathbf{T}\cdot)$$

Step 5: The number r of 'slow' modes plays a crucial role in this approach. The criterion suggested by Deuflhard and Heroth (1996) is adapted (compare also Zobeley *et al*. 2005): for a given error tolerance $\mathrm{tol} > 0$, the number r of slow modes is iteratively decreased

as long as the corresponding decomposition satisfies

$$\varepsilon \cdot \left| \mathbf{g}_{\text{slow}}\left(\mathbf{x}_{0,\text{slow}}, \mathbf{x}_{0,\text{fast}}\right) - \mathbf{g}_{\text{slow}}\left(\mathbf{x}_{0,\text{slow}}, \mathbf{x}^{\text{alg}}_{0,\text{fast}}\right) \right| < \text{tol} \qquad (3)$$

where the vector $\left(\mathbf{x}_{0,\text{slow}}, \mathbf{x}_{0,\text{fast}}\right)$ is the transformed initial vector $\mathbf{c}_0$ of ODE system (1) and $\left(\mathbf{x}_{0,\text{slow}}, \mathbf{x}^{\text{alg}}_{0,\text{fast}}\right)$ denotes its closest approximation that is consistent with the algebraic Equation (2). $\varepsilon = \tau_{r+1} = 1/|\text{Re}\,\lambda_{r+1}|$ is a singular perturbation parameter.

Step 6: The matrix $\mathbf{T}$ relates modes $\mathbf{x}$ to the original variables $\mathbf{c}$. In order to complete our investigation, we perform the study of modes in terms of contributions of all species concentrations by analyzing the entries of transformation matrices $\mathbf{T} = (\beta_{ik})_{1 \le i,k \le n}$ and $\mathbf{T}^{-1} = (\gamma_{ik})_{1 \le i,k \le n}$ (see Surovtsova *et al*. 2006; Zobeley *et al*. 2005).

Vector 'Slow space' ('Fast space', respectively) indicates the contribution of each concentration variable to the set of all slow (fast) modes:

$$\text{'Slow space'}(i) := \frac{\sum_{k=1}^{r} \gamma_{ik}}{\sum_{j,k=1}^{n} \gamma_{jk}}, \quad \text{'Fast space'}(i) := \frac{\sum_{k=r+1}^{n} \gamma_{ik}}{\sum_{j,k=1}^{n} \gamma_{jk}}$$

2.3 Modified time-dependent ILDM method

For proposing a selection of fast and slow metabolites, we use essentially the same numerical tools of local Schur decomposition of the Jacobian as for the classical variant above. The vector 'Fast space' specifies the metabolites making the largest contributions to the set of fast modes. Hence, it provides the next candidate of concentration which is then tested whether its additional nomination as fast metabolite still satisfies the criterion of error tolerance. The algorithm of our modified method could be described as follows:

Step 1: linearize the dynamical system at $\mathbf{c}_0$.

Step 2: detect all time scales of the system and transform the ODEs (sorted by absolute eigenvalues alias time scales) using Schur decomposition.

Step 3: analyze the Schur transformation matrix $\mathbf{Q}$ for obtaining the contribution of each metabolite to the fast space. The vector 'Fast space' specifies the metabolites making the largest contributions to the set of fast modes.

Step 4: check whether the QSSA for the most dominant metabolite still satisfies the tolerance criterion, i.e. the dimension of fast space (and the number of fast metabolites) is increased as long as

$$\varepsilon \cdot \left| \mathbf{f}_{\text{slow}}\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}_{0,\text{fast}}\right) - \mathbf{f}_{\text{slow}}\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}^{\text{alg}}_{0,\text{fast}}\right) \right| < \text{tol} \qquad (4)$$

holds. Here, $\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}_{0,\text{fast}}\right)$ denotes the initial state vector $\mathbf{c}_0$ of ODE (1) with respect to the current distinction between 'slow' and 'fast' metabolites and, the vector $\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}^{\text{alg}}_{0,\text{fast}}\right)$ is its consistent initial value of DAEs. As before, $\varepsilon = \tau_{r+1} = 1/|\text{Re}\,\lambda_{r+1}|$ represents the slowest time scale of the fast space.

2.3.1 Remark In this modified ILDM method, we dispense with solving the Sylvester equation and hence with decoupling $\mathbf{S}_{\text{slow}}$ and $\mathbf{S}_{\text{fast}}$. Thus in general, we cannot exclude that the slow modes depend very much on the exact values of fast modes. Nevertheless, the non-zero matrix $\mathbf{S}_{\text{coup}}$ does not influence the number of fast metabolites. We also note that the non-zero matrix $\mathbf{S}_{\text{coup}}$ does not influence the derivation of error estimate (see Deuflhard and Heroth, 1996).

2.4 Tolerance criterion for the new method

Deuflhard and Heroth (1996) provide a mathematical justification for their tolerance criterion (3) on the basis of asymptotic analysis (with respect to the perturbation parameter ε).

Since ε is not accessible in our case, we suggest an alternative justification of the tolerance criterion (4) whose mathematical simplicity might be regarded as an advantage.

