Journal of Mathematical Chemistry, Vol. 43, No. 4, May 2008 (© 2007)
DOI: 10.1007/s10910-007-9248-4

Quasi steady-state approximations in complex intracellular signal transduction networks – a word of caution

Morten Gram Pedersen*

Department of Physics, Technical University of Denmark, Fysikvej 309, DK-2800 Kgs. Lyngby, Denmark
E-mail: gram@fysik.dtu.dk

Alberto Maria Bersani

Department of Mathematical Methods and Models, "La Sapienza" University, I-00161 Rome, Italy

Enrico Bersani

ISMAC Institute, CNR, I-16149 Genova, Italy

Received 7, December 2006; Revised 24 January 2007

Enzyme reactions play a pivotal role in intracellular signal transduction. Many enzymes are known to possess Michaelis–Menten (MM) kinetics and the MM approximation is often used when modeling enzyme reactions. However, it is known that the MM approximation is only valid at low enzyme concentrations, a condition not fulfilled in many *in vivo* situations. Recently the total quasi steady-state approximation (tQSSA) has been developed for enzymes with MM kinetics. This new approximation is valid not only whenever the MM approximation is, but moreover in a greatly extended parameter range. Starting from a single reaction and arriving at the mitogen activated protein kinase (MAPK) cascade, we give several examples of biologically realistic scenarios where the MM approximation leads to quantitatively as well as qualitatively wrong conclusions, and show that the tQSSA improves the accuracy of the simulations greatly.

KEY WORDS: Michaelis–Menten kinetics, quasi steady-state assumption, enzyme signaling networks, double phosphorylation, MAPK cascade

AMS MSC codes: 92C40, 92C45

1. Introduction

Every living cell responds to external stimuli, like hormones, ions, heat shock, etc., which are transduced by a complex intracellular molecular network. When an external ligand binds a plasma membrane receptor, intracellular second

*Corresponding author.

0259-9791/08/0005-1318/0 © 2007 Springer Science+Business Media, LLC

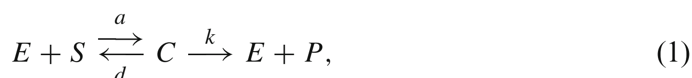
messengers interacting with membrane receptors are activated, and by means of biochemical reactions transduce the signal.

In the last decade many mathematical models have been formulated to investigate the behavior of complex intracellular biochemical networks. Many of those are based on the well-studied mitogen activated protein kinase (MAPK) cascade (see for example [1–5]), and although not crucial for the results presented here, this ubiquitous signaling pathway will be given special attention in the following.

The aim of such modeling (which is an integral part of the 'Systems Biology' large scale project) is roughly twofold: to reproduce and study some particular phenomena observed experimentally (like bistability, oscillations, ultrasensitivity, hysteresis, etc.) and to investigate the properties of these networks as information processing and transducing devices. As a hope for the future, this modeling could be used for pharmaceutical scopes (first of all drug discovery) as a reliable tool to make predictions about the effects of drugs on the biochemical networks, thus shortening the preclinical phase. This goal is related to the ambitious project of a "Virtual Cell" ([6], http://www.vcell.org/) or "Silicon Cell" ([7], http://www.siliconcell.net/), which aims at simulating the behavior of whatever cell as closely as possible to the physiological reality: "A silicon cell is a precise replica of (part of) a living cell" (cited from http://www.siliconcell.net/).

Surprisingly, the mathematical formulation of these highly interconnected enzyme reactions is usually based on *in vitro* studies of isolated reactions, without a serious criticism of the delicate passage from the kinetics of simple reactions to the kinetics of a network of reactions shared by several cascades in a crowded molecular environment [8]. This can be justified when analyzing underlying mechanisms (e.g., the importance of feedback or the creation of oscillations), where the exact kinetic expressions and parameters are less important since one is usually only interested in the qualitative behavior that the system can perform. However, in the light of the Silicon Cell project, which aims at being a both qualitative as well as quantitative precise representation of the living cell, the use of correct parameters, kinetic expressions, and initial conditions (i.e., steady-state concentrations of molecular species) becomes crucial. This is the subject of the present work.

One of the principal components of the mathematical approach to Systems Biology is the model of biochemical reactions set forth by Henri in 1901 [9–11] and Michaelis and Menten in 1913 [12], and further investigated by Briggs and Haldane in 1925 [13]. This formulation considers a reaction where a substrate S binds an enzyme E reversibly to form a complex C. The complex can then decay irreversibly to a product P and the enzyme, which is then free to bind another molecule of the substrate. This process is summarized in the scheme



$$E + S \underset{d}{\overset{a}{\rightleftarrows}} C \overset{k}{\longrightarrow} E + P, \quad (1)$$

where a, d, and k are kinetic parameters (supposed constant) associated with the reaction rates.

This scheme is mathematically represented by a system of two nonlinear ordinary differential equations (ODEs), corresponding initial conditions and two conservation laws. The initial conditions give the concentrations of S and C at the beginning of the reaction, and their time development is described by the ODEs, while E and P are linked to S and C through the conservation laws.

Assuming that the complex concentration is approximately constant after a short transient phase leads to the usual Michaelis–Menten (MM) approximation [or *standard quasi steady-state assumption* or *approximation* (standard QSSA, sQSSA)], which is valid when the enzyme concentration is much lower than either the substrate concentration or the Michaelis constant K_M [14, 15]. This condition is usually fulfilled for *in vitro* experiments, but sometimes breaks down *in vivo* [16–18]. We refer to the next section for the mathematical formulation of scheme (1), and to [19] for a nice, general review of the kinetics and approximations of (1).

The advantage of a quasi steady-state approximation is that it reduces the dimension of the system, passing from two equations (*full system*) to one (*MM approximation* or *sQSSA*) and thus speeds up numerical simulations greatly, especially for large networks as found *in vivo*. Moreover, the kinetic constants in (1) are usually not known, whereas finding the kinetic parameters for the MM approximation is a standard *in vitro* procedure in biochemistry [20]. However, to simulate physiologically realistic *in vivo* scenarios, one faces the problem that the MM approximation is no longer valid as mentioned above. Hence, even if the kinetic constants such as K_M are identical *in vivo* and *in vitro*, they need to be implemented in an approximation which is valid for the system under investigation.

Approximations such as the *total QSSA* (tQSSA) [21, 22], which is valid for a broader range of parameters covering both high and low enzyme concentrations, have been introduced recently. Tzafriri [22] showed that the tQSSA is at least roughly valid for any set of parameters in the case of the reaction in (1). Importantly, the tQSSA uses the same parameters (V_{max}, K_M) as the sQSSA. Hence, the parameters found *in vitro* from the MM approach can be used by the tQSSA for modeling *in vivo* scenarios.

The roles of V_{max}, the maximal reaction velocity, and K_M, the Michaelis constant describing the concentration of the substrate at which the reaction rate is half maximal, become essential when characterizing biochemical reactions *in vitro* as well as *in vivo*. Moreover, descriptions of cooperative reactions, inhibition, and many other biochemical processes have exploited the fundamental ideas of the MM scheme, i.e., the sQSSA and the parameters V_{max} and K_M (see, e.g., [20]). However, since these approximations cannot be expected to be valid *in vivo*, employing the tQSSA to these more complex situations would be preferable. Tzafriri and Edelman [23] studied the completely reversible enzyme reaction in

terms of the tQSSA. We have recently derived the tQSSA for fully competitive reactions [24].

In this paper, we compare various approximation schemes of the time concentration development of the chemical species involved in the reactions. This is done by numerically solving the system of ODEs derived from the reaction scheme using the methods and approximations within the various approaches. Our investigation applies to every biochemical network which includes enzyme reaction cascades. We show that the use of the sQSSA can lead to gross quantitative as well as qualitative wrong conclusions even in the case of simple networks. The tQSSA is shown to estimate the behavior significantly better, and therefore we propose to use this approximation when modeling intracellular signaling networks.

2. Mathematical background

Modeling all of the intermediate steps of enzymatic reactions, including binding, dissociation, and release of the product using mass action and conservation laws, leads to a system of ODEs, one for each involved complex and substrate. We refer to this as the full system. For (1) the equations are

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -a(E_\mathrm{T} - C)S + d\,C, \tag{2a}$$

$$\frac{\mathrm{d}C}{\mathrm{d}t} = a(E_\mathrm{T} - C)S - (d+k)C \tag{2b}$$

with the initial conditions

$$S(0) = S_\mathrm{T}, \qquad C(0) = 0, \tag{3}$$

and the conservation laws

$$E + C = E_\mathrm{T}, \qquad S + C + P = S_\mathrm{T}. \tag{4}$$

Here E_T is the total enzyme concentration assumed to be free at time $t = 0$. Also the total substrate concentration, S_T, is free at $t = 0$. This is the mathematical formulation of the so-called MM kinetics [12, 14, 20].

The next, well-known and widely used step is that of the Henri–Michaelis–Menten–Briggs–Haldane approximation [9–15]. It leads to an ODE for each substrate while the complexes are assumed to be in a quasi-steady state (i.e., $\frac{\mathrm{d}C}{\mathrm{d}t} \approx 0$). See, e.g., [20] for a general introduction to this approach. We refer to this as the sQSSA. For (1) it is given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} \approx -\frac{V_\mathrm{max} S}{K_\mathrm{M} + S}, \qquad S(0) = S_\mathrm{T}, \tag{5}$$

$$E(0) = E_\mathrm{T}, \quad V_\mathrm{max} = k\,E_\mathrm{T}, \quad K_\mathrm{M} = \frac{d+k}{a}.$$

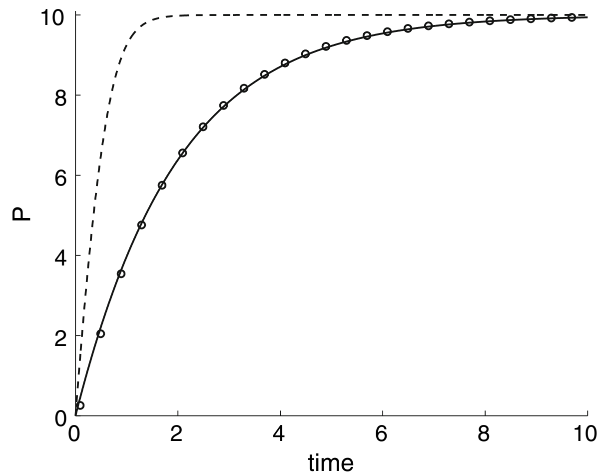


Figure 1. Temporal evolution of the product P at high enzyme concentrations for the single reaction (1). In this case, the solution of the full system [(2), circles] is badly approximated by the MM approximation [sQSSA (5), dashed curve], while the tQSSA [(6), full curve] estimates the behavior very well. Parameters are $k = 0.6$, $K_M = 8$, $E(0) = E_T = 50$, and $S(0) = S_T = 10$, all in arbitrary units.

