Available online at www.sciencedirect.com
ScienceDirect
Computers and Mathematics with Applications 55 (2008) 879–888
An International Journal
computers & mathematics
with applications
www.elsevier.com/locate/camwa

Deterministic and stochastic models of enzymatic networks—applications to pharmaceutical research

A.M. Bersani[a,*], E. Bersani[b], L. Mastroeni[c]

[a] *Dipartimento di Metodi e Modelli Matematici, Università "La Sapienza" di Roma, Italy*
[b] *via Biferno 4, 00199 Rome, Italy*
[c] *Dipartimento di Economia, Università "Roma Tre" di Roma, Italy*

Abstract

The intracellular transducing device consists of complex networks of enzymatic reactions. Unfortunately, the mathematical models commonly used to describe them are still unsatisfactory and unreliable, even at the level of reproducing simple reaction schemes. The improvement of mathematical models is necessary and can follow different approaches still poorly employed, such as the modeling of spatial structures and phenomena, time delays, stochastic perturbations, only to cite the most relevant ones. In this paper we show some recent results related to the total quasi-steady-state approximation (tQSSA), in a deterministic scenario. Moreover, we show some possible applications of the tQSSA in a stochastic scheme.

Keywords: Michaelis–Menten kinetics; Quasi-steady-state approximation; Deterministic and stochastic processes; Protein–protein interactions

1. Introduction

Every year, pharma companies invest billions of dollars on lead compound research, with the aim of drug discovery and commercialization. Most of these costs concern laboratory experiments (*in vitro* and *in vivo*) and clinical tests. One of the main goals of the so-called 'Systems Biology' is to help the pharmaceutical research in dramatically reducing the costs of the research phases (mostly preclinical ones) preceding drug commercialization and the time needed to complete all these phases.

The completion (in some sense) of the Human Genome Project, aimed at sequencing the complete human genome, raised great excitation in the genetics community, because this "primary" (genomic) information was believed to be the first step towards the unification between proteomic and genomic data and its use in medical, clinical and pharmaceutical fields. This effort would have brought complete information not only to basic research but also to medical applications to find the best treatment for human disease.

* Corresponding address: Dipartimento Me.Mo.Mat., Università "La Sapienza" di Roma, via A. Scarpa 16, 00161 Rome, Italy. Tel.: +39 06 49766681; fax: +39 06 49766684.
E-mail addresses: bersani@dmmm.uniroma1.it (A.M. Bersani), enrico.bersani@gmail.com (E. Bersani), mastroen@uniroma3.it (L. Mastroeni).

0898-1221/$ - see front matter
doi:10.1016/j.camwa.2006.12.092

The expected results, human gene mapping and specific drug design to modulate gene expression and protein targets, targeted gene engineering and so on, have been disappointing. As can been seen in [1] (see Fig. 1 of that paper), a constant decline in pharmaceutical production began at the middle of the seventies and persisted throughout the "genomic era". The goal of having a complete list of human genes and, consequently, of the complete list of potential gene and protein targets has not been reached, because of the relatively low number of human genes (more or less one-fourth of the expected number) and of the high complexity of gene regulation and protein domain folding. The proteomic level is, in fact, more complex than expected and many "emergent" properties of protein interactome cannot be inferred from the knowledge of the primary structure of a genome. Another level of complexity is the gene transcription, whose general rules escape our efforts in the study of regulation regions surrounding the genes. The potential targets of drugs have exponentially grown in number, and the genomic information is largely insufficient to lead to drug discovery and therapy design.

Succinctly put, the genomic approach alone is not a successful way to deal with the drug design task. At least, we can say that no beneficial boost has been experienced by drug research during the period (1985–2005) of human genome sequencing. The expected result was the dramatic reduction of the time and cost of the so-called preclinical phase of drug production, the one in which bioinformatic, computational and theoretical methods would have helped to select the promising gene and protein targets.

The recent estimate of the costs for new drug development (see [2], Fig. 2), based on data of the eighties, for the preclinical and clinical studies and trials, is of more than 800M US $ for the total course of drug commercialization (a period of about ten years). But a more recent estimate raises these costs to more than 1000M US $ for each successful drug.

Finally, the success rate of the trials is, in the best actual cases, no more than 30% (see [2], Table 1). It can be estimated that a mean of only 10% of the tested new compounds will be commercialized (with a good profit for the pharmas and a beneficial effect on public healthy), and the amount of nonsuccessful research is easily evaluated.