$\mathbf{c}_{0,\text{slow}} + (t - t_0) \cdot \mathbf{f}_{\text{slow}}\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}_{0,\text{fast}}\right)$ is the (probably simplest) approximation of the exact concentration $\mathbf{c}_{\text{slow}}(t)$ after a short period $t - t_0 \ge 0$ as the integral of the derivative easily reveals. Similarly, $\mathbf{c}_{0,\text{slow}} + (t - t_0) \cdot \mathbf{f}_{\text{slow}}\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}^{\text{alg}}_{0,\text{fast}}\right)$ approximates this concentration resulting from the corresponding DAE and the initial state $\left(\mathbf{c}_{0,\text{slow}}, \mathbf{c}^{\text{alg}}_{0,\text{fast}}\right)$. Hence, the right-hand side of inequality (4) estimates the absolute error of the 'slow' metabolites up to time ε.

2.5 Implementation in COPASI

The Jacobian is calculated by using finite differences. The Schur transformation and the solution of Sylvester equation are performed using CLAPACK. The options for both methods are the same:

- Intervals: the user specifies the number of intervals to which the method is applied. COPASI performs the analysis at each initial point of these intervals. The interval size should be large in comparison with the time scale interesting for the user.
- Deuflhard tolerance: this parameter is a positive numeric value specifying the maximal tolerated error of slow modes (Deuflhard and Heroth, 1996). The default value is 1×10^{-6}. We note that reasonable values of Deuflhard tolerance should be of the same order as the time scale the user is interested in.

2.6 Results and graphical visualization in COPASI

As described above results are displayed in two matrices and two vectors. The matrix '*Metabolites*' provides the contribution of each metabolite to every mode, whereas the table '*Modes*' summarizes the mode distribution for each metabolite.

In the classical variant [called ILDM(Deuflhard)], the vector '*Slow space*' ('*Fast space*', respectively) indicates the relative contribution of each concentration variable to the set of all slow (fast) modes. The metabolites with the largest contribution to the fast space could be supposed to be 'fast' and thus, its ODE can be replaced by the corresponding algebraic equation (i.e. with the same right-hand side). In the specific case that a subset of species does not contribute to the slow space (but contributes only to the fast space), the time-scale decomposition results can be used for a dissection of the reaction network.

The results of the modified variant consist of two similar matrices and two similar vectors, which describe the contribution of metabolites to slow (fast) modes (spaces). However, the corresponding numbers of slow and fast components now refer to the distinction between slow and fast metabolites (not modes).

All these matrices and vectors can either be exported to a text file or displayed in the graphical user interface as tables. In this case, a color coding is used where the numbers are additionally visualized by different shades of color. This makes it easy to immediately spot, e.g. the most important contributions to a specific mode for a large model (where the result tables are correspondingly large).

We also use 3D bar graphs for visualizing the matrices (see also Fig. 5). These bar graphs can be turned and zoomed interactively.

Furthermore, single rows or columns of the matrix can be highlighted. An additional diagram shows the distribution of the time scales of the different modes at chosen points of time (see Fig. 4).

In the graphical user interface, it is very simple to switch between the results for different points of time. Therefore, the user can easily get an overview over the time-dependent changes of the time-scale separation.

2.7 ILDM and classical QSS and QE assumptions

Two of the commonly used reduction methods, based on time-scale analysis, employ the quasi steady-state (QSS) approximation and quasi equilibrium (QE) assumption. QSSA identifies species whose production and destruction rate are in approximate balance, mathematically, it means that the right-hand side f_i of the corresponding equation is zero. QE assumption corresponds to reactions whose forward and reverse rates are nearly equal. Nevertheless, the identification of such fast reactions and steady-state species is normally based on experience and intuition.

The ILDM methods implemented in COPASI enable the automatic decoupling of fast and slow processes. The analysis of transformation matrices allows the detection of QE reactions and QSS metabolites.

The $(n-r)$ species dominating in the vector 'Fast space' in the modified ILDM satisfy tolerance criterion (4). It justifies the QSS approximation for corresponding metabolites. The fast reactions could be discovered from the analysis of product between stoichiometric and transformation matrices in 'classical ILDM'. The reactions which dominate the fast modes and which are less important for the slow modes could be regarded as QE.