When we have more than one reaction in the system we denote the Michaelis constant for reaction i by K_i^M, and the reaction constants by a_i, d_i, and k_i.

For the single reaction represented in (1) it has been known for many years that the sQSSA (MM approximation) holds when the initial substrate concentration is much higher than the initial enzyme concentration ($S_T \gg E_T$), but it was later realized that this is not a necessary condition; if the enzyme concentration is much less than the Michaelis constant, then the sQSSA also holds [14, 15]. This is summarized in the validity criterion $E_T \ll S_T + K_M$, which loosely says that the sQSSA (5) holds at low enzyme concentrations (with respect to either the substrate concentration or the K_M value).

As mentioned in the introduction, *in vivo* we cannot in general assume a low enzyme concentration and hence, the MM approximation cannot be expected to hold. A recent approach to resolve this problem is that of the tQSSA. It was introduced by Borghans et al. [21] and refined by Tzafriri [22] for isolated reactions.

This approximation holds for a much larger region of parameter space, and is in fact always roughly valid [22]. Importantly, the tQSSA coincides with the sQSSA when the latter is expected to hold, i.e., at low enzyme concentrations. Figure 1 shows that the tQSSA approximates the full system very well also for high enzyme concentrations where the sQSSA fails.

The tQSSA [21, 22] arises by introducing the total substrate $\bar{S} = S + C$, and assuming that the complex is in a quasi-steady state as for the sQSSA. For (1)

it gives [22]

$$\frac{d\bar{S}}{dt} \approx -k\,C_-(\bar{S}), \quad \bar{S}(0) = S_T, \tag{6}$$

where

$$C_-(\bar{S}) = \frac{(E_T + K_M + \bar{S}) - \sqrt{(E_T + K_M + \bar{S})^2 - 4E_T\bar{S}}}{2}. \tag{7}$$

Numerical integration of (6) easily gives the time behavior of $\bar{S}$, C [by (7)], and S (by the relation $S = \bar{S} - C$). Tzafriri [22] found a criterium for the validity of the tQSSA (6) expressed as a certain $\epsilon \ll 1$, and showed that this is always roughly valid in the sense that

$$\epsilon \leqslant \frac{K}{4K_M} \leqslant \frac{1}{4}. \tag{8}$$

This means that for *any* combination of parameters and initial conditions, (6) is a decent approximation to the full system (2).

As a first order approximation to (6), Tzafriri [22] found the expression, obtained originally in [21] by different techniques,

$$\frac{d\bar{S}}{dt} \approx -\frac{V_{max}\bar{S}}{K_M + E_T + \bar{S}}, \quad \bar{S}(0) = S_T. \tag{9}$$

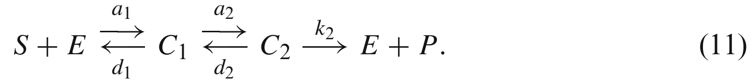
This approximation is valid at low enzyme concentrations $E_T \ll S_T + K_M$, where it reduces to the MM expression (5), but holds moreover at low substrate concentrations $S_T \ll E_T + K_M$ [22]. Thus, with minimal effort, performing the substitutions of S by $\bar{S}$ and of K_M by $K_M + E_T$, one obtains a significantly improved MM-like approximation, without any need of more advanced mathematics. In this first order approximation $C = E_T\bar{S}/(K_M + E_T + \bar{S})$ and then $S = \bar{S} - C$ can be calculated.

Finally, while every reaction is characterized by three constant rates (a, d, k), its QSSA works with only two parameters: V_{max} and K_M. In general, posing $d = \alpha k$, we have

$$k = \frac{V_{max}}{E_T}, \quad d = \frac{\alpha V_{max}}{E_T}, \quad a = \frac{k + d}{K_M} = \frac{(1 + \alpha)V_{max}}{E_T K_M} \tag{10}$$

with one degree of freedom related to the value of α. Consequently, it is possible to vary the triplet (a, d, k) obtaining the same pair (V_{max}, K_M). However, the different choices of (a, d, k) could produce significantly different outputs, and thus predict completely different behavior in the solutions of the full system and of its QSSAs, respectively, as we will show in section 4.

Let us remark that several biochemical and mathematical schemes for the phosphorylation have been suggested in the literature [25–27]. In particular, Salazar and Höfer [27] consider phosphorylation as a double step mechanism, which can be represented by the following scheme:

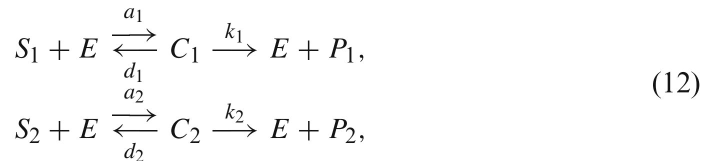


$$S + E \underset{d_1}{\overset{a_1}{\rightleftarrows}} C_1 \underset{d_2}{\overset{a_2}{\rightleftarrows}} C_2 \overset{k_2}{\longrightarrow} E + P. \tag{11}$$

The sQSSA of this reaction was already studied from a mathematical point of view by Frenzen and Maini [28]. The reason for different approaches to phosphorylation is that there is evidence or reaction schemes different from the MM kinetics for this kind of processes [29, 30].

3. Competing substrates and the double phosphorylation mechanism

A theoretically well-studied example of a slightly more complicated network than (1) is the case of fully competitive reactions [14, 24, 31], i.e., reactions with competing substrates, S_1 and S_2,



$$\begin{aligned} S_1 + E &\underset{d_1}{\overset{a_1}{\rightleftarrows}} C_1 \overset{k_1}{\longrightarrow} E + P_1, \\ S_2 + E &\underset{d_2}{\overset{a_2}{\rightleftarrows}} C_2 \overset{k_2}{\longrightarrow} E + P_2, \end{aligned} \tag{12}$$

where S_i, C_i, and P_i represent substrate, enzyme-substrate complex, and product $(i = 1, 2)$ for the two competing reactions. Note that this reaction scheme also covers competitive inhibition (for $k_2 = 0$ with S_2 being the inhibitor).

The system (12) is governed by the coupled ODEs and conservation laws similar to the scenario of a single reaction (1) described above [14, 31, 32]. These equations are given in appendix A. The sQSSA of this system is [14, 32]

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} \approx -\frac{k_i E_{\mathrm{T}} S_i}{K_i^{\mathrm{M}}(1 + S_j/K_j^{\mathrm{M}}) + S_i}, \qquad S_i(0) = S_{i,\mathrm{T}}, \quad i = 1, 2, \quad j \neq i, \tag{13}$$

which basically holds at low enzyme concentrations [31] as in the case of a single reaction, which can be seen as a special case of (12) with negligible inhibitor concentration.

We have recently improved these results [24], extending the region of validity, by applying the tQSSA to the reactions (12). We showed that the tQSSA is given by finding C_1 as the unique biologically acceptable root $(0 < C_1 <$

$\min\{E_T, \bar{S}_1\}$) of the third degree polynomial

$$\begin{aligned}\psi_1(C_1) = &-(K_1^M - K_2^M)C_1^3 \\ &+\left[(E_T + K_1^M + \bar{S}_1)(K_1^M - K_2^M) - (\bar{S}_1 K_2^M + \bar{S}_2 K_1^M)\right] C_1^2 \\ &+\left[-E_T(K_1^M - K_2^M) + (\bar{S}_1 K_2^M + \bar{S}_2 K_1^M) + K_2^M(E_T + K_1^M)\right] \bar{S}_1 C_1 \\ &-E_T K_2^M \bar{S}_1^2, \end{aligned} \tag{14}$$

and similarly finding C_2 as the root in the polynomial ψ_2 obtained by interchanging the indices 1 and 2 in (14). Then the evolution of the system can be studied by means of the tQSSA

$$\frac{d\bar{S}_i}{dt} \approx -k_i C_i(\bar{S}_1, \bar{S}_2), \quad \bar{S}_i(0) = S_{i,T}. \tag{15}$$

This approach extends both the sQSSA for competitive reactions (13) as well as the tQSSA for isolated reactions (6) as shown in [24]. This is confirmed by figure 2(a), which shows that the tQSSA (15) approximates the full system very well. This holds over a wide range of parameters covering both low and high enzyme concentrations [24].

Our results are immediately applicable to, e.g., successive reactions catalyzed by the same enzyme, such as non-processive or distributive double phosphorylation or dephosphorylation processes, as seen for example in the MAPK cascade [25, 26, 33–35]. The reaction scheme for this case can be seen as a special case of (12) with $P_1 = S_2$ and is summarized as

$$S_1 \xrightarrow{E} S_2 \xrightarrow{E} P, \tag{16}$$

where it is usually assumed that at the beginning only S_1 is present. Here S_1 and S_2 compete for the same enzyme, E. For the case of the MAPK cascade one can think of, e.g., diphosphorylated and thus activated MAPKK (MAPKK-PP, here E) phosphorylating MAPK (here S_1) twice, producing first monophosphorylated MAPK (MAPK-P, here S_2) and then double-phosphorylated MAPK (MAPK-PP, here P). See also scheme (17) below. The reason for our results carrying over to this scenario is that they have the same conservation law for the enzyme concentration, $E_T = E + C_1 + C_2$. See appendix A for the equations describing (16). The applicability of the competitive tQSSA to this scenario is confirmed in figure 2(b).

However, it should be remarked that, up to now, there is no theoretical proof of the applicability of the tQSSA for fully competitive enzymes to the case of successive reactions. The problem is that there is no S_2 at time $t = 0$, and hence the timescales cannot be found following [14] because the definition of the transient phase no longer holds.