These data highlight the need for new conceptual tools for effective drug discovery. In other words, it is not the large-scale extension of (otherwise) successful experimental and computational tools to improve the pharmaceutical research. A conclusion can be drawn from these data: since the research on new drugs is guided by empirical and trial methods, it would be more efficient to develop new tools for the preclinical phase of the research, i.e. the introduction of new powerful theoretical schemes, based on the huge quantity of experimental data on cell biology and biochemistry. These new tools, which are of most interest in the field of Systems Biology, are mathematical models and computer software to accurately reproduce the intracellular and intercellular phenomena in a quantitative way. A robust predictive mathematical model must be used to address the experimental activity following initial research stages. In other words, the experimental activity must be targeted on a more massive production of quantitative data (kinetic constants, molecular concentrations, geometrical data on the cell structure, etc.) which will be used for the mathematical model construction. The accuracy of these data, together with reliable proteomic data (for example, protein interaction map data) will allow the formulation of mathematical models of increasing complexity and large-scale on the various aspects of the intra and intercellular phenomena. This "virtual cell" project, which has been successfully begun at the end of nineties [3,4] on some selected well-known intracellular signal transduction pathways (MAP kinase cascade, NFkB pathway and others), needs an extension of the present understanding of protein interaction events, for example a reconsideration of the well-known Michaelis–Menten reaction scheme is needed to be reconciled with the highly connected structure of intracellular signal transduction pathways.

2. The systems biology approach

In the last decade, many mathematical models have been formulated to investigate the behavior of complex intracellular biochemical networks. 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.

Surprisingly, the mathematical formulation of these highly interconnected enzyme reactions is usually based on *in vitro* studies of isolated reactions, without 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 [5].

This fact 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 several "virtual cell" projects, which aim at being both a qualitative and a quantitative precise representation of the living cell, the use of correct parameters, kinetic schemes and initial conditions i.e., steady-state concentrations of molecular species become crucial. The mathematical efforts in this direction are still largely unsatisfactory, due to the complexity and variety of the involved phenomena. Most of the mathematical models reproducing the signal transduction pathways neglect the spatial and stochastic features of the cell, which is seen as a homogeneous medium in which a lot of randomly interacting molecules are present.

In many cases, the spatial distribution of molecules is important (for example, some proteins can be activated and interact only on the plasma membrane), while some phenomena (like the gene expression and protein production) are intrinsically stochastic events, in which few molecules are engaged at a time. After the successes obtained by means of the mathematical models formulated in terms of ordinary differential equations (ODEs), there is the need of extending these models to the use of partial differential equations (PDEs, in which the nonhomogeneous spatial structure of the cell is explicitly introduced and the subcellular compartments are treated as quantitatively different subcellular regions) and stochastic differential equations (SDEs) to have a more complete theoretical framework and reliable predictions. In this way, the pharmaceutical and basic research would improve the lead compound discovery and speed up the preclinical phase of drug development, dramatically reducing the costs and accelerating the commercialization of drugs, with enormous benefits for public health and the health national programs.

3. The deterministic approach—theoretical background

One of the principal components of the mathematical approach to Systems Biology is the model of biochemical reactions set forth by Henri in 1901 [6–8] and Michaelis and Menten in 1913 [9], and further developed by Briggs and Haldane in 1925 [10]. 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

$$E + S \underset{}{\overset{d}{\longleftarrow\!\!\!\!\longrightarrow}}{}^{a} C \xrightarrow{k} E + P, \tag{1}$$

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

The fundamental step is modeling all of the intermediate reactions, including binding, dissociation and release of the product using mass action and conservation laws. This leads to an ODE for each involved complex and substrate. We refer to this as the full system. For (1) the equations are

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

$$\frac{\mathrm{d}C}{\mathrm{d}t} = a(E_T - C)S - (d + k)C \tag{3}$$

with the initial conditions

$$S(0) = S_T, \qquad C(0) = 0, \tag{4}$$

and the conservation laws

$$E + C = E_T, \qquad S + C + P = S_T. \tag{5}$$

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. 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 so-called Michaelis–Menten (MM) kinetics [9,11,12]. Let us observe that the system (2) and (3) admits an asymptotic solution for $t \to \infty$, obtained by setting the derivatives equal to zero. This solution is given by $C = S = 0$, so that from the conservation laws $P = S_T$ and $E = E_T$. This means that all the substrate eventually becomes product due to the irreversibility, while the enzyme eventually is free and the complex concentration tends to zero.

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 approximation* (standard QSSA, sQSSA). It leads to an ODE for each substrate while the complexes are assumed to be in a quasi-steady-state (i.e., $\frac{dC}{dt} \approx 0$):

$$\frac{dS}{dt} \approx -kC \approx -\frac{V_{\max} S}{K_M + S}, \qquad S(0) = S_T, \tag{6}$$

$$V_{\max} = kE_T, \qquad K_M = \frac{d+k}{a}. \tag{7}$$

The advantage of a quasi-steady-state approximation is that it reduces the dimensionality 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. See e.g. [12] for a general introduction to this approach. We stress here that this is an approximation to the full system, and that it is only valid when the enzyme concentration is much lower than either the substrate concentration or the Michaelis constant K_M, i.e., $E_T \ll S_T + K_M$ [11,13]. This condition is usually fulfilled for *in vitro* experiments, but often breaks down *in vivo* [14,15]. We refer to [16] for a nice, general review of the kinetics and approximations of (1).

However, as mentioned above, to simulate physiologically realistic *in vivo* scenarios, one faces the problem that the MM approximation is no longer valid. Hence, even though 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.