3 EXAMPLES

3.1 Michaelis–Menten kinetics

We start our discussion with the simplest enzymatic reaction mechanism: Michaelis–Menten irreversible kinetics, described by

$$\underset{S}{\text{Substrate}} + \underset{E}{\text{Enzyme}} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \underset{C}{\text{Complex}} \xrightarrow{k_2} \underset{P}{\text{Product}} + \underset{E}{\text{Enzyme}}$$

Palsson and Lightfoot (1984) explored the properties of the model via improved scaling and linearization about a characteristic steady state. Here, we use their analytical approach for testing our time-scale separation methods. A slightly different scaling of model variables is used:

$$u=\frac{S}{S_0},\quad v=\frac{C}{E_0},\quad e=\frac{E}{E_0},\quad \tau=k_1E_0t,$$

where E_0 and S_0 are the corresponding initial concentrations. Substitution of these scaled variables into the differential equations leads to the following scaled equations:

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = -u(\tau)+\left(u(\tau)+\frac{Q}{1+S_t}\right)v(\tau)$$
$$\epsilon\frac{\mathrm{d}v}{d\tau} = u(\tau)-\big(u(\tau)+Q\big)v(\tau)$$

Table 1. Michael–Menten model: $\epsilon\to 0$

	Classical ILDM		Modified method	
TS	C	S	C	S
4.2626	39.68	60.32	0.23	99.77
0.0034	99.77	0.23	99.77	0.23

The table displays metabolite contribution to modes derived at time $t=0.25$ by ILDM analysis. The model parameters are $S_0=100,\ E_0=1,\ k_1=1,\ k_{-1}=k_2=100$.

with three dimensionless parameters:

$$\epsilon=\frac{E_0}{S_0},\quad S_t=\frac{k_2}{k_{-1}},\quad Q=\frac{K_m}{S_0}=\frac{k_{-1}+k_2}{k_1S_0}$$

Our methods are not restricted to the local analysis near the steady state, so we can investigate the dynamical interaction between species in the modes additionally to the analytical results of Palsson and Lightfoot (1984).

3.1.1 ILDM analysis in COPASI Following the approach of Palsson and Lightfoot, we consider five possible combinations of dimensionless parameters. In all considered cases, the clear time-scale separation occurs.

(i) $\epsilon\to 0$: this is the classical situation of Michaelis–Menten kinetics. In this limit, the motion on the fast time scale is dominated by the complex C decoupled from the substrate S. On the slow part, the changes of substrate and complex are balanced (Palsson and Lightfoot, 1984). Here, the QSS assumption is valid.

Our time-dependent classical ILDM leads to the distinction between slow and fast modes (for chosen $tol=10^{-3}$) at times $t>0.001$. The analysis of the transformation matrix $\mathbf{T}^{-1}$ shows that the complex C dominates the fast mode. The contributions of both metabolites to the slow mode are comparable and so, the changes of substrate and complex are balanced [see Table 1, which displays the matrix 'Modes(i,j)' at time point $t=0.25$].

The analysis via the modified method leads to the classification of the complex C as a fast metabolite (however, for time $t>0.017$ first). Thus, the QSSA for metabolites C is justified in this case. The last method does not indicate the interaction between metabolites in the slow mode (compare results in Table 1).

(ii) $Q\to 0$ (Fig. 1): this limit is very interesting in order to consider the dynamical interaction between metabolites in slow and fast modes. For our analysis in COPASI, we choose $\epsilon=0.5,\ S_t=1,\ Q=0.05$ and tolerance is $tol=10^{-2}$.

No reduction is possible by modified ILDM method until $t=2.6$. For $t>2.6$, the substrate is distinguished as fast, its contribution to the fast space is >95%. The analysis of the classical ILDM shows that after the short transient phase ($t<0.025$) the complex C dominates the fast space. This dominance decreases with time and for $t\approx 1.13$, both substrate and complex are balanced also in the fast mode (their contributions to the fast space are equal). Then the contribution of the substrate increases and finally, the substrate dominates the fast space near the equilibrium. Contributions of both

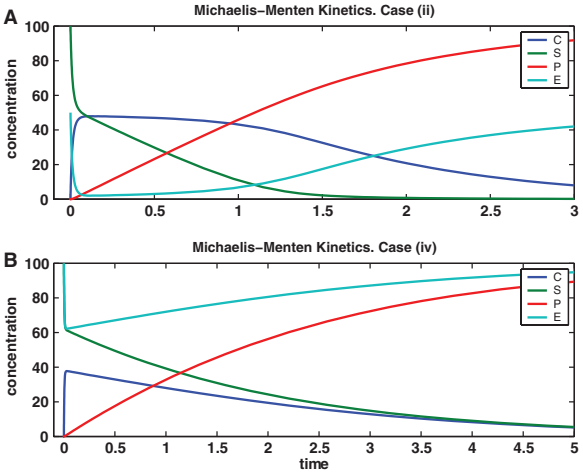

Fig. 1. Simulation of Michaelis–Menten model in Matlab. (**A**) $\mathbf{Q} \to 0$ and (**B**) $\mathbf{S_t} \to \mathbf{0}$.

Table 2. Michaelis–Menten model: $\mathbf{Q} \to 0$

t = 0.15			t = 2.99		
TS	C	S	TS	C	S
23.87	49.49	50.51	1.03	49.41	50.59
0.02	95.81	4.19	0.02	2.84	97.16

The table represents the matrix of metabolite contribution to modes by classical ILDM. The calculations are performed for $S_0 = 100$, $E_0 = 50$, $k_1 = k_{-1} = k_2 = 1$ and $tol = 10^{-2}$.

variables to the slow mode are approximately equal for all times (Table 2).