In the MAPK cascade literature every single reaction is often treated by a MM approximation for an isolated reaction of the form (1), not only without any *a priori* examination of its applicability, but also neglecting the other terms involved in the double reaction and, in particular, the important fact that, for example, MAPK and MAPK-P are competing substrates for MAPKK-P (however, see [25, 26]). This means that even when the sQSSA for (16) holds, the neglect of the competition leads to wrong estimations of the behavior, and can only be expected to be an even greater problem when the sQSSA breaks down, in which case the tQSSA should be used. This situation is illustrated in figure 2(b), which shows that even when both the non-competitive sQSSA and tQSSA as well as the competitive sQSSA fail, the competitive tQSSA is an excellent approximation.

The double phosphorylation as well as double dephosphorylation of MAPK was recently modeled taking into consideration the competition between the pools of MAPK with different phosphorylation states [25, 26]. We model this process by assuming that (16) holds for both the phosphorylation as well as the dephosphorylation processes as in [25]. In [26] both (16) as well as a more complicated process of phosphorylation were considered, but this further step is not of our interest here although applying the tQSSA to this more complicated scheme would be interesting. Similarly, we follow [25] and model the dephosphorylation by (16) instead of the slightly more complicated scheme from [26] for the sake of simplicity.

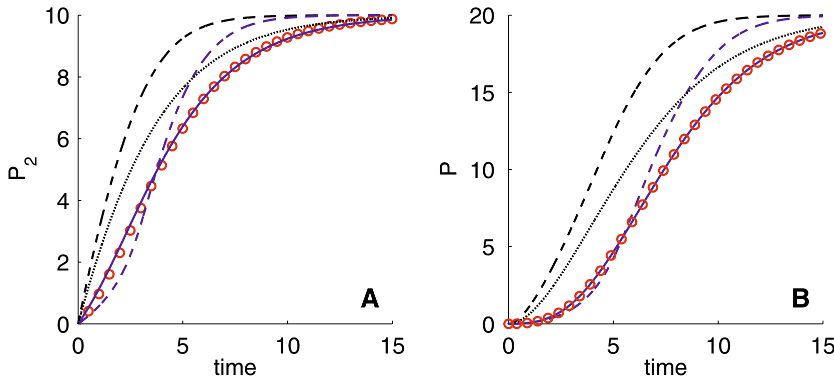
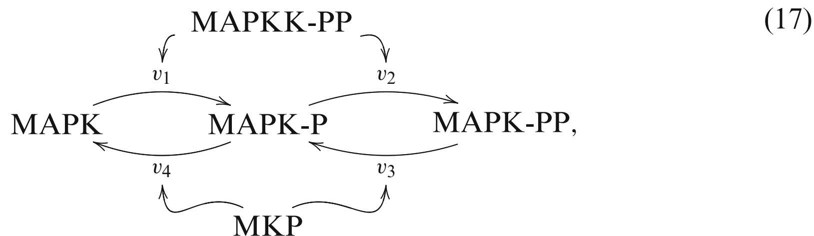


Figure 2. Competitive systems. Panel A: a simulation of competing substrates [scheme (12)]. Panel B: a simulation of two successive reactions catalyzed by the same enzyme (scheme (16)). In both cases the full system [red circles; panel A: (25), panel B: (28)] is estimated very well by the competitive tQSSA [blue, full curve; panel A: (15), panel B: (30)], while the competitive sQSSA [blue, dashed curve; panel A: (13), panel B: (29)] as well as the non-competitive sQSSA [black, dashed curve; (5) for each reaction] and tQSSA [black, dotted curve; (6) for each reaction] do not fit. The parameters are in both panels: $k_1 = 0.5, k_2 = 0.6, K_1^M = 0.75, K_2^M = 8, E_{tot} = 10$, and $S_1(0) = S_2(0) = 10$ in panel A, $S_1(0) = 20, S_2(0) = 0$ in panel B. All units are arbitrary.

Thus, we are studying the scheme



(17)

where the reaction rates $v_1 - v_4$ are assumed to follow MM kinetics with competition between MAPK and MAPK-P for activated MAPKK (MAPKK-PP), and between MAPK-PP and MAPK-P for the generic phosphatase MKP [36].

For the full system and the sQSSA of (17), the equations are basically as above for schemes (12) and (16), see appendix A.

For the tQSSA special care must be used compared to the previous case of (16). The problem is that we still do not have the exact tQSSA for reactions of the given type, because the total substrate corresponding to MAPK-P involves the two complexes of the reactions v_2 and v_4, not both relevant when calculating one of the reaction rates. Therefore, we find a first approximation to each of these complexes and subtract it from the total substrate, see (21)–(24). New tQSSAs should be developed to improve on this *ad hoc* approach.

We use the notation M = [MAPK], M_{p}=[MAPK-P], M_{pp}=[MAPK-PP], E = [MAPKK-PP], F = [MKP] and the concentration of the complex of reaction v_i is denoted by C_i, $i = 1, 2, 3, 4$. The total substrates are defined by

$$\bar{M} = M + C_1, \quad \bar{M}_{\mathrm{p}} = M_{\mathrm{p}} + C_2 + C_4, \quad \bar{M}_{\mathrm{pp}} = M_{\mathrm{pp}} + C_3,$$

and are described by

$$\frac{\mathrm{d}\bar{M}}{\mathrm{d}t} \approx v_4 - v_1, \quad \bar{M}(0) = [\mathrm{MAPK}]_{\mathrm{T}}, \tag{18}$$

$$\frac{\mathrm{d}\bar{M}_{\mathrm{p}}}{\mathrm{d}t} \approx v_1 - v_2 + v_3 - v_4, \quad \bar{M}_{\mathrm{p}}(0) = 0, \tag{19}$$

$$\frac{\mathrm{d}\bar{M}_{\mathrm{pp}}}{\mathrm{d}t} \approx v_2 - v_3, \quad \bar{M}_{\mathrm{pp}} = 0, \tag{20}$$

where the rates are given by

$$v_1 = k_1 C_1(\bar{M}, \bar{M}_{\mathrm{p}} - C_4(\bar{M}_{\mathrm{p}}, \bar{M}_{\mathrm{pp}}\,;F_{\mathrm{T}})\,;E_{\mathrm{T}}), \tag{21}$$

$$v_2 = k_2 C_2(\bar{M}_{\mathrm{p}} - C_4(\bar{M}_{\mathrm{p}}, \bar{M}_{\mathrm{pp}}\,;F_{\mathrm{T}}), \bar{M}\,;E_{\mathrm{T}}), \tag{22}$$

$$v_3 = k_3 C_3(\bar{M}_{\mathrm{pp}}, \bar{M}_{\mathrm{p}} - C_2(\bar{M}_{\mathrm{p}}, \bar{M}\,;E_{\mathrm{T}})\,;F_{\mathrm{T}}), \tag{23}$$

$$v_4 = k_4 C_4(\bar{M}_{\mathrm{p}} - C_2(\bar{M}_{\mathrm{p}}, \bar{M}\,;E_{\mathrm{T}}), \bar{M}_{\mathrm{pp}}\,;F_{\mathrm{T}}). \tag{24}$$

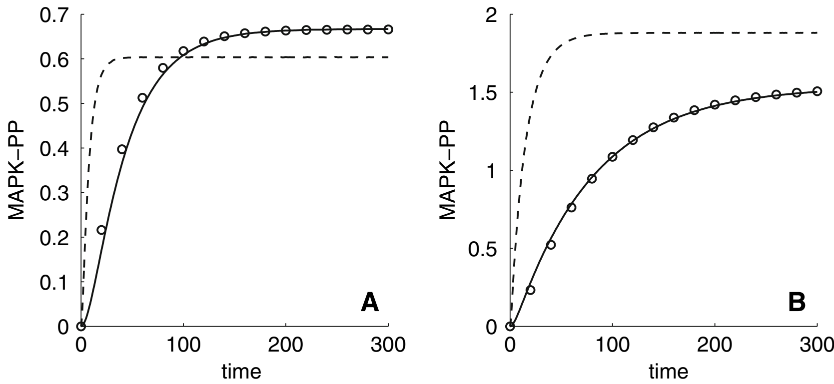


Figure 3. Phosphorylation and dephosphorylation. The sQSSA [(32), dashed line] leads to a wrong estimation of both the transient behavior as well as steady-state levels, while the tQSSA ((18)–(24), full line) fits well, for the double phosphorylation/dephosphorylation (17) modeled with MM kinetics. The full system (31) is shown as circles. Parameters: $[\mathrm{MAPK}]_\mathrm{T}=500$ (panel A), $[\mathrm{MAPK}]_\mathrm{T} = 50$ (panel B), $[\mathrm{MKP}]_\mathrm{T} = 100$, $[\mathrm{MAPKK}]_\mathrm{T} = 50$, $K_1^\mathrm{M} = 50$, $K_2^\mathrm{M} = 500$, $K_3^\mathrm{M} = 22$, $K_4^\mathrm{M} = 18$, $k_1 = 0.01$, $k_2 = 15$, $k_3 = 0.084$, $k_4 = 0.06$. (Concentrations and time in arbitrary units, but for consistence with [26] one can think of nM and seconds).

Here, $C_i(X, Y; Z_\mathrm{T})$ indicates the complex of reaction i with substrate X, competing substrate Y and enzyme Z, and is found as indicated for scheme (12). Specifically, in (14) the following substitutions should be made, and the root then found: C_1 by C_i, $\bar{S}_1$ by X, $\bar{S}_2$ by Y, E_T by Z_T, K_1^M by K_i^M, and K_2^M by K_j^M, where j indicates the competing reaction, see scheme (17).

The parameters are taken from [26, figure 1] and are given in the caption of figure 3. They are given in arbitrary units, but for consistency with [26] one can think of nM for all concentrations, and seconds for all time units. Other parameters should then be expressed in these units.