4. The total quasi-steady-state approximation

As mentioned in the previous section, *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 total quasi-steady-state assumption (tQSSA).

It was introduced by Borghans et al. [17] and refined by Tzafriri [18] for isolated reactions. It arises by introducing the total substrate

$$\overline{S} = S + C, \tag{8}$$

and assuming that the complex is in a quasi-steady-state as for the sQSSA. For (1) it gives [18]

$$\frac{d\overline{S}}{dt} \approx -kC_-(\overline{S}), \qquad \overline{S}(0) = S_T, \tag{9}$$

where

$$C_-(\overline{S}) = \frac{(E_T + K_M + \overline{S}) - \sqrt{(E_T + K_M + \overline{S})^2 - 4E_T\overline{S}}}{2}. \tag{10}$$

Numerical integration of (9) easily gives the time behavior of $\overline{S}$, C (by (10)) and S (by the relation $S = \overline{S} - C$).

Tzafriri [18] showed that the tQSSA (9) is valid whenever

$$\epsilon_{Tz} := \frac{K}{2S_T}\left(\frac{E_T + K_M + S_T}{\sqrt{(E_T + K_M + S_T)^2 - 4E_T S_T}} - 1\right) \ll 1, \quad \text{where } K = \frac{k}{a}, \tag{11}$$

and that this is at least roughly valid for any set of parameters, in the sense that

$$\epsilon_{Tz} \le \frac{K}{4K_M} \le \frac{1}{4}. \tag{12}$$

This means that, for *any* combination of parameters and initial conditions, (9) is a decent approximation to the full system (2) and (3). The parameter K is known as the Van Slyke–Cullen constant.

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, the description of cooperative reactions, inhibition and many other biochemical processes have up to now exploited the fundamental ideas of the MM scheme, i.e., the sQSSA and the parameters $V_{\max}$ and K_M (see, e.g., [12]). 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 [19] studied the completely reversible enzyme reaction in terms of the tQSSA. In [20] the tQSSA was derived for fully competitive reactions.

As a first-order approximation to (9) Tzafriri [18] found the expression, obtained originally in [17] by different techniques,

$$\frac{\mathrm{d}\overline{S}}{\mathrm{d}t} \approx -\frac{V_{\max}\overline{S}}{K_M + E_T + \overline{S}}, \qquad \overline{S}(0) = S_T. \tag{13}$$

This approximation is valid at low enzyme concentrations $E_T \ll S_T + K_M$, where it reduces to the MM expression (6), but holds moreover at low substrate concentrations $S_T \ll E_T + K_M$ [18]. Thus, with minimal effort in performing the substitutions of S by $\overline{S}$ and of K_M by $K_M + E_T$ one obtains a significantly improved MM-like approximation, without any need for more advanced mathematics.

5. The total quasi-steady-state approximation for complex reactions

The paper [20] investigated the system

$$\begin{aligned} S_1 + E &\underset{d_1}{\overset{a_1}{\rightleftarrows}} C_1 \xrightarrow{k_1} E + P_1, \\ S_2 + E &\underset{d_2}{\overset{a_2}{\rightleftarrows}} C_2 \xrightarrow{k_2} E + P_2, \end{aligned} \tag{14}$$

which consists of two reactions catalyzed by the same enzyme, i.e., a system with competing substrates, which is governed by the coupled ODEs [11,21,22], $i = 1, 2$,

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = -a_i E \cdot S_i + d_i C_i, \qquad S_i(0) = S_{i,T}, \tag{15}$$

$$\frac{\mathrm{d}C_i}{\mathrm{d}t} = a_i(E \cdot S_i - K_i C_i), \qquad C_i(0) = 0, \qquad K_i = \frac{d_i + k_i}{a_i} \tag{16}$$

and the conservation laws

$$S_{i,T} = S_i + C_i + P_i, \quad i = 1, 2, \qquad E_T = E + C_1 + C_2. \tag{17}$$

The sQSSA of this system is [11,21]

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} \approx -\frac{k_i E_T S_i}{K_i(1 + S_j/K_j) + S_i}, \qquad S_i(0) = S_{i,T}, \quad i = 1, 2,\ j \neq i, \tag{18}$$

which is valid when [22]

$$\frac{E_T}{K_i(1 + S_{j,T}/K_j) + S_{i,T}} \ll 1, \quad i = 1, 2,\ j \neq i. \tag{19}$$

As in the noncompetitive case, it says that the sQSSA holds at low enzyme concentrations.