(iii) $Q \to \infty$: here, we have a situation similar to the limit (i). Namely, the motion on the fast time-scale is dominated by the complex C decoupled from the substrate. On the slow time-scale, the changes of both variables are balanced. Hence the classical QSS approximation is possible.
Applying the modified method also leads to the classification of the complex C as fast after some transient phase.

(iv) $S_t \to 0$ (Fig. 1): this limit means that an equilibrium between the enzyme E, the substrate S and the enzyme–substrate complex C is established quickly. Originally, this QE assumption was proposed by Michaelis and Menten (1913) for deriving their classical equation:

$$\frac{dP}{dt} = \frac{k_2 E_0 S}{K_s + S}, \qquad K_s = \frac{k_{-1} + k_2}{k_1}.$$

They assume that the reaction $S + E \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} C$ is always practical in equilibrium, and K_s is its equilibrium constant, i.e. k_2 is negligible in comparison with k_{-1}.

The classical ILDM analysis leads to the occurrence of two independent dynamical modes (cf. Palsson and Lightfoot, 1984). After the short transient phase ($t < 0.006$, when no reduction is possible), the contribution of C to the fast mode is larger than the one of S. Nevertheless, during that time the contribution of both variables becomes equal. Table 3 shows

Table 3. Michaelis–Menten model: $\mathbf{S_t} \to 0$

t= 0.01			t = 10		
TS	C	S	TS	C	S
3.54	49.89	50.11	2.01	49.87	50.13
0.004	71.72	28.28	0.005	50.35	49.65

Metabolite contribution to modes by classical ILDM. Two separate modes occur which however could not be related to one distinct metabolite. The calculations are performed for $S_0 = 100$, $E_0 = 100$, $k_1 = 1$, $k_{-1} = 100$, $k_2 = 1$ and $tol = 10^{-3}$.

the contribution of each metabolites to the slow and fast modes simulated in COPASI for two time steps $t = 0.01$ and $t = 10$. Our modified method does not lead to any reduction for single metabolites up to time $t = 5.4$ (at least for $tol < 0.01$).

Moreover, according to Delgado and Liao (1992), the criterion for identifying the fast reactions of the system is the existence of a partition of the modal matrix $\mathbf{T}^{-1} = [\mathbf{T}_s^{-1}, \mathbf{T}_f^{-1}]$ such that $\mathbf{T}_s^{-1} \cdot \mathbf{N_R} \approx 0$, where $\mathbf{N_R}$ corresponds to fast reaction terms in the stoichiometric matrix. At time $t = 0.01$ the simulations in COPASI lead to

$$\mathbf{T}_s^{-1} \cdot \mathbf{R_1} = 49.89 - 50.11 = -0.22, \quad \mathbf{T}_s^{-1} \cdot \mathbf{R_2} = -49.89,$$

where stoichiometric vector $\mathbf{R_1} = [1 \ -1]$ corresponds to the first reaction $S + E \underset{k_{-1}}{\overset{k_1}{\Longleftrightarrow}} C$ and $\mathbf{R_2} = [-1 \ 0]$ to the product formation.

Thus for the case $S_t \to 0$, the QSSA for complex C is incorrect. Nevertheless, the QEA leads to similar equations as for 'standard' Michaelis–Menten system (compare Palsson and Lightfoot, 1984).

(**v**) $\mathbf{S_t} \to \infty$. In this case, no rational reduction is possible by means of both methods.

3.2 Calcium oscillations

The second example discussed here is the 'one-pool model' of Ca^{2+} oscillations analyzed by Schuster and Mahrl (2001). The model includes six processes: an influx of Ca^{2+} into the cell ($v_1 = \text{const}$), the efflux of Ca^{2+} out of the cell (rate v_2), the pumping of Ca^{2+} from the cytosol into an intracellular calcium store (rate v_3), the efflux out of the store through channels (v_4), and the leak out of the store (rate v_5):

$$v_1 = k_1, \ v_2 = k_2 S_1, \ v_3 = k_3 S_1, \ v_4 = \frac{k_4 S_2 S_1^n}{K^n + S_1^n}, \ v_5 = k_5 S_2,$$

where S_1 and S_2 denote the Ca^{2+} concentrations in the cytosol and in the Ca^{2+} store, respectively. The corresponding equations are:

$$\frac{dS_1}{dt} = v_1 - v_2 - v_3 + v_4 + v_5,$$
$$\frac{dS_2}{dt} = v_3 - v_4 - v_5.$$

The system exhibits typical spiking oscillations with phases when the free cytosolic calcium concentration remains nearly constant (numerical simulations in Matlab are shown in Fig. 2 for $k_1 = 0.323$, $k_2 = 1$, $k_3 = 2$, $k_4 = 1$, $n = 4$, $K = 1$, $k_5 = 0.01$).