As seen in figure 3(a), in this case the competitive sQSSA underestimates the duration of the transient phase before reaching the steady state. Furthermore, it underestimates the steady state level of MAPK-PP. However, this underestimation is not a feature of the sQSSA, since lowering the total MAPK concentration to $[\mathrm{MAPK}]_\mathrm{T} = 50$ results in an (even more pronounced) overestimation of the steady state level (figure 3(b)), which of course can be of equal importance as an underestimation. Notably, the tQSSA (18)–(24) fits both the dynamic behavior as well as steady-state levels very well in both cases. Remark the counter-intuitive result that a ten times *lower* total MAPK concentration in figure 3(b) yields a more than two times *higher* level of activated MAPK, showing the strength and utility of mathematical modeling. To illustrate the importance of a reliable estimation of the MAPK levels, we remark that is has been shown experimentally that the dynamics of MAPK activity is crucial for the fate of the cell [37–40]. For example, PC12 cells proliferate in response to transient MAPK activation,

while they differentiate when the activated MAPK levels are sustained [41]. We follow this up in the next section and in the discussion.

4. The MAPK cascade

In the MAPK pathway (figure 4), the upstream kinase (denoted MKKK, i.e., MAP kinase kinase kinase; for example Raf), when activated, phosphorylates the immediately downstream target, which is also a kinase (MAPKK, i.e., MAP kinase kinase, for example MEK) successively on two specific sites, eventually activating it. This last double-phosphorylated kinase (MAPKK-PP) acts on the MAPK (for example ERK) through specific phosphorylation events on two distinct sites. The activated MAPK is then responsible for further downstream signalling. The activated cascade is shut down by the reverse action of specific phosphatases [36, 42], whose outcome is the time modulation of the signal, probably through the regulation of the active kinase (for example, transient versus sustained activation). Moreover, the phosphatase controls the steady state level of activated MAPK, which, in turn, controls downstream processes as mentioned in the previous section. Looking at the complete MAPK cascade, shown in figure 4, it is clear that all the problems arising in the simpler cases described in the previous sections may occur.

Recently, oscillatory phenomena have been investigated theoretically for signal transduction networks like MAPK cascade [3]. Several authors suppose that MAPK-PP acts, by means of a feedback mechanism, on the first layer of the MAPK cascade, and in some cases this feedback has been shown experimentally, for example in NIH3T3 cells, where Raf-1, a MKKK, was found to be

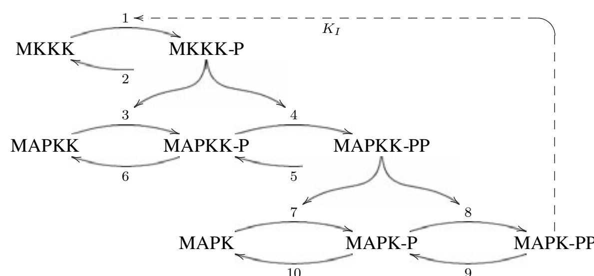


Figure 4. The MAPK cascade. The diagram is based on [3]. Each of the reactions is assumed to follow MM kinetics, but there are competitive reactions since MKKK-P catalyzes both reactions 3 and 4 and MAPKK-PP catalyzes both reactions 7 and 8. Similarly, reactions 5 and 6 are assumed to compete for a phosphatase, and both reactions 9 and 10 to be catalyzed by another phosphastase (MKP in (17)). The phosphatases are not shown for clarity of the figure. The dashed line indicates inhibition of reaction 1 by MAPK-PP as in [3]. However, we assume that this inhibition is competitive. M-P and M-PP represent, respectively, monophosphorylated and diphosphorylated M, where M is either MKKK, MAPKK or MAPK.

Table 1
Regions of oscillations for the MAPK cascade with feedback expressed by the inhibition constant K_I.

Method	Oscillations
Full system	$K_I < 0.18$
Competitive tQSSA	$K_I < 0.86$
Competitive sQSSA	$K_I < 3.02$

inactivated by ERK, a MAPK [43]. See also [39, 44] for reviews. Kholodenko [3] introduced a non-competitive inhibition of this kind. The mathematical model of this complex network (with or without feedback) was built using the MM approximation without considering substrate competition, and it was shown that oscillations could occur for several parameter values.

However, the appearance of oscillations could depend on the way in which the model has been formulated. We compare the network with the full system of reactions to both the competitive sQSSA [25, 26] and the competitive tQSSA [24]. In contrast to Kholodenko [3], we model the negative feedback as a competitive inhibition with inhibition constant K_I to allow the use of the competitive tQSSA (see figure 4 and appendix A). The differential equations describing the full system and the competitive sQSSA are found as for the reactions described in the previous sections, and the equations for the tQSSA are found as for scheme (17). Since MKKK-P and MAPKK-PP have the function of products, substrates, and enzymes, we have non-constant enzyme concentrations. Moreover, special care must be used in the tQSSA, since the total MKKK-P, MAPKK-PP, and MAPK-PP concentrations are not easily defined, in addition to the problems for MAPKK-P and MAPK-P mentioned above for scheme (17). All the differential equations describing the MAPK cascade and further explanations for the tQSSA are given in appendix A. The parameter values used for the MAPK cascade are given in figure captions or in table 1.

Our simulations confirm that the cascade can reach a steady state as well as oscillate also in this case. However, with parameters very similar to [3] the (competitive) sQSSA approximation can lead to qualitatively wrong conclusions such as oscillations when the full system is steady [figures 5(b) and (c)], or quantitative wrong estimations of, e.g., the amplitude of the oscillations [figure 5(a)], or the steady-state levels of MAPK-PP [figure 5(d)], while the full system is in general much better approximated by the (competitive) tQSSA. However, for some parameters the tQSSA approach also fails qualitatively [figure 5(b)], or with respect to the period of the oscillations [figure 5(a)].

In table 1, we summarize the ranges of the inhibition constant K_I for which MAPK-PP oscillates in the three cases. It is seen that the solution of the full system undergoes oscillations for a very narrow range of this parameter, while

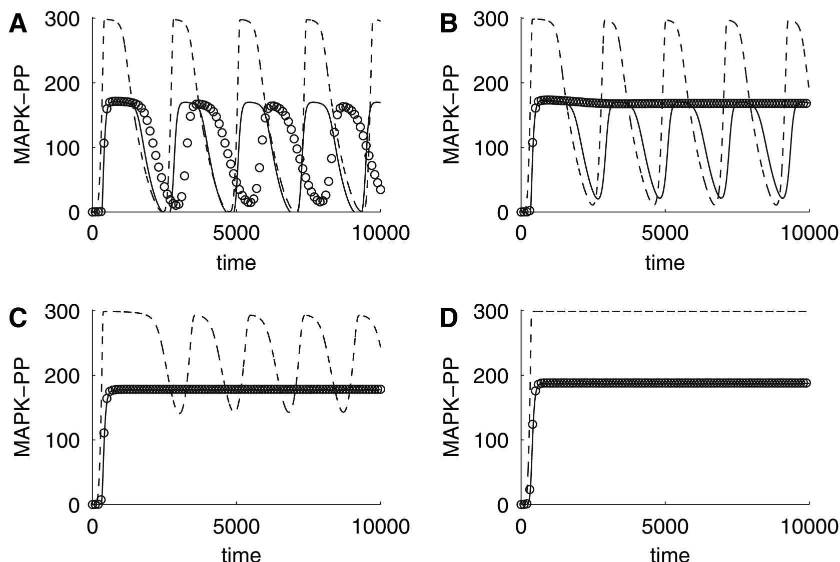


Figure 5. Simulations of the MAPK cascade with feedback as in figure 4. The computed MAPK-PP concentration is shown following the legends in figure 3. The values of the inhibition constant are as follows: Panel A: $K_{\mathrm{I}} = 0.1$. Panel B: $K_{\mathrm{I}} = 0.5$. Panel C: $K_{\mathrm{I}} = 2.5$. Panel D: $K_{\mathrm{I}} = 20$. At low values of K_{I} (panel A), all the three schemes, full system (circles), tQSSA (full curve) and sQSSA (dashed curve), produce oscillations, but the tQSSA follows the solution much better than the sQSSA, especially with respect to the amplitude of the oscillations. In panel C, although the MAPK-PP modelled by the full system (circles) almost immediately reaches a steady-state, the sQSSA (dashed curve) shows oscillations. On the other hand, the full system is followed very well by the tQSSA (full curve). However, this is not always the case, since the tQSSA can also predict oscillations when the full system is stable (panel B). Finally, at high values of K_{I} all the three approaches go to a steady state, but the sQSSA overestimates the MAPK-PP level significantly (panel D).

the use of the sQSSA yields oscillations for a much larger range, also for values for which the solution of the full system does not perform rhythmic behavior [Figures 5(b) and (c)]. However, the competitive tQSSA also fails to predict the behavior for some parameters, but the range for which this occurs is markedly reduced compared to the sQSSA (table 1). New improved tQSSAs should be developed in order to get a better representation of the full system.

The great majority of authors using the sQSSA usually neglects the concentration of the complexes, as expressed, e.g., in the conservation law $[\mathrm{MAPK}]_{\mathrm{T}} = [\mathrm{MAPK}]+[\mathrm{MAPK\text{-}P}]+[\mathrm{MAPK\text{-}PP}]$ [3, 26], but this is only valid at low enzyme concentrations. We suppose that this is the major reason for the poor prediction of the sQSSA. Figure 6(a) shows the complex (MAPK-P)–(MAPKK-PP) of reaction 8 in figure 4, the substrate MAPK-P and the free enzyme MAPKK-PP. In contrast with the sQSSA, the tQSSA considers the complex concentrations, and it is seen from figure 6(a) that this is necessary, since the total substrate concentration is comparable with the complex concentration.

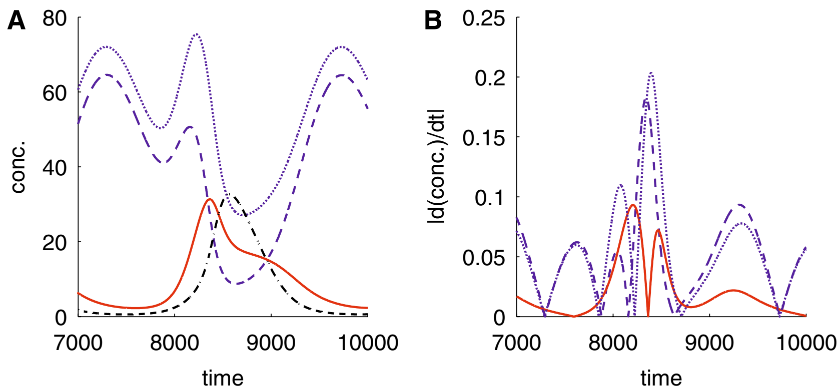


Figure 6. Non-neglectible and non-constant complex concentrations in the MAPK cascade. In the full system describing the MAPK cascade with feedback, the complex (MAPK-P)–(MAPKK-PP) is neither negligible (panel A) nor approximately constant (panel B). Panel A shows the concentrations of the complex (MAPK-P)–(MAPKK-PP) (red, full curve), MAPK-P (blue, dashed curve), "total MAPK-P" ($\overline{\text{MAPK-P}}$; blue, dotted curve) and MAPKK-PP (black, dash-dot curve) during the last part of the simulation of the full system from figure 5(b). Panel B shows the absolute value of the time derivative of the complex, MAPK-P and $\overline{\text{MAPK-P}}$ (same legends as in panel A).