These results were improved in [20], applying the tQSSA to both reactions and showing that the tQSSA is given by finding C_1 as the unique biologically acceptable root ($0 < C_1 < \min\{E_T, \overline{S}_1\}$) of the third degree polynomial

$$\begin{aligned} \psi_1(C_1) = &-(K_1 - K_2)C_1^3 + [(E_T + K_1 + \overline{S}_1)(K_1 - K_2) - (\overline{S}_1 K_2 + \overline{S}_2 K_1)]C_1^2 + [-E_T(K_1 - K_2) \\ &+ (\overline{S}_1 K_2 + \overline{S}_2 K_1) + K_2(E_T + K_1)]\overline{S}_1 C_1 - E_T K_2(\overline{S}_1)^2 \end{aligned} \tag{20}$$

and similarly finding C_2 as the root in the polynomial ψ_2 obtained by interchanging the indices 1 and 2 in (20). After a short transient phase the complex concentrations are assumed to equal the quasi-steady-state concentrations, $C_i = C_i(\overline{S}_1, \overline{S}_2)$, given by the roots in the respective polynomials as discussed above. Then the evolution of the system can be studied by means of the tQSSA

$$\frac{\mathrm{d}\overline{S}_i}{\mathrm{d}t} \approx -k_i C_i(\overline{S}_1, \overline{S}_2), \qquad \overline{S}_i(0) = S_{i,T}. \tag{21}$$

This approach extends both the sQSSA for competitive reactions (18) as well as the tQSSA for isolated reactions (9), as shown in [20].

Let us remark that in the classical literature every single reaction of a network is often treated by an MM approximation for an isolated reaction of the form (1), not only without any *a priori* examination of its applicability, but also neglecting the complexes and the other terms involved in double reactions or in competitions. This means that even when the sQSSA holds for every noncompetitive reaction in (14), the neglect of the complexes and of the competition leads to wrong estimations of the behavior.

To show this fact, let us consider two important reactions: the Goldbeter–Koshland switch and the double phosphorylation mechanism.

In [23] the tQSSA was applied to the so-called Goldbeter–Koshland switch [24]

$$\begin{aligned} &S + E_1 \underset{a_1}{\overset{d_1}{\longleftrightarrow}} C_1 \xrightarrow{k_1} E_1 + P_1, \\ &P + E_2 \underset{a_2}{\overset{d_2}{\longleftrightarrow}} C_2 \xrightarrow{k_2} E_2 + S, \end{aligned} \tag{22}$$

which describes, for example, the cycle of phosphorylation and dephosphorylation of a substrate, by means of a kinase E_1 and a phosphatase E_2. This reaction is very important in every intracellular pathway, because the process of phosphorylation and dephosphorylation is one of the most important process in activating and inactivating enzymes.

The reaction is governed by the coupled ODEs

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= -a_1 E_1 \cdot S + d_1 C_1 + k_2 C_2, \qquad S(0) = S_T, \\ \frac{\mathrm{d}C_1}{\mathrm{d}t} &= a_1 E_1 \cdot S - (d_1 + k_1) C_1, \\ \frac{\mathrm{d}C_2}{\mathrm{d}t} &= a_2 E_2 \cdot P - (d_2 + k_2) C_2, \qquad C_i(0) = 0 \end{aligned} \tag{23}$$

and the conservation laws

$$S_T = S + C_1 + C_2 + P, \qquad E_{i,T} = E_i + C_i, \quad i = 1, 2. \tag{24}$$

Introducing the total substrates $\overline{S} = S + C_1$, $\overline{P} = P + C_2$, we rewrite the Eq. (23) in the following way:

$$\begin{aligned} \frac{\mathrm{d}\overline{S}}{\mathrm{d}t} &= k_2 C_2 - k_1 C_1 = -\frac{\mathrm{d}\overline{P}}{\mathrm{d}t}, \qquad \overline{S}(0) = S_T \\ \frac{\mathrm{d}C_1}{\mathrm{d}t} &= a_1 (E_{1,T} - C_1) \cdot (\overline{S} - C_1) - (d_1 + k_1) C_1, \\ \frac{\mathrm{d}C_2}{\mathrm{d}t} &= a_2 (S_T - \overline{S} - C_2) \cdot (E_{2,T} - C_2) - (d_2 + k_2) C_2, \qquad C_i(0) = 0. \end{aligned} \tag{25}$$

Assuming the tQSSA $\frac{\mathrm{d}C_i}{\mathrm{d}t} \approx 0$ and considering only the biologically significant roots C_i, we arrive at the following equation

$$\frac{\mathrm{d}\overline{S}}{\mathrm{d}t} \approx k_2 C_2^- - k_1 C_1^-, \qquad \overline{S}(0) = S_T \tag{26}$$