Schuster and Mahrl (2001) showed analytically that the oscillations involve fast and slow processes and that the separation

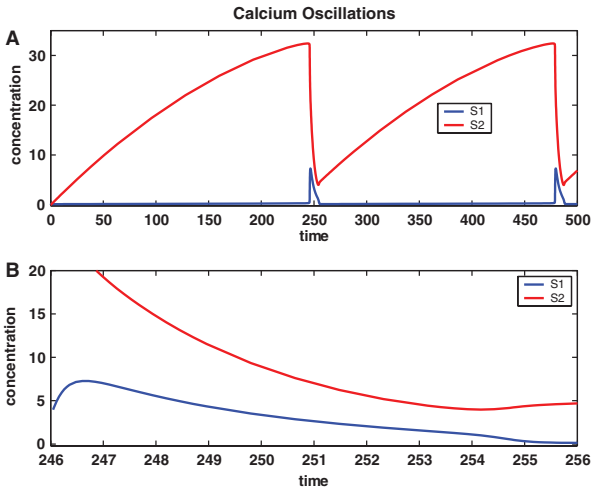

Fig. 2. Numerical simulation of calcium oscillations in Matlab. (**B**) shows an enlargement of (**A**) close to $t = 250$

Table 4. Calcium oscillations: classical ILDM

$t = 2$			$t = 234$		
TS	S_1	S_2	TS	S_1	S_2
296.4	40.02	59.98	22.39	64.99	35.01
0.33	99.66	0.34	2.86	95.53	4.47

Metabolite contribution to modes. The parameters of calculation are given in the text.

of time scales can be observed. For cytosolic Ca^+ concentrations which are below the half-saturation constant K ($S_1 < K$), the QSSA for S_1 could be justified. For the regions $S_1 > K$ and $S_1 \approx K$, the QSSA is not as well-suited as for low S_1.
Here, we analyze this model by means of our methods.

3.2.1 Time-dependent classical ILDM During the initial phase (Table 4), the time scales are very well separated (with the factor $\tau_1/\tau_2 \approx 0.0001$ between eigenvalues of Jacobian for $t = 2$). The cytosolic Ca^+ concentration S_1 dominates the fast mode. The contribution of both variables to the slow mode are balanced and are approximately equal. During the time course, the separation of time scales decreases and also the dominance of S_2 in the slow mode increases. For $t = 216$, both variables S_1 and S_2 are very well separated in fast and slow modes. Moreover, the cytosolic Ca^+ concentration S_1 dominates the fast mode, the store Ca^+ concentration S_2 dominates the slow one.

Then the influence of S_1 on the slow mode increases and, the separation of time scale keeps diminishing (Table 4). For $t \in (235, 245)$, no reduction is possible. In this period , the eigenvalues of the Jacobian become purely imaginary first and then, they even have positive real parts.

The reduction is possible again for $t \in (246, 253)$. According to the lower graph in figure 2, both variables are very well balanced in the modes. For $t = 249$, for instance, the contribution of each metabolite to the slow mode (with time scale 3.67 s) is 42.17% for S_1 and 57.83% for S_2. There is a slight dominance of cytosolic Ca^+ concentration in the fast mode (with TS = 0.27 s). Its contribution is ~73%.

No reduction is possible for $t \in (253, 255)$. For $t > 256$, the system repeats the behavior of t between 0 and 256 periodically.

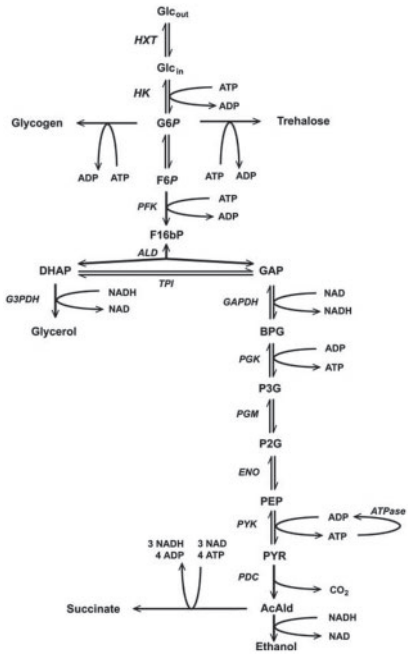

Fig. 3. Glycolytic pathway in yeast.

3.2.2 Modified time-dependent ILDM Again, there is a clear time-scale separation for time $t < 234$. The cytosolic Ca^+ concentration S_1 dominates the fast mode and can be described as a fast variable. According to our method, no reduction is possible between $t = 234$ and $t = 255$. For time interval $t \in (234, 245)$, the concentration of S_1 is between 0.27 and 0.36, which increases then very quickly between $t = 245$ and $t = 246.5$ with maximum $S_1 = 7.3$ and then falls down to the 'near stationary' values $S_1 = 0.2$ at $t = 255$.

In summary, the analytical predictions of Schuster and Mahrl (2001) correspond to the results of our modified time-dependent ILDM method: if the cytosolic Ca^+ concentration is smaller than K ($S_1 > 0.2 < K = 1$), then a QSSA for S_1 is appropriate. For the values $S_1 > K$ and $S_1 \approx K$, the QSSA could not be justified. The classical ILDM implemented in COPASI, however, derives additional information about the system behavior in the 'peak' state (at times between 246 and 253).