Taking the complex concentrations into account is not only important for the generation of oscillations, but, as it could be expected, also for the steady-state concentrations obtained for the full system, the sQSSA and the tQSSA at high values of the inhibition constant K_I (table 1). With the parameters used here the competitive sQSSA overestimates the steady-state level of activated MAPK [figure 5(d)] as in the simpler case considering only the last level of the cascade [figure 3(b)]. On the other hand, the competitive tQSSA estimates this level well [figures 5(c) and (d)]. As mentioned in the previous section, the correct estimation of activated MAPK-PP has important implications for predicting further downstream effects. We follow this question up in the discussion.

A second problem of both the QSSAs is the fact that the complex never enters a steady state during the oscillations. In figure 6(b), we show the time derivatives of the concentrations from panel A, which measure the rate of change. The assumption that the complex concentration changes much more slowly than the substrate lies at the heart of the QSSAs, but this does not hold in general; as seen in figure 6(b) the rate of change of the complex is comparable to that of the substrate, the total substrate and the kinase. We believe that this is why the tQSSA also fails for some parameters and, moreover, sometimes estimates the period of the oscillations badly [figure 5(b)]. Consequently, it would in some cases be preferable to model the network by means of the full system. The implications of this approach and its flaws will be faced in the discussion.

5. Conclusions and perspectives

Far from being a collection of serial chemical reactions, the higher eukaryotic intracellular signal transduction networks are highly complex. The increased amount of data and knowledge about these networks has made mathematical modeling and computational methods increasingly important in Systems Biology, and has led to projects such as the Silicon Cell, which aims at being a precise replica of the living cell. This means using experimentally found data and reproducing both qualitative and quantitative behavior of the cell.

In a recent paper [45] the so-called Van Slyke–Cullen mechanism, i.e., an irreversible enzyme reaction, with inflow and outflow was studied. It was shown that the sQSSA of the system would always approach a stable steady state even though the full system permitted an unstable steady state and even a limit cycle. It is readily shown that these observations can be extended to an open MM mechanism by changing some of the constants in the computations. Thus, in line with our results the authors showed that the sQSSA must be used with care [45].

So far most of the models describing enzyme reactions, e.g., in the MAPK cascade, have been based on the classical MM approximation (sQSSA) and many of these did not consider competition between substrates. These approaches were taken, although parameters and initial conditions were chosen so that the validity criterion for the sQSSA no longer held and the competition could not be neglected. As exceptions, we mention Hatakeyama et al. [25] and Markevich et al. [26], who treated the problem of substrate competition in terms of the sQSSA for two substrates competing for the same kinase, but they did not consider the region of validity of the sQSSA and neglected the enzyme–substrate complexes.

Although it was known that the sQSSA will often be invalid *in vivo*, the sQSSA approach was necessary for many years, since no better approximations were known, but this has changed recently with the introduction of the tQSSA. This approach was first applied to the simplest reactions [21, 22], and later to increasingly more complex schemes such as reversible reactions [23] and fully competing systems [24].

We have here presented the application of the tQSSA to biologically realistic networks, and shown that it is superior to the sQSSA in all the presented cases. We did not formally investigate the validity of the tQSSA for all the reaction networks examined, and found in fact that the tQSSA has its limitations as well (table 1), probably related to the fact that the complexes do not always enter a quasi-steady state (figure 6b). However, based on our simulations we feel confident in saying that compared to the sQSSA it provides a more accurate estimate of the behavior of enzyme networks. For example, it was found that the tQSSA estimates the steady state levels of activated MAPK very well [figures 3 and 5(a)], while the sQSSA often fails dramatically. We believe that the main reason for this is the fact that the tQSSA incorporates the complex concentrations while the sQSSA does not, as stated for example in the conservation law

$[\text{MAPK}]_\text{T} = [\text{MAPK}] + [\text{MAPK-P}] + [\text{MAPK-PP}]$ [3, 26]. We also showed that the choice of the approximation scheme could dramatically change the size of the parameter range in which oscillations occur in the MAPK cascade with a competitive, negative feedback [figure 5(a) and table 1].

As already noted in section 3, to illustrate the importance of a reliable estimation of the MAPK dynamics and steady-state levels, we remark that is has been observed that both the duration and intensity of the activated MAPK is crucial for the fate of the cell ([37, 40, 41, 46] and references therein). For example, rat PC12 cells differentiate if stimulated by nerve growth factor (NGF) and proliferate if stimulated by epidermal growth factor (EGF) [47–50], although the cognate receptors use the same signaling cytoplasmic network to transduce the signal to the nucleus. In the two cases, the most evident difference is that NGF induces a sustained MAPK (ERK) activity, while EGF induces a transient MAPK (ERK) activity (see [41] for a review). Recently, it was also shown that PC12 [48] and Kaposi Sarcoma [51] cells are sensitive to the strength of the MAPK signal indicating a threshold phenomenon, which means that even minor changes in the levels of activated MAPK can have dramatic consequences.

If any Silicon Cell should help to discover pharmaceutically sensitive targets and reproduce the effects of drugs on these targets, the quantitative aspects of the model would have to be carefully studied and resolved, for example in estimating the size of the above parameter windows. For instance, continuing the example of the MAPK cascade with inhibitory feedback, assume that we wish to apply a drug in order to create oscillations. Lowering the K_I value would appear promising on the basis of the model using the sQSSA, since this model predicts oscillations in a rather wide parameter range (table 1). However, this could encourage a waste of resources searching for an appropriate pharmaceutical compound, since the drug would have to be very finely tuned and, hence, difficult to find, because the full system has a very narrow parameter range yielding oscillations. Thus, one might be better off looking for a drug acting elsewhere in the network.

Since the tQSSA, although superior to the sQSSA, also does not always work, one could suggest to use the alternative of simulating each step of the reaction by means of the full system of ODEs, which means describing every reaction in terms of two equations, and facing three instead of two parameters for every reaction, as it has been done for example for the MAPK cascade [4, 5, 52, 53], However, more equations would mean, especially for larger systems, that this approach quickly would become computer expensive.

A more serious problem is the fact that the three rate constants [a, d, and k in (1)] are usually unknown, while finding the QSSA parameters K_M and V_max (or $k_\text{cat} = V_\text{max}/E_\text{T}$) is a standard procedure in biochemistry. Thus, the reduction obtained from the QSSA is in this sense an advantage compared to the full system. We could in any case rebuild the parameters a, d, k starting from the

MM parameters, but, as shown in the Introduction, we then introduce a degree freedom. Bhalla and Iyengar [2] try to overcome this problem supposing that $d = 4k$, but this hypothesis seems to us a bit arbitrary without any strong experimental support, as already remarked by the authors. However, we have applied this assumption through out this work when modeling the full system.

The validity of the tQSSA depends on the precise values of a, d, and k as stated for example in (8) for the case of a single reaction: the smaller the ratio K/K_{M}, i.e., the larger the ratio d/k, the better the approximation. However, for any choice with a large ratio d/k, the tQSSA holds. A similar result holds for fully competitive reactions [24]. This is consistent with the choice of $d = 4k$ and supported by the fact that for many enzymes the parameter d is much greater than k [54, 55].

From a theoretical point of view, the application of the tQSSA in this way makes the actual parameter values of a, d, and k less important. When we *a priori* know that the system can be well-approximated by the tQSSA, all the possible choices of a, d, k will give approximations near each other, and hence, near the true solution, assuming that the true parameters are such that the tQSSA is valid. This can be used in cases where only the parameters K_{M} and V_{max} are available, the sQSSA is known not to hold, and only a very complicated tQSSA, too complicated to implement effectively on a computer, exists. One can then choose any relation between a, d, and k giving the correct values for K_{M} and V_{max}, check that the tQSSA holds using a theoretically founded validity criteria, and then do the simpler implementation of the full system of equations.

Related to the above, but from another point of view, is the lack of reliable experimental data about the kinetic constants of the intracellular biochemical reactions, including K_{M} and V_{max} values. To reconstruct these missing parameter values, some authors rely on the so-called reverse engineering (or inverse problem). The classical approach to reverse engineering is based on least square techniques with the aim to find the set of parameters that gives the best fitting curve, i.e., the curve passing “as close as possible” to the experimental data. This is done searching for the global minimum of a function of as many variables as there are unknown parameters. To find the global minimum of these functions is in general far from trivial, for example due to the risk of finding only local, not global, minima. Furthermore the uniqueness of this minimum cannot, in general, be guaranteed; several sets of parameters could give the global minimum. This is the question of *a priori identifiability* [56].

As shown in the present work, the misuse of the sQSSA can lead to large quantitative and qualitative errors. However, even when the sQSSA is not a good approximation of the system, we can still find parameters for which the sQSSA does fit the data (the full system), by minimizing, e.g., the least square error. This would inevitably lead to wrongly estimated parameters, since the original ones did not provide a good approximation [57].

From these considerations it follows that the ability of the model to fit a certain data set cannot be used to test whether a certain approximation holds. Applying reverse engineering for the sQSSA, without any *a priori* examination of its validity, one could argue that the (mis)use of the sQSSA causes no problems, since we can obtain a good fit anyway. However, one would prefer to have a model that works under many different conditions, not only in a certain experimental setting. If fitting the sQSSA model to the data yields wrong estimates of the parameters, then it is likely that the predicted behavior using these parameters would be far from the true behavior. The same would be true if the model was later used as a subsystem of an enlarged model. For example, the estimation of the Michaelis or inhibition constant relying on a wrong model formulation could be crucial, as seen in the following example.