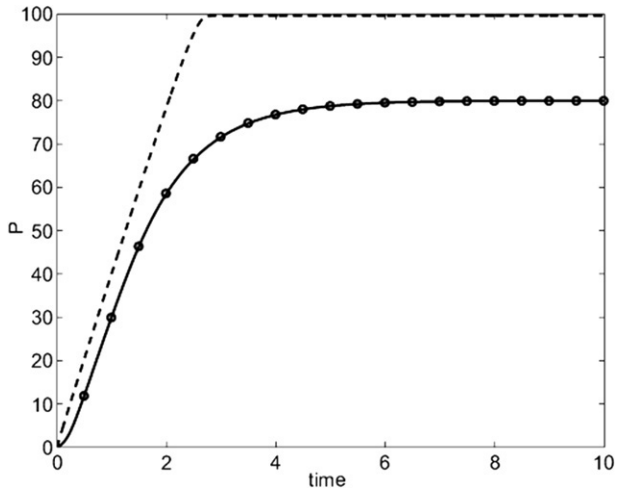

Fig. 1. The tQSSA estimates well the development of the product P of the Goldbeter–Koshland switch (scheme (22) [23]). Legends are: Full system (circles), sQSSA (dashed line) and tQSSA (full curve), where we show $P_{\text{tQSSA}} := \overline{P} - C_2^-$ for a correct comparison between sQSSA and tQSSA. The parameters are $S_T = 100,\ E_{1,T} = 500,\ E_{2,T} = 10,\ a_1 = a_2 = 4,\ d_1 = d_2 = 5,\ k_1 = k_2 = 1$, all in arbitrary units.

where

$$C_1^- = \frac{(\overline{S} + E_{1,T} + K_1) - \sqrt{(\overline{S} + E_{1,T} + K_1)^2 - 4\overline{S}E_{1,T}}}{2}, \tag{27}$$

$$C_2^- = \frac{(S_T - \overline{S} + E_{2,T} + K_2) - \sqrt{(S_T - \overline{S} + E_{2,T} + K_2)^2 - 4(S_T - \overline{S})E_{2,T}}}{2} \tag{28}$$

and $K_i = \frac{d_i + k_i}{a_i}$.

Formulas (27) and (28) show that, differently from the case of a single phosphorylation reaction, in this situation the quasi-steady-state does not imply that the complexes tend to be negligible. In Fig. 1 [23] the simulations of the full system and of its QSSAs are compared.

The double phosphorylation as well as double dephosphorylation was recently modeled in the context of the so-called mitogen-activated protein kinase cascade (MAPK cascade) (see, for example, [25–29]).

As far as we know, differently from the previous examples, there is no explicit formula for the tQSSA of the double phosphorylation mechanism

$$\begin{aligned} &S + E \underset{a_1}{\overset{d_1}{\longleftarrow\!\!\longrightarrow}} C_1 \xrightarrow{k_1} E + S_p, \\ &S_p + E \underset{a_2}{\overset{d_2}{\longleftarrow\!\!\longrightarrow}} C_2 \xrightarrow{k_2} E + S_{pp}, \end{aligned} \tag{29}$$

where the activating enzyme E is a kinase, which phosphorylates the substrate S. The phosphorylated molecule S_p can bind the same enzymes, producing the double phosphorylated form S_{pp}. The reaction scheme can be seen as a special case of (14) with $S_1 = S$, $P_1 = S_2$ and $P_2 = S_{pp}$, because S and S_p compete for the same enzyme, E, but differently from the fully competitive reaction, it is usually assumed that at the beginning only S is present.

Consequently in [20] the tQSSA for competitive reactions was applied to this case, too, though it should be remarked that the theoretical investigation of the validity of the tQSSA does not work in the case of successive reactions. The problem is that there is no S_2 at time $t = 0$, and hence the time scales cannot be found following [11] because the definition of the transient phase no longer holds.

Nevertheless, it seems that the conclusions concerning the validity of the tQSSA from [20] carry over to this scenario (see the three panels of Fig. 2).

As remarked above, the double phosphorylation–dephosphorylation mechanism is central in many reaction networks and, in particular, in the MAPK cascade. This is a ubiquitous cascade, present in many cells and in many

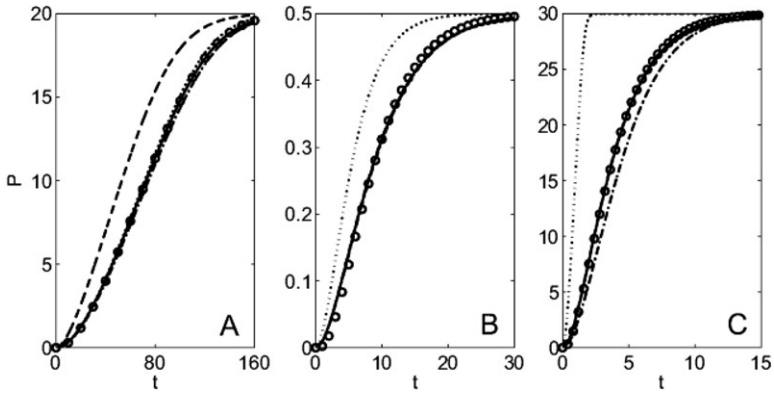

Fig. 2. A simulation of two successive reactions catalyzed by the same enzyme (scheme (29) [20]). In the three cases under investigation the full system (circles) is estimated very well by the competitive tQSSA (full curve), even when its first-order approximation (dash–dot curve), the competitive sQSSA (dotted curve) as well as the noncompetitive tQSSA (dashed curve) do not fit (panel C). Parameters are $a_1 = a_2 = 0.2$; $d_1 = d_2 = 1$; $k_1 = 0.6$; $k_2 = 0.5$ ($K_1 = 8$; $K_2 = 7.5$; $K = 3$). $S_p(0) = S_{pp}(0) = 0$. In A: $S(0) = S_T = 20$. In B: $S(0) = S_T = 0.5$. In C: $S(0) = S_T = 30$. All units are arbitrary.

transduction pathways, and is one of the most important and most studied subnetworks, not only by a biochemical point of view, but also by means of mathematical models (see, for example, [30–33]).