3.3 Yeast glycolysis

Our last example is a model of branched yeast glycolysis developed by Teusink *et al.* (2000) and modified by Pritchard and Kell (2002). A reaction system is illustrated schematically in Figure 3. We will not discuss this model in detail and the interested reader is directed to the original reference (Pritchard and Kell, 2002).

The differential equations describing the model as well as the kinetic parameters used are given in Supplementary Material. The set of ODEs consists of 17 equations. From the reaction stoichiometries, two moiety conservations were

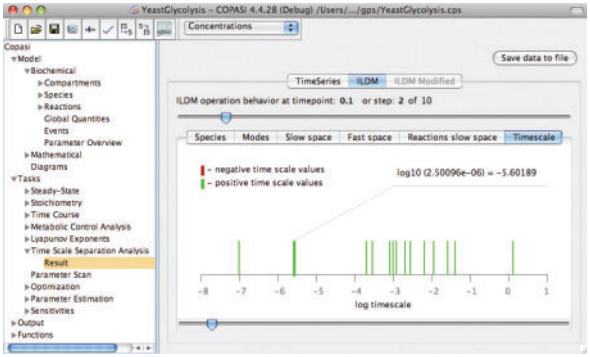

Fig. 4. The distribution of the eigenvalues for $t=0.1$

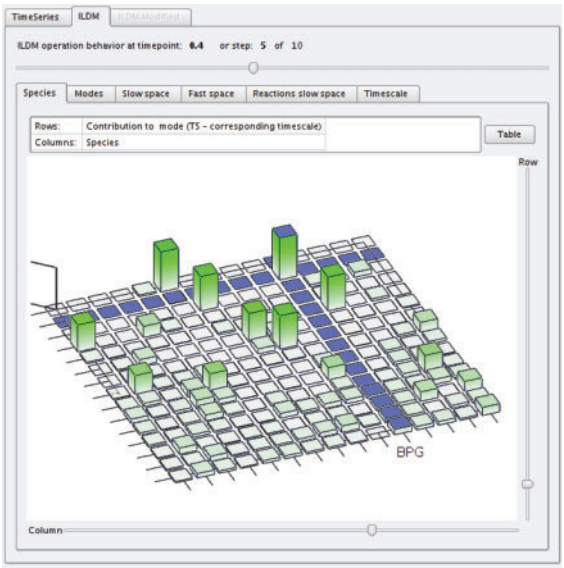

Fig. 5. Metabolite contribution to modes in yeast glycolysis model.

derived: ATP+ADP+AMP=const, NAD+NADH=const. The model shows a stable steady state.

Our time-scale separation methods are not restricted to the steady state. We analyze 15 differential equations for independent metabolites. The Jacobian at all investigated points of time has 11 purely real and 4 complex eigenvalues; the real parts of all eigenvalues are negative. We have two clearly separated fast time scales (of the order 10^{-6} and 10^{-8}). The slowest mode is clearly separated as well. The remaining 10 time scales are distributed between 10^{-4} and 10^{-1}. The distribution of the eigenvalues is graphically represented in Figure 4.

3.3.1 Classical time-dependent ILDM We have chosen $tol=10^{-6}$. At the starting point $t=0$, the time-scale analysis leads to 10 slow modes, this number reduces with time. For $t>1$, only one mode with time-scale 0.14 is active, the system reaches the steady state. The diagram in Figure 5 shows the metabolite distribution in single modes. The fastest mode is dominated by GAP (with 95.3%) and DHAP (with 4.3%), indicating the rapid conversion of GAP and the fact that the TPI reaction is a near equilibrium reaction.

Table 5. Product of $\mathbf{T}_s^{-1}$ by $\mathbf{N}$ for model of Yeast glycolysis

	HXT	HK	PGI	ALD	TPI	GAPDH	PGK	PGM	ENO	PYK
t=0	5.71	−0.19	0.066	8.59	−0.002	−0.975	−5.33	0.24	8.136	−7.305
t=0.3	6.11	−5.8	−0.95	5.76	−0.001	−2.77	1.19	−0.22	5.07	2.63
t=0.5	7.78	−5.99	−0.367	−0.37	9.9e-05	2.58	2.23	0.036	0.164	2.458
t=0.9	11.38	−6.45	−0.46	2.9	0.0001	0.87	2.61	−0.11	−0.6	2.8

The table indicates that TPI as well as PGI and PGM reactions do not contribute to the slow space significantly. Thus, the QEA for these reactions is valid.

The second mode is dominated by the metabolite BPG (with >99%). Again, and not unexpectedly, one can conclude that this metabolite is transformed very rapidly during glycolysis. The same holds true for P2G dominating the third fastest mode. Interestingly, P3G has a much more distributed contribution to the different time scales. This indicates that even though P3G is produced and consumed by two fast reactions, the equilibrium constant of PGM which favors the substrate side, causes P3G not to rapidly convert in all cases.