Assume that all the parameters except the inhibition constant K_I were known for the model illustrated by figure 4. If we had a data set for this model showing stable behavior, according to table 1, using the sQSSA we would estimate a value of K_I greater than 3.02, even though the true value of K_I could be between 0.18 and 3.02. Assume now that we obtain a drug capable of lowering the K_I value according to some known mechanisms, and that we decide to administrate the drug to lower K_I with the aim to let the system oscillate. Believing that the sQSSA estimated K_I is the true value, we would apply a certain amount of the drug in order to get below the threshold value at $K_I = 3.02$. But the actual value of K_I could be completely different from the wrongly estimated one and such that the drug administration, though lowering K_I, would leave the system stable.

Similar problems can be expected to occur in metabolic control analysis [52, 58, 59], which is used to find the steps in the network that controls some output behavior, e.g., the concentration of a certain biochemical species. It seems likely that an invalid sQSSA model might predict that a certain step is the most important, while the full system or the corresponding tQSSA model finds that step to be less important. In the light of applications for the pharmaceutical industry, this could lead to a waste of money and energy focusing on an apparently sensitive target, which then turns out to be unimportant or, vice versa, the neglect of an important target that apparently seems unimportant.

In conclusion, we have shown that the use of the classical MM approach (sQSSA) should be done with much care, since it can lead to both quantitative and qualitative errors. This has further impact on techniques such as reverse engineering and metabolic control analysis. Finding approximations improving the sQSSA for complex reactions such as successive reactions, open systems, loops such as the Goldbeter–Koshland switch [60], feedback systems, etc., and investigating their validity, should be of great interest for further investigations and simulations of such reactions *in vivo*, where the MM description can be expected to break down.

Acknowledgment

M. G. Pedersen was supported by BioSim (EU contract No. 005137).

Appendix A All equations and parameters

- The equations for the single, non-competitive reaction (1) are described in the Introduction [equations (2), (5), and (6)].
  The parameters used are given in the caption of figure 1 using the relation $d = 4k$ for the full system as described in the Introduction.
- The equations for the sQSSA and the tQSSA of the fully competitive reactions (12) are given in section 2 [(13) and (15)].
  The full system is given by $i = 1, 2$,

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = -a_i E \cdot S_i + d_i C_i, \quad S_i(0) = S_{i,\mathrm{T}}, \tag{25a}$$

$$\frac{\mathrm{d}C_i}{\mathrm{d}t} = a_i(E \cdot S_i - K_i^{\mathrm{M}} C_i), \quad C_i(0) = 0, \quad K_i^{\mathrm{M}} = \frac{d_i + k_i}{a_i}, \tag{25b}$$

  and the conservation laws

$$S_{i,\mathrm{T}} = S_i + C_i + P_i, \quad i = 1, 2, \tag{26}$$

$$E_{\mathrm{T}} = E + C_1 + C_2. \tag{27}$$

  The parameters used are given in the caption of figure 2 using the relation $d = 4k$ for the full system as described in the Introduction.
- The equations for the two successive reactions (16) are descibed by:
  Full system:

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = -a_1 E \cdot S_1 + d_1 C_1, \quad S_1(0) = S_{\mathrm{T}}, \tag{28a}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = k_1 C_1 - a_2 E \cdot S_2 + d_2 C_2, \quad S_2(0) = 0, \tag{28b}$$

$$\frac{\mathrm{d}C_i}{\mathrm{d}t} = a_i(E \cdot S_i - K_i^{\mathrm{M}} C_i), \quad C_i(0) = 0, \quad K_i^{\mathrm{M}} = \frac{d_i + k_i}{a_i}, \quad i = 1, 2, \tag{28c}$$

$$S_{\mathrm{T}} = S_1 + S_2 + C_1 + C_2 + P, \tag{28d}$$

$$E_{\mathrm{T}} = E + C_1 + C_2. \tag{28e}$$

  sQSSA:

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} \approx -\frac{k_1 E_{\mathrm{T}} S_1}{K_1^{\mathrm{M}}(1 + S_2/K_2^{\mathrm{M}}) + S_1}, \quad S_1(0) = S_{\mathrm{T}}, \tag{29a}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} \approx \frac{k_1 E_{\mathrm{T}} S_1}{K_1^{\mathrm{M}}(1 + S_2/K_2^{\mathrm{M}}) + S_1} - \frac{k_2 E_{\mathrm{T}} S_2}{K_2^{\mathrm{M}}(1 + S_1/K_1^{\mathrm{M}}) + S_2}, \quad S_2(0) = 0, \tag{29b}$$

$$P = S_{\mathrm{T}} - S_1 - S_2. \tag{29c}$$

tQSSA:

$$\frac{d\bar{S}_1}{dt} \approx -k_1\, C_1(\bar{S}_1, \bar{S}_2), \quad \bar{S}_1(0) = S_T, \tag{30a}$$

$$\frac{d\bar{S}_2}{dt} \approx k_1\, C_1(\bar{S}_1, \bar{S}_2) - k_2\, C_2(\bar{S}_1, \bar{S}_2), \quad \bar{S}_2(0) = 0, \tag{30b}$$

$$P = S_T - \bar{S}_1 - \bar{S}_2, \tag{30c}$$

where the expression for $C_i(\bar{S}_1, \bar{S}_2), i = 1, 2$, are given by finding the root of ψ_1 given by (14) and the corresponding expression for C_2.
The parameters used are given in the caption of figure 2 using the relation $d = 4k$ for the full system as described in the Introduction.

- The double phosphorylation and dephosphorylation scheme (17) is described by:
Full system:

$$\frac{dM}{dt} = -a_1 E \cdot M + d_1 C_1 + k_4 C_4, \quad M(0) = [\text{MAPK}]_T, \tag{31a}$$

$$\frac{dM_p}{dt} = -a_2 E \cdot M_p + d_2 C_2 + k_1 C_1 - a_4 F \cdot M_p + d_4 C_4 + k_3 C_3, \quad M_p(0) = 0, \tag{31b}$$

$$\frac{dM_{pp}}{dt} = -a_3 F \cdot M_{pp} + d_3 C_3 + k_2 C_2, \quad M_{pp}(0) = 0, \tag{31c}$$

$$\frac{dC_1}{dt} = -(d_1 + k_1)\, C_1 + a_1 E \cdot M, \quad C_1(0) = 0, \tag{31d}$$

$$\frac{dC_2}{dt} = -(d_2 + k_2)\, C_2 + a_2 E \cdot M_p, \quad C_2(0) = 0, \tag{31e}$$

$$\frac{dC_3}{dt} = -(d_3 + k_3)\, C_3 + a_3 F \cdot M_{pp}, \quad C_3(0) = 0, \tag{31f}$$

$$\frac{dC_4}{dt} = -(d_4 + k_4)\, C_4 + a_4 F \cdot M_p, \quad C_4(0) = 0, \tag{31g}$$

where M=[MAPK], M_p=[MAPK-P], M_{pp}=[MAPK-PP], E=[MAPKK-PP], F=[MKP], and the concentration of the complex of reaction v_i is denoted by C_i, $i = 1, 2, 3, 4$. The conservation laws are (the first one is a consequence of the ODEs, and thus redundant and only included for clarity):

$$[\text{MAPK}]_T = M + M_p + M_{pp} + C_1 + C_2 + C_3 + C_4,$$

$$[\text{MAPKK-PP}]_T = E + C_1 + C_2 = E_T,$$

$$[\text{MKP}]_T = F + C_3 + C_4 = F_T.$$

sQSSA:

$$\frac{dM}{dt} \approx v_4 - v_1, \quad M(0) = [\text{MAPK}]_T, \tag{32a}$$

$$\frac{dM_p}{dt} \approx v_1 - v_2 + v_3 - v_4, \quad M_p(0) = 0, \tag{32b}$$

$$\frac{dM_{pp}}{dt} \approx v_2 - v_3, \quad M_{pp} = 0, \tag{32c}$$

with the rates $v_1 - v_4$ given by:

$$v_1 = \frac{k_1 E_T M}{K_1^M(1 + M_p/K_2^M) + M}, \quad v_2 = \frac{k_2 E_T M_p}{K_2^M(1 + M/K_1^M) + M_p},$$
$$v_3 = \frac{k_3 F_T M_{pp}}{K_3^M(1 + M_p/K_4^M) + M_{pp}}, \quad v_4 = \frac{k_4 F_T M_p}{K_4^M(1 + M_{pp}/K_3^M) + M_p}.$$

tQSSA: The equations are described in the main text (18)–(24).
The parameters are taken from [26, figure 1] and are given in the caption of figure 3. They are given in arbitrary units, but for consistency with [26] one can think of nM for all concentrations, and seconds for all time units. Other parameters should then be expressed in these units.

- The MAPK cascade of figure 4 is modeled by:
  Full system:

$$\begin{aligned}
\frac{dS_2}{dt} &= k_1C_1 - a_2S_2F_2 + d_2C_2 \\
&\quad +(d_3 + k_3)C_3 - a_3S_2S_3 + (d_4 + k_4)C_4 - a_4S_2S_4, \quad S_2(0) = 0,\\
\frac{dS_4}{dt} &= k_3C_3 - a_4S_4S_2 + d_4C_4 + k_5C_5 - a_6S_4F_5 + d_6C_6, \quad S_4(0) = 0,\\
\frac{dS_5}{dt} &= k_4C_4 - a_5S_5F_5 + d_5C_5 \\
&\quad +(d_7 + k_7)C_7 + (d_8 + k_8)C_8 - a_7S_7S_5 - a_8S_8S_5, \quad S_5(0) = 0,\\
\frac{dS_8}{dt} &= k_7C_7 - a_8S_8S_5 + d_8C_8 + k_9C_9 - a_{10}S_8F_9 + d_{10}C_{10}, \quad S_8(0) = 0,\\
\frac{dS_9}{dt} &= k_8C_8 - a_9S_9F_9 + d_9C_9 + k_-^IC_{\text{inh}} - k_+^IS_9E_1, \quad S_9(0) = 0,\\
\frac{dC_1}{dt} &= a_1S_1E_1 - (d_1 + k_1)C_1, \quad C_1(0) = 0,\\
\frac{dC_2}{dt} &= a_2S_2F_2 - (d_2 + k_2)C_2, \quad C_2(0) = 0,
\end{aligned}$$