In [34] a comparison between the application of the sQSSA and of the tQSSA to the MAPK cascade computational model has been performed. Also in this case the double phosphorylation–dephosphorylation loops were approximated by the tQSSA for fully competitive reactions. The tQSSA was shown to be much more effective, though better approximations are needed for similar pathways, characterized by high complexity.

6. Numerical results

In the previous section we have shown that the use of the sQSSA can lead to gross quantitatively as well as qualitatively wrong conclusions even in the case of simple networks. Simulations show that the tQSSA can estimate the behavior of the systems under investigation significantly better, and therefore we propose to use this approximation when modeling intracellular signalling networks.

Fig. 1 (see [23]) and Fig. 2 (see [34]) show various approximation schemes of the time concentration development of the chemical species involved respectively in the Goldbeter–Koshland switch and in the double phosphorylation. This was done by numerically solving the systems of ordinary differential equations (ODEs) derived from the reaction schemes (22) and (29), using the methods and approximations within the various approaches.

As shown in Fig. 1, neglecting the complexes in the sQSSA brings about not only qualitatively, but also quantitatively bad approximations of the temporal behavior of the concentrations. On the other hand, the tQSSA fits in a very satisfactory way the solution of the full system.

Fig. 2 shows that, though without any theoretical investigation of its validity for the scheme (29), applying the tQSSA for fully competitive reactions to this scheme brings in much better approximations than the sQSSA.

7. The stochastic approach—problems and perspectives

The small dimensions of cells and proteins imply the fact that enzymes inside the cell are subject to random fluctuations due to Brownian motion. The randomness of enzyme displacements affects their interaction rates, mainly when the number of proteins is very low and any description of the reactions in terms of continuous fluxes of matter is expected not to hold anymore.

In this case the deterministic differential equations must be substituted by discrete, stochastic equations.

Starting from the pioneering works by Delbrück [35] and Kramers [36] for general chemical reactions and Bartholomay [37–39] for the Michaelis–Menten kinetics, much literature has been produced on this topic. Let us, for instance, mention the paper [40] by Qian and Elson, where a single molecule kinetics is studied in a stochastic framework.

Let us consider $n \geq 1$ different molecules $S_1, S_2, \ldots, S_n$, interacting through $m \geq 1$ reactions $R_1, R_2, \ldots, R_m$ and introduce the vector $\mathbf{X}(t)$, describing the dynamical state of the system, whose components $X_i(t)$ represent the number of S_i molecules in the system at time t.

We can consider the probability $a_j(\mathbf{x})\mathrm{d}t$ given $\mathbf{X}(t) = \mathbf{x}$, that one R_j reaction will occur in the next infinitesimal time interval $[t, t + \mathrm{d}t)$ and the *state-change vector* $\boldsymbol{\nu}_\mathbf{j}$, whose ith component ν_{ji} is defined by the change in number of S_i molecules produced by one R_j reaction.

The time evolution for the probability $\mathbf{P}(\mathbf{x}, t|\mathbf{x_0}, t_0)$ that, for $t \geq t_0$, $\mathbf{X}(t)$ will equal $\mathbf{x}$ given that $\mathbf{X}(t_0) = \mathbf{x_0}$ is described by the so-called *chemical master equation* (CME) [41]:

$$\frac{\partial}{\partial t}\mathbf{P}(\mathbf{x}, t \mid \mathbf{x_0}, t_0) = \sum_{j=1}^{m}[a_j(\mathbf{x} - \nu_\mathbf{j})\mathbf{P}(\mathbf{x} - \boldsymbol{\nu}_\mathbf{j}, t \mid \mathbf{x_0}, t_0) - a_j(\mathbf{x})\mathbf{P}(\mathbf{x}, t|\mathbf{x_0}, t_0)]. \tag{30}$$

The functions a_j are known as *propensity functions.*

When we deal with enzyme kinetics and consider the probability $\mathbf{P} := \mathbf{P}(E, S, C, P, t)$ that there will be present, at time t, E molecules of the activating enzyme, S molecules of the substrate, P molecules of the product and C molecules of the complex, the CME assumes the following form

$$\begin{aligned}\frac{\partial \mathbf{P}}{\partial t} &= a(S+1)(E+1)\mathbf{P}(S+1, E+1, C-1, P, t) + d(C+1)\mathbf{P}(S-1, E-1, C+1, P, t)\\ &\quad + k(C+1)\mathbf{P}(S, E-1, C+1, P-1, t) - [aS \cdot E + kC + dC]\mathbf{P}(S, E, C, P, t),\end{aligned} \tag{31}$$

where the constant rates a, d, k of the reaction play the role of propensity functions.