In order to detect the QE reactions, we analyze also in COPASI the product $\mathbf{T}_s^{-1}\cdot\mathbf{N_R}$, where we have chosen as $\mathbf{T}_s^{-1}$ the vector 'slow space' (the contribution of metabolites to the set of all slow modes). The corresponding analysis of the most important reversible reactions shown in Table 5 confirms that TPI is the fastest reaction at all considered time points. This rapid equilibration leads also to a straight line in the phase plane between GAP and DHAP. The trajectory has a positive slope which corresponds to the equilibrium constant of the TPI reaction: $\mathrm{DHAP}=\mathrm{GAP}\cdot k_{\mathrm{eq}}^{\mathrm{TPI}}$.

The comparison with direct Matlab simulation of initial ODE and reduced ODE corresponding to QEA for TPI reaction leads to the relative error of $\sim 10^{-5}$ for whole time interval $t\in(0,1)$.

The next two candidates for QE are the PGI and PGM reactions. The corresponding relative error between full system and reduced systems simulated with Matlab is $\sim 10^{-2}$.

3.3.2 Modified time-dependent ILDM The number of slow and fast metabolites depends clearly on the user-defined tolerance. For instance, for $tol=10^{-6}$ at $t=0.1$, we have only two fast metabolites (GAP and BPG). For $tol=10^{-4}$, four metabolites (GAP, BPG, NAD and P2G) are fast.

We compare shortly the results of our time-scale analysis with the direct simulation of the corresponding DAE system with Matlab. Figure 6 shows the relative error between the solution of the initial ODE system and the corresponding DAE system with consistent initial values for chosen steady-state metabolites. The bars on the graph represent the magnitude of error for slow metabolites. The time $t=0.1$ serves as initial point of analysis. In the case of two 'fast' metabolites (GAP and BPG) at $t=0.1001$ (at distance $0.001\gg\tau_{\mathrm{fast}}\approx 10^{-6}$), the maximal relative error (2×10^{-4}) corresponds to the metabolite DHAP, which has the second important contribution to the fastest mode. After a longer time interval ($t=0.2$), the relative error of two steady-state metabolites is of the order 10^{-3} for most of the slow species (Fig. 6A).

Assuming four steady-state metabolites (Fig. 6B), the time scale of the fastest slow mode is about 3×10^{-4}. The maximum error at $t=0.11$ is $\sim 5\times 10^{-3}$ and corresponds to PEP and P3G. These are immediate neighbors of QSS metabolite P2G in the reaction schema.

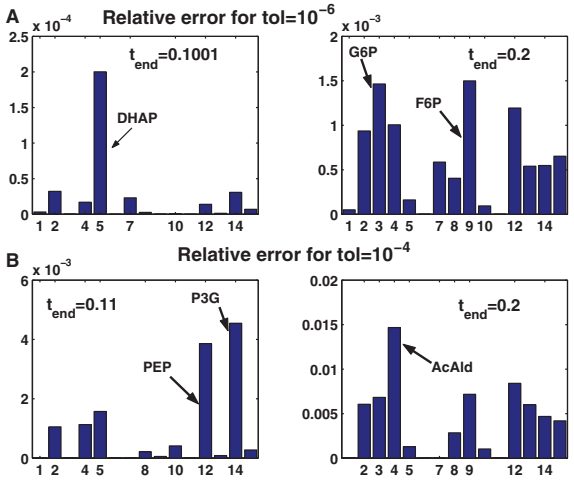

Fig. 6. Comparison of solutions to original ODEs and approximative DAEs. The bars represent the magnitude of the relative error of slow metabolites at time $t = t_{end}$. (**A**). For chosen $tol = 10^{-6}$ the modified ILDM justifies QSSA for GAP and BPG. (**B**). For $tol = 10^{-4}$ four metabolites GAP, BPG, NAD and P2G are fast.

After longer time periods most of the slow metabolites have errors of the order 10^{-2}.

4 DISCUSSION AND CONCLUSIONS

In this article, we present the implementation of two complexity reduction methods for biochemical systems. The algorithms provide an automated dynamical network decomposition with respect to time-scale decoupling in non-linear reaction systems. Both methods work independently of assumptions about the specific dynamic regime. Since the local ILDM analysis can be performed at every user-chosen point of the trajectory, the time-dependent reduction is achieved.

For complex biochemical systems, the concept of the classical, time-dependent variant of the ILDM method allows in principle the dimension reduction of the model by a systematic decomposition of the dynamics in the different time scales. However, for the case of metabolic networks, the fast dynamics cannot always be associated with specific state variables (specific metabolites) of the model, in general. Most of the species participate in multiple reactions, which may evolve in vastly different time scales. The classical ILDM 'modal' analysis implemented in COPASI gives a rationalization of QE assumption [as shown in the case (iv) of Michaelis–Menten reaction system and in the case of TPI, PGI and PGM reactions in yeast glycolysis], as well as a more accurate QSSA interpretation, which should be related to reduced dynamic interactions between concentration variables rather than classical heuristic notions stating that production and destruction rates of the intermediate metabolites are approximately equal. In this sense, we confirmed the results of Palsson and Lightfoot (1984) analysing Michaelis–Menten systems.