$$\begin{aligned}
\frac{dC_3}{dt} &= a_3 S_3 S_2 - (d_3 + k_3)C_3, \quad C_3(0) = 0,\\
\frac{dC_4}{dt} &= a_4 S_4 S_2 - (d_4 + k_4)C_4, \quad C_4(0) = 0,\\
\frac{dC_5}{dt} &= a_5 S_5 F_5 - (d_5 + k_5)C_5, \quad C_5(0) = 0,\\
\frac{dC_6}{dt} &= a_6 S_4 F_5 - (d_6 + k_6)C_6, \quad C_6(0) = 0,\\
\frac{dC_7}{dt} &= a_7 S_7 S_5 - (d_7 + k_7)C_7, \quad C_7(0) = 0,\\
\frac{dC_8}{dt} &= a_8 S_8 S_5 - (d_8 + k_8)C_8, \quad C_8(0) = 0,\\
\frac{dC_9}{dt} &= a_9 S_9 F_9 - (d_9 + k_9)C_9, \quad C_9(0) = 0,\\
\frac{dC_{10}}{dt} &= a_{10} S_8 F_9 - (d_{10} + k_{10})C_{10}, \quad C_{10}(0) = 0,\\
\frac{dC_{\text{inh}}}{dt} &= k_+^{\text{I}} E_1 S_9 - k_-^{\text{I}} C_{\text{inh}}, \quad C_{\text{inh}}(0) = 0,
\end{aligned}$$

with conservation laws

$$\begin{aligned}
S_1 &= S_{1,\text{T}} - S_2 - C_1 - C_2 - C_3 - C_4,\\
S_3 &= S_{3,\text{T}} - S_4 - S_5 - C_3 - C_4 - C_5 - C_6 - C_7 - C_8,\\
S_7 &= S_{7,\text{T}} - S_8 - S_9 - C_7 - C_8 - C_9 - C_{10} - C_{\text{inh}},\\
F_2 &= F_{2,\text{T}} - C_2,\\
F_5 &= F_{5,\text{T}} - C_5 - C_6,\\
F_9 &= F_{9,\text{T}} - C_9 - C_{10},\\
E_1 &= E_{1,\text{T}} - C_1 - C_{\text{inh}}.
\end{aligned}$$

Here S_1=[MKKK], S_2=[MKKK-P], S_3=[MAPKK], S_4=[MAPKK-P], S_5=[MAPKK-PP], S_7=[MAPK], S_8=[MAPK-P], S_9=[MAPK-PP]; E_1 the concentration of a kinase (e.g., Ras), while F_2, F_5, and F_9 are concentrations of phosphatases; C_j, $j = 1, \ldots, 10$, the complex of reaction i and C_{inh} the inhibition complex E_1–(MAPK-PP).
sQSSA:
Introducing

$$K_i^{\text{M}} = \frac{d_i + k_i}{a_i}, \quad i = 1, \ldots, 10, \quad K_{\text{I}} = \frac{k_-^{\text{I}}}{k_+^{\text{I}}}, \tag{33}$$

the sQSSA is

$$
\begin{aligned}
\frac{dS_2}{dt} &\approx v_1 - v_2, \quad S_2(0) = 0,\\
\frac{dS_4}{dt} &\approx v_3 - v_4 + v_5 - v_6, \quad S_4(0) = 0,\\
\frac{dS_5}{dt} &\approx v_4 - v_5, \quad S_5(0) = 0,\\
\frac{dS_8}{dt} &\approx v_7 - v_8 + v_9 - v_{10}, \quad S_8(0) = 0,\\
\frac{dS_9}{dt} &\approx v_8 - v_9, \quad S_9(0) = 0,\\
[\mathrm{MKKK}]_\mathrm{T} &= \bar{S}_1 + \bar{S}_2,\\
[\mathrm{MAPKK}]_\mathrm{T} &= \bar{S}_3 + \bar{S}_4 + \bar{S}_5,\\
[\mathrm{MAPK}]_\mathrm{T} &= \bar{S}_7 + \bar{S}_8 + \bar{S}_9
\end{aligned}
$$

with the rates given by

$$
\begin{aligned}
v_1 &= \frac{k_1 E_{1,\mathrm{T}} S_1}{K_1^\mathrm{M}(1 + S_9/K_\mathrm{I}) + S_1}, \quad v_2 = \frac{k_2 F_{2,\mathrm{T}} S_2}{K_2^\mathrm{M} + S_2},\\
v_3 &= \frac{k_3 S_2 S_3}{K_3^\mathrm{M}(1 + S_4/K_4^\mathrm{M}) + S_3}, \quad v_4 = \frac{k_4 S_2 S_4}{K_4^\mathrm{M}(1 + S_3/K_3^\mathrm{M}) + S_4},\\
v_5 &= \frac{k_5 F_{5,\mathrm{T}} S_5}{K_5^\mathrm{M}(1 + S_4/K_6^\mathrm{M}) + S_5}, \quad v_6 = \frac{k_6 F_{5,\mathrm{T}} S_4}{K_6^\mathrm{M}(1 + S_5/K_5^\mathrm{M}) + S_4},\\
v_7 &= \frac{k_7 S_5 S_7}{K_7^\mathrm{M}(1 + S_8/K_8^\mathrm{M}) + S_7}, \quad v_8 = \frac{k_8 S_5 S_8}{K_8^\mathrm{M}(1 + S_7/K_7^\mathrm{M}) + S_8},\\
v_9 &= \frac{k_9 F_{9,\mathrm{T}} S_9}{K_9^\mathrm{M}(1 + S_8/K_{10}^\mathrm{M}) + S_9}, \quad v_{10} = \frac{k_{10} F_{9,\mathrm{T}} S_8}{K_{10}^\mathrm{M}(1 + S_9/K_9^\mathrm{M}) + S_8}.
\end{aligned}
$$

tQSSA:
The total substrates are

$$
\begin{aligned}
&\bar{S}_1 = S_1 + C_1, \quad \bar{S}_2 = S_2 + C_2 + C_3 + C_4,\\
&\bar{S}_3 = S_3 + C_3, \quad \bar{S}_4 = S_4 + C_4 + C_6, \quad \bar{S}_5 = S_5 + C_5 + C_7 + C_8,\\
&\bar{S}_7 = S_7 + C_7, \quad \bar{S}_8 = S_8 + C_8 + C_{10}, \quad \bar{S}_9 = S_9 + C_9 + C_\mathrm{inh},
\end{aligned}
$$

which are governed by the equations

$$
\begin{aligned}
\frac{d\bar{S}_2}{dt} &\approx v_1 - v_2, \quad \bar{S}_2(0) = 0,\\
\frac{d\bar{S}_4}{dt} &\approx v_3 - v_4 + v_5 - v_6, \quad \bar{S}_4(0) = 0,\\
\frac{d\bar{S}_5}{dt} &\approx v_4 - v_5, \quad \bar{S}_5(0) = 0,\\
\frac{d\bar{S}_8}{dt} &\approx v_7 - v_8 + v_9 - v_{10}, \quad \bar{S}_8(0) = 0,\\
\frac{d\bar{S}_9}{dt} &\approx v_8 - v_9, \quad \bar{S}_9(0) = 0,\\
[\mathrm{MKKK}]_\mathrm{T} &= \bar{S}_1 + \bar{S}_2,\\
[\mathrm{MAPKK}]_\mathrm{T} &= \bar{S}_3 + \bar{S}_4 + \bar{S}_5,\\
[\mathrm{MAPK}]_\mathrm{T} &= \bar{S}_7 + \bar{S}_8 + \bar{S}_9
\end{aligned}
$$

with rates

$$
\begin{aligned}
v_1 &= k_1\, C_1(\bar{S}_1, \bar{S}_9 - C_9(\bar{S}_9, \bar{S}_8; F_9); E_1),\\
v_2 &= k_2\, C_2(\bar{S}_2 - C_3(\bar{S}_3, \bar{S}_4; \bar{S}_2) - C_4(\bar{S}_4, \bar{S}_3; \bar{S}_2), 0; F_{2,\mathrm{T}}),\\
v_3 &= k_3\, C_3(\bar{S}_3, \bar{S}_4 - C_6(\bar{S}_4, \bar{S}_5; F_{5,\mathrm{T}}); \bar{S}_2 - C_2(\bar{S}_2, 0; F_{2,\mathrm{T}})),\\
v_4 &= k_4\, C_4(\bar{S}_4 - C_6(\bar{S}_4, \bar{S}_5; F_{5,\mathrm{T}}), \bar{S}_3; \bar{S}_2 - C_2(\bar{S}_2, 0; F_{2,\mathrm{T}})),\\
v_5 &= k_5\, C_5(\bar{S}_5 - C_7(\bar{S}_7, \bar{S}_8; \bar{S}_5) - C_8(\bar{S}_8, \bar{S}_7; \bar{S}_5), \bar{S}_4 - C_4(\bar{S}_4, \bar{S}_3; \bar{S}_2); F_{5,\mathrm{T}}),\\
v_6 &= k_6\, C_6(\bar{S}_4 - C_4(\bar{S}_4, \bar{S}_3; \bar{S}_2), \bar{S}_5 - C_7(\bar{S}_7, \bar{S}_8; \bar{S}_5) - C_8(\bar{S}_8, \bar{S}_7; \bar{S}_5); F_{5,\mathrm{T}}),\\
v_7 &= k_7\, C_7(\bar{S}_7, \bar{S}_8 - C_{10}(\bar{S}_8, \bar{S}_9; F_{9,\mathrm{T}}); \bar{S}_5 - C_5(\bar{S}_5, \bar{S}_4; F_{5,\mathrm{T}})),\\
v_8 &= k_8\, C_8(\bar{S}_8 - C_{10}(\bar{S}_8, \bar{S}_9; F_{9,\mathrm{T}}), \bar{S}_7; \bar{S}_5 - C_5(\bar{S}_5, \bar{S}_4; F_{5,\mathrm{T}})),\\
v_9 &= k_9\, C_9(\bar{S}_9 - C_{\mathrm{inh}}(\bar{S}_9, \bar{S}_1; E_1), \bar{S}_8 - C_8(\bar{S}_8, \bar{S}_7; \bar{S}_5); F_{9,\mathrm{T}}),\\
v_{10} &= vk_{10}\, C_{10}(\bar{S}_8 - C_8(\bar{S}_8, \bar{S}_7; \bar{S}_5), \bar{S}_9 - C_{\mathrm{inh}}(\bar{S}_9, \bar{S}_1; E_1); F_{9,\mathrm{T}}),
\end{aligned}
$$

$C_i(X, Y; Z_\mathrm{T})$ indicates the complex of reaction i with substrate X, competing substrate Y and enzyme Z, and is found as indicated for scheme (12). Specifically, in (14) the following substitutions should be made, and the root then found: C_1 by C_i, $\bar{S}_1$ by X, $\bar{S}_2$ by Y, E_T by Z_T, K_1^M by K_i^M, and K_2^M by K_j^M, where j indicates the competing reaction, see scheme (17). For C_1, K_2^M should be substituted by K_I because of the competitive inhibition, while for C_{inh}, K_1^M should be substituted by K_I and K_2^M by K_1^M.