Since in general it is not possible to analytically or numerically solve the CME, stochastic simulation algorithms (SSAs) [42] were formulated and further improved and/or adapted to different scenarios (see [43,44] for clear and complete reviews).

Many of these algorithms use the sQSSA, which has been shown to be inadequate in the deterministic approach even in the simplest of reactions. It seems much more appropriate to apply the tQSSA to any SSA. This is one of the main goals of our future research.

The stochastic nature of the enzymatic reactions leads to many other, interesting challenges for the future researches.

As remarked above, the subnetworks we are studying, like the MAP kinase cascade, can be characterized, in particular, by multistability, hysteresis, zero–one responses etc. The crucial role of multistationarity and ultrasensitivity has been remarked upon by many authors (see, for example, [30–32,45,33]).

The introduction of even small stochastic fluctuations could bring about dramatic changes, like the passage of the system from one steady-state to another one. Considering, for instance, the localization property of many reactions in a network, we could interpret the concentration fluctuations as the changes caused by the random movements of the enzymes, which enter in and exit from the cell compartment where the reaction is localized, changing the probabilities of interactions.

On the other hand, since the constant rates are never really constant in time, depending on many factors, like, for example, the molecular crowding (see, for example, [46,47,5,48,49]), we want in the future to study the effects of even small fluctuations of the constant rates on the enzyme concentrations. In general, we may expect that, in the presence of multistable systems, small stochastic perturbations can even bring about drastic changes in the cell behavior.

Finally, some particular subnetworks, like the above quoted double phosphorylation–dephosphorylation loop or some more complex feedback loops, like in the MAPK cascade, can show the presence of oscillations of some enzyme concentrations. Under suitable conditions, which must still be investigated, we could expect some extremal phenomena, like the so-called stochastic resonance (see [50] for a complete review), where optimal small stochastic perturbations can bring about a dramatic enhancement of the oscillation amplitude.

References

[1] B. Booth, R. Zemmel, Prospects for productivity, Nat. Rev. Drug Discov. 3 (2004) 451–456.
[2] J. DiMasi, R.W. Hansen, H.G. Grabowski, The price of innovation: New estimates of drug development costs, J. Health Econ. 22 (2003) 151–185.
[3] U.S. Bhalla, Models of cell signaling pathways, Curr. Opin. Genet. Dev. 14 (2004) 375–381.