Moreover, the lack of explicit separation of fast and slow species concentrations stimulated us to develop a modified method with the main goal of distinguishing clearly between fast and slow metabolites (rather than modes).

In summary, the complexity reduction methods introduced here:

(1) identify the fast species in the biochemical network;

(2) clarify the dynamical interaction between metabolites in slow and fast spaces;

(3) support the user in detecting QE reactions.

ACKNOWLEDGEMENTS

We would like to thank our collaborators Pedro Mendes, Stefan Hoops (both on COPASI) and Jurgen Zobeley, Dirk Lebiedz and Julia Kammerer (development of the original ILDM method).

Funding: Klaus Tschira Foundation; BMBF HepatoSys initiative.

Conflict of Interest: none declared.

REFERENCES

Delgado,J. and Liao,J.C. (1992) Control of metabolic pathways by time-scale separation. *Biosystems*, **36**, 55–70.

Deuflhard,P. and Heroth,J. (1996) Dynamic dimension reduction in ODE models. In Keil,F. *et al*. (eds) *Scientific Computing in Chemical Engineering*. Springer, Heidelberg, pp. 29–43.

Golub,G.H. and van Loan,C.F. (1996) *Matrix Computations*. Johns Hopkins University Press, Baltimore.

Goussis,D.A. and Najm,H.N. (2006) Model reduction and physical understanding of slowly oscillating processes: the circadian cycle. *Multiscale Model. Simul.*, **5**, 1297–1332.

Holme,P. *et al*. (2003) Subnetwork hierarchies of biochemical pathways. *Bioinformatics*. **19**, 532–538.

Hoops,S. *et al*. (2006) COPASI - a complex pathway simulator. *Bioinformatics*, **22**, 3067–3074.

Kauffman,K.J. *et al*. (2002) Description and analysis of metabolic connectivity and dynamics in the human red blood cell. *Biophys. J.*, **83**, 646–662.

Lam,S.H. and Goussis,D.A. (1994) The CSP method for simplifying kinetics. *Int. J. Chem. Kinetics*, **26**, 461–486.

Lebiedz,D. *et al*. (2008) Automatic complexity analysis and model reduction of nonlinear biochemical systems. In *Proceedings of CMSB 2008*, Vol. 5307 of *LNBI*, Springer, Heidelberg, pp. 123–140.

Lovrics,A. *et al*. (2006) Time scale and dimension analysis of budding yeast cell cycle model. *BMC Bioinformatics*, **7**,494.

Maas,U. and Pope,S.B. (1992) Simplifying chemical reaction kinetics: intrinsic low-dimensional manifolds in composition space. *Combust. Flame*, **88**, 239–264.

Michaelis,L. and Menten,M.L. (1913) Die Kinetik der Invertinwirkung. *Biochem. Z.*, **49**, 333–369.

Palsson,B.O. and Lightfoot,E.N. (1984) Mathematical modeling of dynamics and control in metabolic networks I. On Michaelis Menten kinetics. *J. Theor. Biol.*, **111**, 273–302.

Price,N.D. *et al*. (2003) Analysis of metabolic capabilities using singular value decomposition of extreme pathway matrices. *Biophys. J.*, **84**, 794–804.

Pritchard,L. and Kell,D.B. (2002) Schemes of flux control in a model of Saccharomyces cerevisiae glycolysis. *Eur. J. Biochem.*, **269**, 3894–3904.

Shaik,O.S. *et al*. (2005) Derivation of a quantitative minimal model for the photosensitive Belousov-Zhabotinsky reaction from a detailed elementary-step mechanism. *J. Chem. Phys.*, **123**, 324103.

Schmidt,D. *et al*. (1998) Intrinsic low-dimensional manifolds of strained and unstrained flames. *Combustion Theory and Modelling*, **2**, 135–152.

Schuster,S. and Mahrl,M. (2001) Bifurcation analysis of calcium oscillations: time-scale separation, canards, and frequency lowering. *J. Biol. Syst.*, **9**, 291–314.

Schuster,S. *et al*. (2002) Exploring the pathway structure of metabolism: decomposition into subnetworks and application to Mycoplasma pneumoniae. *Bioinformatics*, **18**, 351–361.

Surovtsova,I. *et al*. (2006) Approaches to complexity reduction in a System Biology Research Environment (SYCAMORE). In *Winter Simulation Conference 2006*, Omnipress, pp. 1683–1989.

Teusink,B. *et al*. (2000) Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.*, **267**, 5313–5329.

Zobeley,J. *et al*. (2005) A new time-dependent complexity reduction method for biochemical systems. In Prami,C. *et al*. (eds) *Transactions on Computational Systems Biology*, vol. 3380 of *Lecture Notes in Computer Science*. Springer, Heidelberg, pp. 90–110.