The idea behind the expressions above is that we still do not have the exact tQSSA for reactions of the given type. The problem is that the total substrates involve several complexes not all relevant for the reaction under consideration, and this is further complicated by the fact that several of the species, e.g., S_2, act both as a substrate and an enzyme. We calculate a first approximation to

Table 2
Parameter values in the MAPK cascade model [3]. For the full system the remaining parameters are calculated as explained above and from (33).

$[\text{MKKK}]_\text{T} = 100$	$[\text{MAPKK}]_\text{T} = 300$	$[\text{MAPK}]_\text{T} = 300$	$E_{1,\text{T}} = 100$
$F_{2,\text{T}} = 10$	$F_{5,\text{T}} = 10$	$F_{9,\text{T}} = 10$	$K_\text{I} =$ varies
$K_1^\text{M} = 10$	$K_2^\text{M} = 8$	$K_j^\text{M} = 15$, $j = 3, \dots, 10$	
$k_+^\text{I} = 1$	$k_j = 0.025$, $j = 1, \dots, 4$	$k_5 = 0.075$	$k_6 = 0.075$
$k_7 = 0.025$	$k_8 = 0.025$	$k_9 = 0.05$	$k_{10} = 0.05$

each of these complexes and subtract it from the total substrate. New tQSSAs should be developed to improve on this *ad hoc* approach.
The parameters are based on [3], but modified to competitive inhibition. In order to have the same magnitude on the k_i values, we chose the level of E_1 and the phosphatases appropriately. The parameters are given in table 2.

References

[1] C.-Y. F. Huang and J. E. Ferrell, Proc. Natl. Acad. Sci. 93 (1996) 10078–10083.
[2] U.S. Bhalla and R. Iyengar, Science 283 (1999) 381–387.
[3] B.N. Kholodenko, Eur. J. Biochem. 267 (2000) 1583–1588.
[4] B. Schoeberl, C. Eichler-Jonsson, E.D. Gilles and G. Muller, Nat. Biotechnol. 20 (2002) 370–375.
[5] S. Sasagawa, Y.-I. Ozaki, K. Fujita and S Kuroda, Nat. Cell Biol. 7 (2005) 365–373.
[6] L.M. Loew and J.C. Schaff, Trends Biotechnol. 19 (2001) 401–406.
[7] J.L. Snoep, Curr. Opin. Biotechnol. 16 (2005) 336–343.
[8] R.J. Ellis, Trends Biochem. Sci. 26 (2001) 597–604.
[9] V. Henri, C.R. Hebd. Acad. Sci. 133 (1901) 891–899.
[10] V. Henri, Z. Phys. Chem. 39 (1901) 194–216.
[11] V. Henri, C.R. Hebd. Acad. Sci. 135 (1902) 916–919.
[12] L. Michaelis and M.L. Menten, Biochem. Z. 49 (1913) 333–369.
[13] G.E. Briggs and J.B.S. Haldane, Biochem. J. 19 (1925) 338–339.
[14] L.A. Segel, Bull. Math. Biol. 50 (1988) 579–593.
[15] L.A. Segel and M. Slemrod, SIAM Rev. 31 (1989) 446–477.
[16] K.R. Albe, M.H. Butler and B.E. Wright, J. Theor. Biol. 143 (1990) 163–195.
[17] O.H. Straus and A. Goldstein, J. Gen. Physiol. 26 (1943) 559–585.
[18] A. Sols and R. Marco, Implications in metabolic regulation. in: *Current topics in Cellular Regulation*, Vol. 2 (Academic Press, New York, 1970).
[19] S. Schnell and P. K. Maini, Comm. Theor. Biol. 8 (2003) 169–187.
[20] H. Bisswanger, *Enzyme Kinetics. Principles and Methods*. (Wiley-VCH, Weinheim, 2002).
[21] J.A.M. Borghans, R.J. de Boer and L.A. Segel, Bull. Math. Biol. 58 (1996) 43–63.
[22] A.R. Tzafriri, Bull. Math. Biol. 65 (2003) 1111–1129.
[23] A.R. Tzafriri and E. R. Edelman, J. Theor. Biol. 226 (2004) 303–313.
[24] M.G. Pedersen, A.M. Bersani and E. Bersani, Bull. Math. Biol. 69 (2007) 433–457.
[25] M. Hatakeyama, S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J.H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama and A. Konagaya, Biochem. J. 373 (2003) 451–463.

[26] N.I. Markevich, J.B. Hoek and B.N. Kholodenko, J. Cell Biol. 164 (2004) 353–359.
[27] C. Salazar and T. Höfer, BioSystems 83 (2006) 195–206.
[28] C.L. Frenzen and P.K. Maini, J. Math. Biol. 26 (1988) 689–703.
[29] L.M. Stevenson, M.S. Deal, J.C. Hagopian and J. Lew, Biochemistry 41 (2002) 8528–8534.
[30] M.C. Morris, C. Gondeau, J.A. Tainer and G. Divita, J. Biol. Chem. 277 (2002) 23847–23853.
[31] S. Schnell and C. Mendoza, Bull. Math. Biol. 62 (2000) 321–336.
[32] S.I. Rubinow and J.L. Lebowitz, J. Am. Chem. Soc. 92 (1970) 3888–3893.
[33] W.R. Burack and T.W. Sturgill, Biochemistry 36 (1997) 5929–5933.
[34] J.E. Ferrell and R.R. Bhatt, J. Biol. Chem. 272 (1997) 19008–19016.
[35] Y. Zhao and Z.-Y. Zhang, J. Biol. Chem. 276 (2001) 32382–32391.
[36] M. Camps, A. Nichols and S. Arkinstall, FASEB J. 14 (2000) 6–16.
[37] M. Ebisuya, K. Kondoh and E. Nishida, J. Cell Sci. 118 (2005) 2997–3002.
[38] L. Chang and M. Karin, Nature 410 (2001) 37–40.
[39] T.S. Lewis, P.S. Shapiro and N.G. Ahn, Adv. Cancer Res. 74 (1998) 49–139.
[40] L.O. Murphy and J. Blenis, Trends Biochem. Sci. 31 (2006) 268–274.
[41] C.J. Marshall, Cell 80 (1995) 179–185.
[42] X.L. Zhan, M.J. Wishart and K.L. Guan, Chem. Rev. 101 (2001) 2477–2496.
[43] M.K. Dougherty, J. Muller, D.A. Ritt, M. Zhou, X.Z. Zhou, T.D. Copeland, T.P. Conrads, T.D. Veenstra, K.P. Lu and D.K. Morrison, Mol. Cell. 17 (2005) 215–224.
[44] J.E. Ferrell Jr, Curr. Opin. Cell Biol. 14 (2002) 140–148.
[45] E.H. Flach and S. Schnell, IEE Proc. Syst. Biol. 153 (2006) 187–191.
[46] C.S. Hill and R. Treisman, Cell 80 (1995) 199–211.
[47] S. Traverse, K. Seedorf, H. Paterson, C.J. Marshall, P. Cohen and A. Ullrich, Curr. Biol. 4 (1994) 694–701.
[48] L. New, Y. Li, B. Ge, H. Zhong, J. Mansbridge, K. Liu and J. Han, J. Cell. Biochem. 83 (2001) 585–596.
[49] Y. Gotoh, E. Nishida, T. Yamashita, M. Hoshi, M. Kawakami and H. Sakai, Eur. J. Biochem. 193 (1990) 661–669.
[50] T.T. Nguyen, J.C. Scimeca, C. Filloux, P. Peraldi, J.L. Carpentier and E. Van Obberghen, J. Biol. Chem. 268 (1993) 9803–9810.
[51] C. Bardelli, M. Sala, U. Cavallazzi and M. Prat, Biochem. Biophys. Res. Comm. 334 (2005) 1172–1179.
[52] J.J. Hornberg, F.J. Bruggeman, B. Binder, C.R. Geest, A.J.M.B. de Vaate, J. Lankelma, R. Heinrich and H.V. Westerhoff, FEBS J. 272 (2005) 244–258.
[53] A. Fujioka, K. Terai, R.E. Itoh, K. Aoki, T. Nakamura, S. Kuroda, E. Nishida and M. Matsuda, J. Biol. Chem. 281 (2006) 8917–8926.
[54] D.E. Atkinson, *Cellular Energy Metabolism and its Regulation* (Academic, New York, 1977).
[55] Y. Cao, D.T. Gillespie and L.R. Petzold, J. Chem. Phys. 123 (2005) 144917.
[56] M.P. Saccomani, L. D'Angio, S. Audoly and C. Cobelli, in: *Modelling Methodology for Physiology and Medicine*, eds. E. Carson and C. Cobelli, (Academic, San Diego, 2001) chapter 4, pp. 77–105.
[57] M.G. Pedersen, A.M. Bersani, E. Bersani and G. Cortese, Math. Comput. Simulat. To appear, (2007).
[58] J.J. Hornberg, B. Binder, F.J. Bruggeman, B. Schoeberl, R. Heinrich and H.V. Westerhoff, Oncogene 24 (2005) 5533–5542.
[59] M. Cascante, L.G. Boros, B. Comin-Anduix, P. de Atauri, J.J. Centelles and P.W.-N. Lee, Nat. Biotechnol. 20 (2002) 243–249.
[60] A. Goldbeter and D.E. Koshland Jr., Proc. Natl. Acad. Sci. 78 (1981) 6840–6844.