[4] R.J. Orton, O.E. Sturm, V. Vyshemirsky, M. Calder, D.R. Gilbert, W. Kolch, Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway, Biochem. J. 392 (2005) 249–261.
[5] R.J. Ellis, Macromolecular crowding: Obvious but underappreciated, Trends Biochem. Sci. 26 (2001) 597–604.
[6] V. Henri, Recherches sur la loi de l'action de la sucrase, C. R. Hebd. Acad. Sci. 133 (1901) 891–899.
[7] V. Henri, Über das Gesetz der Wirkung des Invertins, Z. Phys. Chem. 39 (1901) 194–216.
[8] V. Henri, Théorie générale de l'action de quelques diastases, C. R. Hebd. Acad. Sci. 135 (1902) 916–919.
[9] L. Michaelis, M.L. Menten, Die Kinetik der Invertinwirkung, Biochem. Z. 49 (1913) 333–369.
[10] G.E. Briggs, J.B.S. Haldane, A note on the kinetics of enzyme action, Biochem. J. 19 (1925) 338–339.
[11] L. Segel, On the validity of the steady state assumption of enzyme kinetics, Bull. Math. Biol. 50 (1988) 579–593.
[12] H. Bisswanger, Enzyme Kinetics: Principles and Methods, Wiley-VCH, Weinheim, 2002.
[13] L.A. Segel, M. Slemrod, The quasi steady-state assumption: A case study in perturbation, SIAM Rev. 31 (1989) 446–477.
[14] O.H. Straus, A. Goldstein, Zone behavior of enzymes, J. Genet. Physiol. 26 (1943) 559–585.
[15] A. Sols, R. Marco, Concentration of metabolites and binding sites. Implications in metabolic regulation, in: Current topics in Cellular Regulation, Academic Press, New York, 1970.
[16] S. Schnell, P.K. Maini, A century of enzyme kinetics. Reliability of the KM and vmax estimates, Comment. Theor. Biol. 8 (2003) 169–187.
[17] J. Borghans, R. de Boer, L. Segel, Extending the quasi-steady state approximation by changing variables, Bull. Math. Biol. 58 (1996) 43–63.
[18] A.R. Tzafriri, Michaelis–Menten kinetics at high enzyme concentrations, Bull. Math. Biol. 65 (2003) 1111–1129.
[19] A.R. Tzafriri, E.R. Edelman, The total quasi-steady-state approximation is valid for reversible enzyme kinetics, J. Theor. Biol. 226 (2004) 303–313.
[20] M.G. Pedersen, A.M. Bersani, E. Bersani, The quasi steady-state approximation for fully competitive enzyme reactions, Bull. Math. Biol. 69 (2007) 433–457.
[21] S.I. Rubinow, J.L. Lebowitz, Time-dependent Michaelis–Menten kinetics for an enzyme-substrate-inhibitor system, J. Amer. Chem. Soc. 92 (1970) 3888–3893.
[22] S. Schnell, C. Mendoza, Time-dependent closed form solutions for fully competitive enzyme reactions, Bull. Math. Biol. 62 (2000) 321–336.
[23] M.G. Pedersen, A.M. Bersani, E. Bersani, G. Cortese, The total quasi-steady-state approximation for complex enzyme reactions, in: I. Troch, F. Breitenecker (Eds.), Proceedings IV MATHMOD Vienna, in: ARGESIM Report n. 30, vol. 2, Vienna University of Technology, Vienna, 2006.
[24] A. Goldbeter, D.E. Koshland Jr., An amplified sensitivity arising from covalent modification in biological systems, Proc. Natl. Acad. Sci. 78 (1981) 6840–6844.
[25] W.R. Burack, T.W. Sturgill, The activating dual phosphorylation of MAPK by MEK is nonprocessive, Biochemistry 36 (1997) 5929–5933.
[26] J.E. Ferrell Jr., R.R. Bhatt, Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase, J. Biol. Chem. 272 (1997) 19008–19016.
[27] Y. Zhao, Z.-Y. Zhang, The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinase phosphatase 3, J. Biol. Chem. 276 (2001) 32382–32391.
[28] M. Hatakeyama, S. Kimura, T. Naka, T. Kawasaki, N. Yumoto, M. Ichikawa, J.H. Kim, K. Saito, M. Saeki, M. Shirouzu, S. Yokoyama, A. Konagaya, A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling, Biochem. J. 373 (2003) 451–463.
[29] N.I. Markevich, J.B. Hoek, B.N. Kholodenko, Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades, J. Cell Biol. 164 (2004) 353–359.
[30] C.-Y.F. Huang, J.E. Ferrell Jr., Ultrasensitivity in the mitogen-activated protein kinase cascade, Proc. Natl. Acad. Sci. 93 (1996) 10078–10083.
[31] U.S. Bhalla, R. Iyengar, Emergent properties of networks of biological signaling pathways, Science 283 (1999) 381–387.
[32] B.N. Kholodenko, Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades, Eur. J. Biochem. 267 (2000) 1583–1588.
[33] X. Wang, N. Hao, H.G. Dohlman, T.C. Elston, Bistability, stochasticity, and oscillations in the mitogen-activated protein kinase cascade, Biophys. J. 90 (2006) 1961–1978.
[34] M.G. Pedersen, A.M. Bersani, E. Bersani, Quasi steady-state approximations in intracellular signal transduction — a word of caution, J. Math. Chem. preprint Me.Mo.Mat. n. 6/2006 (accepted for publication).
[35] M. Delbrück, Statistical fluctuations in autocatalytic reactions, J. Chem. Phys. 8 (1940) 120–124.
[36] H.A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica 7 (1940) 284–304.
[37] A. Bartholomay, A stochastic approach to chemical reaction kinetics, Ph.D. Thesis, Harvard University, 1957.
[38] A. Bartholomay, Enzyme reaction rate theory: A stochastic approach, Ann. N.Y. Acad. Sci. 96 (1962) 897–912.
[39] A. Bartholomay, A stochastic approach to statistical kinetics with application to enzyme kinetics, Biochemistry 1 (1962) 223–230.
[40] H. Qian, E.L. Elson, Single-molecule enzymology: Stochastic Michaelis–Menten kinetics, Biophys. Chem. 101–102 (2002) 565–576.
[41] D.A. McQuarrie, Stochastic approach to chemical kinetics, J. Appl. Probab. 4 (1967) 413–478.
[42] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem. 81 (1977) 2340–2361.
[43] T.E. Turner, S. Schnell, K. Burrage, Stochastic approaches for modelling in vivo reactions, Comput. Biol. Chem. 28 (2004) 165–178.
[44] J. Puchałka, A.M. Kierzek, Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks, Biophys. J. 86 (2004) 1357–1372.
[45] R. Thomas, M. Kaufman, Multistationarity, the basis of cell differentiation and memory. I. Structural conditions of multistationarity and other nontrivial behavior, Chaos 11 (2001) 170–179.
[46] R. Kopelman, Rate-processes on fractals: Theory, simulations, and experiments, J. Stat. Phys. 42 (1986) 185–200.
[47] R. Kopelman, Fractal reaction kinetics, Science 241 (1988) 1620–1626.
[48] S. Schnell, T.E. Turner, Reaction kinetics in intracellular environments with macromolecular crowding: Simulations and rate laws, Prog. Biophys. Mol. Biol. 85 (2004) 235–260.
[49] D.V. Nicolau Jr., K. Burrage, Stochastic simulation of chemical reactions in spatially complex media, Comput. Math. Appl. 55 (5) (2008) 1007–1018.
[50] L. Gammaitoni, P. Hanggi, P. Jung, F. Marchesoni, Stochastic resonance, Rev. Modern Phys. 70 (1998) 315–322.